

5-2018

# Prisons, Genres, and Big Data: Understanding the Language of Corrections in America's Prisons

Eric Stephens

Clemson University, stephens.ej@gmail.com

Follow this and additional works at: [https://tigerprints.clemson.edu/all\\_dissertations](https://tigerprints.clemson.edu/all_dissertations)

---

## Recommended Citation

Stephens, Eric, "Prisons, Genres, and Big Data: Understanding the Language of Corrections in America's Prisons" (2018). *All Dissertations*. 2103.

[https://tigerprints.clemson.edu/all\\_dissertations/2103](https://tigerprints.clemson.edu/all_dissertations/2103)

This Dissertation is brought to you for free and open access by the Dissertations at TigerPrints. It has been accepted for inclusion in All Dissertations by an authorized administrator of TigerPrints. For more information, please contact [kokeefe@clemson.edu](mailto:kokeefe@clemson.edu).

PRISONS, GENRES, AND BIG DATA: UNDERSTANDING THE LANGUAGE OF  
CORRECTIONS IN AMERICA'S PRISONS

---

A Dissertation  
Presented to  
the Graduate School of  
Clemson University

---

In Partial Fulfillment  
of the Requirements for the Degree  
Doctor of Philosophy  
Rhetorics, Communication, and Information Design

---

by  
Eric James Stephens  
April 2018

---

Accepted by:  
Dr. Victor Vitanza, Committee Chair  
Dr. David Blakesley  
Dr. Michael Meng  
Dr. Alexander Herzog

## ABSTRACT

This dissertation seeks to answer one fundamental question: How can I as a researcher conduct social justice research that is ethical, durable, and portable? As social justice research becomes more prominent in the field of technical and professional communication, ethical research practices must be maintained to avoid an unintentional wounding of the subjects for whom researchers hope to advocate. The dissertation is divided into five sections, each written as a stand-alone article that builds on the principles of the section before it. Each section addresses a key question:

- 1) How do I ethically engage in social justice research?
- 2) How do I ethically engage with big data and algorithmic rhetorics?
- 3) How do I frame my research to have the most impact outside my home discipline?
- 4) What does an ethical, computational content analysis look like?
- 5) How do these principles translate into the classroom?

Together, these articles identify a methodology called Institutional Genre Analysis, which focuses on text as data that was produced by an institution rather than individual users, avoiding many of the pitfalls of big data research while providing a means for what Vitanza calls “intellectual guerilla warfare conducted by [marginalized individuals]” (1987, p. 52).

## TABLE OF CONTENTS

	Page
TITLE PAGE .....	i
ABSTRACT .....	ii
LIST OF TABLES .....	v
LIST OF FIGURES .....	vi
ARTICLES	
1. SOCIAL JUSTICE AND <i>THE BELMONT REPORT</i> : A WICKED PROBLEM	
Introduction.....	1
Justice.....	5
Identity .....	13
Research .....	16
Concluding Thoughts.....	25
References.....	28
2. ALGORITHMIC AURAS AND INSTITUTIONAL GENRE ANALYSIS: A CASE STUDY FOR WORKING ETHICALLY WITH BIG DATA	
Introduction.....	31
Replicating the Unrepeatable, Unique Individual.....	33
Big Data Methodologies as Ideologies .....	39
Establishing Best Practices for Big Data Methodologies .....	46
Conclusion .....	53
References.....	55
3. RHETORICAL CONVERSATIONS AND RADICAL COLLABORATIONS: BUILDING PORTABLE, DURABLE RESEARCH IN TECHNICAL AND PROFESSIONAL COMMUNICATION	
Introduction.....	59
Defining Durable, Portable Research as Rhetorical Conversations.....	61
Defining Radical Collaboration .....	65
TPC in Prisons: An Example of Radical Collaboration in Action.....	74
Concluding Thoughts.....	82
References.....	84

Table of Contents (Continued)	Page
4. FROM AN ETHICS OF CARE TO EXPEDIENCY IN AMERICA’S PRISONS: USING INSTITUTIONAL GENRE ANALYSIS TO UNDERSTAND THE LANGUAGE OF CORRECTIONS IN ADULT INMATE HANDBOOKS	
Introduction.....	87
Methods.....	98
Results.....	105
Analysis and Discussion .....	112
Conclusion .....	114
References.....	115
5. THE WICKED AND THE MUNDANE: AN ALTERNATIVE PROJECT-BASED COURSE FOR THE MULTI-MAJOR TECHNICAL AND PROFESSIONAL COMMUNICAITON CLASSROOM	
Introduction.....	118
Wicked Problems and Mundane Artifacts .....	120
Frameworks for TPC Pedagogy.....	122
PESTEL and Design Thinking.....	128
The Course Design.....	134
Conclusion .....	141
References.....	142
APPENDIX	
A. Python Code.....	145

## LIST OF TABLES

Tables	Page
1.1 The doctrine of double-effect.....	18
2.1 Establishing best practices for big data methodologies .....	49
3.1 Three categories for content analysis .....	81
3.2 A sample of coded content .....	81
5.1 A PESTEL breakdown of an episode of <i>The Walking Dead</i> .....	136
5.2 Combining layered literacies and design thinking.....	139

## LIST OF FIGURES

Figures	Page
3.1 Machine learning generated synonyms for “discipline” in two penal studies journals .....	79
3.2 Machine learning generated synonyms for “punishment” in two penal studies journals .....	80
4.1 Screenshot of the listed locations of federally run prisons .....	101
4.2 Screenshot of USP Leavenworth’s landing page .....	102
4.3 Locations of each handbook .....	103
4.4 Map showing the sizes of unique n-grams by operational level .....	103
4.5 Breaking down “rehabilitation” in 347 inmate handbooks .....	107
4.6 Breaking down “discipline” in 347 inmate handbooks .....	108
4.7 Breaking down “work” in 347 inmate handbooks .....	109
4.8 Triangulating “rehabilitation,” “punishment,” and “work,” along with subject-verb n-grams .....	111

## ARTICLE 1

### SOCIAL JUSTICE RESEARCH AND *THE BELMONT REPORT*: A WICKED PROBLEM

“The examination and understanding of one’s own activity and consciousness, the ‘return of consciousness to its own center,’ is, as Walter Ong has suggested, the central impulse of the humanities.”

—C. Miller, “A Humanistic Rationale for Technical Writing”

“And if you ask me whether the disenfranchised can think this critique, I would say yes. It is the disenfranchised who teaches us most often by saying: I do not recognize myself in the object of your benevolence, I do not recognize my share in your naming.”

—G.C. Spivak, “Feminism and Deconstruction, Again”

“We need data, ideas, plans and strategies, but we need to see them coming from people like us, people who don’t, right now, seem to make it into your little position of power.”

—A. Tagonist, “Fuck You and Fuck Your Fucking Thesis”

#### **Introduction**

In his essay, “The ‘Q’ Question,” Richard Lanham (2010) addresses and readdresses a question originally posed by Quintilian: “Does a good orator need to be a good person?” Or to rephrase the question: Does a good rhetorician need to be a good person? Lanham goes on to describe how one should go about answering the “Q” Question by describing two defenses he calls Weak and Strong. The Weak Defense is most simplistic and relies on the context: it “argues that there are two kinds of rhetoric, good and bad. The good kind used in good causes, the bad kind in bad causes” (p. 155). Essentially, if rhetoric is used for something good, then it is good rhetoric and, therefore, a good person using it; if it’s used for something bad, then it is bad rhetoric and, therefore, a bad person using it. For obvious reasons, the Weak Defense is a poor defense because problems are often too complex; these complex problems are what Rittel & Webber (1973) call “wicked problems.” Lanham’s Strong Defense, however, is one that



“assumes that truth is determined by social dramas, some more formal than others but all man-made” (p. 156). While the Strong Defense has more appeal to myself as an academic in the humanities, I’m not convinced it’s the best defense. I wonder why he even uses the word *defense*—what is he defending exactly? In his preface to the essay, Lanham explains he’s wrestling with what the humanities are trying to protect when they want to protect books: “And defining that *center* is now an exigent task, which [he] tries to begin in this essay” (p. 154; emphasis added). If one were to turn to Adrianna Cavarero’s collective work, however, one might say that Lanham’s defenses used to seek out an intangible center has the same phallogentric intentions as Western philosophy in general—being so focused on the abstract that the “unrepeatable, unique individual” is lost.

For me, the “Q” question doesn’t seem to be a question that any one scholar can answer for another. It’s something personal. This alters how I approach the “Q” Question, breaking it down to two questions when designing and conducting research: 1) Do scholars have an ethical obligation to do social justice research? 2) How can social justice researchers do their research without further exploiting those for whom they advocate? In the fields of technical and professional communication (TPC), these questions are necessary—especially if one is inclined towards social justice research. In her essay, Hopton (2013) describes the tools that effective technical communicators have (or should have) as they “will play the roles of information developer, communicator, interpreter, and usability expert” (p. 66); technical communicators have the ability to persuade, to communicate well, and to think critically. Hopton (2013) cautions that these tools and

abilities are not to be abused (p. 67), arguing that “the technical communicator of today must be technically competent, ethically bound, critically conscious and situated with enough institutional power to halt the wheels of production when necessary, putting the common good over institutional gain” (p. 66). While her argument focuses on those working in the industry, the same principle applies to those researching in the academy. Her essay is not, however, a prescription to the industry or the academy—it’s a self-reflection. Her title repeats the question posed by the Nobel Peace Prize winner Andrei Sakharov, “If not me, then who?” For me, this is why Lanham’s “Q” Question is personal. *I am a rhetorician. I am an academic who wants to do good. How do I do that in the best possible way?* Colton, Holmes, and Walwema contend that technical communicators “must consider those who are affected by their practices, recognizing that an act of care for one may be an act of wounding another” (p. 65). ~~But what about those acts that care for one while wounding them?~~ At its core, social justice work is about other people. Other unrepeatabe, unique individuals. Social justice researchers set out to do good for and with other unrepeatabe, unique individuals to tackle some of the wicked problems in the world as they relate to TPC, but we face a glaring problem: we benefit from the research disproportionately compared to those for whom we advocate. What kind of benefits? Tenure. Salaries. Retirement. Health insurance. More research funding. Job security. Social justice research *exploits* humans for personal gain. Social justice research is itself a wicked problem.

Jones, Moore, and Walton (2016) boldly claim that as a field, we need to “unabashedly embrace” social justice research (p. 212). This article extends their call,

asking individual researchers to unabashedly embrace social justice *as a wicked problem*. Following Hopton's lead (who followed Sakharov's lead), I ask myself, "If not me, then who?" but I also seek to take this self-reflection a step further by asking *how*. How do I do research in the most ethical way possible? This is not a new question. This question prompted *The Belmont Report* (1979), which outlines and discusses the ethics of human subject research, and it has become a required reading for anyone doing research with people. Ethics, as philosophy indicates, is no simple thing, and *The Belmont Report* sought to prevent the unethical treatment of human subjects. In their introduction they admit the complicated nature of their task:

Such rules are often inadequate to cover complex situations; at times they come into conflict, and they are frequently difficult to interpret or apply. Broader ethical principles will provide a basis on which specific rules may be formulated, criticized, and interpreted. . . . The objective [of this report] is to provide an analytical framework that will guide the resolution of ethical problems arising from research involving human subjects. (p. 3)

Who we are as researchers, the work we engage with as researchers, and the people for whom we wish to advocate as a researchers all present an ethical problem when it comes to social justice research. The first formulation on which the authors of *The Belmont Report* use to define justice is "to each person an equal share." Equal share of what exactly? In discussing the systematic assessment of risks and benefits, the report proclaims, "This ideal requires those making decisions about the justifiability of research to be thorough in the accumulation and assessment of information about all aspects of the

research, and to consider alternatives systematically” (p. 8). I make decisions about the research I do. I need to consider those alternatives. Using *The Belmont Report* as an analytical framework, this article acts as a self-reflective, research-based guide for what I feel I need to do as a researcher. This article is not a prescription telling you what you should do as a social justice researcher. Instead, I offer three areas with accompanying questions designed as a heuristic to prompt self-reflection as others move forward in their own work working with other unrepeatably, unique individuals: justice, identity, and research.

## **Justice**

### **What has our field said about ethics?**

Part B of *The Belmont Report* claims there are three basic ethical principles: respect for persons, beneficence, and justice (p. 4). Historically, TPC has been concerned with process, but as it grew into a discipline scholars began asking critical questions about how they wanted to do the work they did (Connors, 1999), which brought questions of ethics into the field. In the same year *The Belmont Report* was released, Miller (1979) published her seminal article, “A Humanistic Rationale for Technical Writing,” helping to center the human in the discussion of ethics in technical communication with a simple truism: “communication occurs in communities” (p. 617). Miller discusses this centering by describing a scene from her own life. Students in her department did not want to take a literature course and wanted to take a technical writing course to fulfil their humanities requirements. Many in the English department felt technical writing courses didn’t fulfil the requirement as a humanities course. What followed was a vigorous debate in the

department about whether or not technical writing would fulfil such a requirement. Miller's argues it does: "Whatever we know of reality is created by individual action and by communal assent. Reality cannot be separated from our knowledge of it; knowledge cannot be separated from the knower; the knower cannot be separated from a community" (p. 615). The act of technical writing is to communicate ideas to an audience, to stakeholders. To do so, technical writers understand that various genres and practices are socially made with their own nuances that depend on certain community expectations. For Miller, reality cannot be separated from the community, so a technical communicator must understand the human element of communication in order to know how to best communicate. The practice of technical writing is to help other humans understand and complete a certain task; this, for Miller, is what makes it a humanistic practice—communicating *to* and *for* the Other.

While Miller establishes humans at the center of TPC, Katz (1992) extends this trajectory by asking how the technical communication we produce impacts those humans. Katz conducts a close rhetorical analysis of a technical document from Nazi Germany. "By any formal criteria in technical communication," writes Katz, "[the memo] is an almost perfect document" (p. 256). This almost perfect document, however, works toward "the mass destruction of human beings" (p. 257). For Katz, technical communication "always leads to action, and thus always impacts human life" (p. 259). With this understanding, he writes that the act of technical communication is not only an epistemological issue (as voiced by Miller) but an ethical one as well. Those decisions, those actions, those deliberations have an impact on human life—in the case of Katz's

example, the destruction of human life. Together, these two essays opened the door to more discussions on the ethics of technical communication and how our work impacts human life. More than a decade after Katz's article, Sapp, Savage, and Mattson (2013) begin the introduction to their special issue of *Rhetoric, Professional Communication and Globalization* with a bold claim: "If published work is at all a reliable indicator, the issue of human rights has not yet emerged as a consistent thread in professional communication scholarship" (p. 1). Using the International Bill of Human Rights of 1948 as their heuristic, the special issue brings the conversations of human rights and human dignity into TPC. Although humans have always been at the center of TPC work (as discussed by Miller and Katz), there is little common or consistent language used to discuss human rights and human dignity.

As the topic of ethics emerged in TPC, Salvo (2001) "explores the shift [in TPC] from observation *of* users to participation *with* users" (p. 273). This new collaborative framework raises new ethical worries for researchers that bleed into the ethical problem of this article. Salvo argues, "As usability issues become more complex, users become increasingly valuable sources of information and guidance in the design process" (p. 274). As we work with humans, Salvo contends that they become more valuable sources. In research, is that increase in value shared in equal parts? How long do participants reap the benefits of the research that may secure my tenure, providing me with job security? The trajectory continues with Walton (2016) asking an imperative question of TPC: "which humans are at the center of our work" (p. 402)? Where Walton asks "which humans," *The Belmont Report* asks, "Who ought to receive the benefits of research and

bear its burdens” (p. 5)? Walton draws upon the closely related field of human-centered design (HCD in an effort promote human dignity. The first principle of human-centered design (as described by Buchanan) is to support individuals as they act out their lives socially, politically, and culturally—it centers on human dignity (p. 409). “In other words,” Walton contends, “HCD is a design philosophy and research-based process that emphasizes the importance of having users themselves provide input to shape the design” (p. 405). According to *The Belmont Report*, human rights and dignity are encompassed by three basic ethical principles mentioned earlier: respect of persons, beneficence, and justice. The field of TPC addresses these principles. Miller establishes that people are at the center of our work, and Katz ensures that those people deserve respect as individuals. Salvo encourages researchers to make sure the work we do benefits those with whom we work. Walton challenges researchers to promote social justice, which “inextricably involves human dignity and human rights, since it is by definition concerned with the agency of oppressed people” (Walton, 2016, p. 411).

### **Whose ethics should be enacted?**

The field of ethics is as large as it is old. While this article cannot and will not delve into the various theories of ethics and justice, I extend the same invitation as Elliot (2016) for reading the canon of ethics, which he summarizes and synthesizes with the ethics of writing assessment: “Whatever the method of reading, cross-reference by sections allows discrete ideas to be examined in detail while allowing connections among them to be established.” No theory of ethics exists in a vacuum, and no theory of ethics

will resolve every issue. In his thorough discussion of justice and rationality, MacIntyre notes:

One of the most striking facts about modern political orders is that they lack institutionalized forums within which these fundamental disagreements can be systematically explored and charted, let alone there being any attempt to solve them. The facts of disagreement themselves frequently go unacknowledged, disguised by a rhetoric of consensus. (p. 2)

As noted in the previous sections, the field of TPC has recently developed discussions on ethics and human dignity, but this article does not prescribe a specific theory of ethics. Whichever ethic a researcher decides to pursue, every researcher working with human subjects ought to closely examine the literature, drawing connections and looking for what resonates. Like Lanham's (2010) "Q" Question, whose ethics should be enacted ought to be a personal decision informed by thorough investigations. After reviewing theories of ethics, I found whose ethic drives my own—Emmanuel Levinas and Adrianna Cavarero, who each grapple with our obligation to the Other.

For Levinas (1979), ethics precedes everything, even ontology (p. 43). Ethics is the first responsibility we have. In his work, *Totality and Infinity*, Levinas claims that an individual comes into being when he or she sees the face of the Other. We exist because we are called upon by the Other. Ethics come first because when we see the face of the Other, we are called upon to respond—this interpellation is what pulls us into Being. Even a non-response is a response. The ethical dilemma occurs when we think about why we need to respond, or rather "I must respond to the Other because . . ." Other



philosophies don't begin until because, after we respond. Levinas' driving question was not "Why is there so much bad in the world?" (which is significant given he was writing from a Nazi work camp), his question was "Why was there any good in the world at all?" His response, simplified, is *not* a prescription of what to do once you respond. Levinas felt that if you already had an idea of what to do or how to help the individual Other, then you were already unethical, projecting your own beliefs onto the situation. One cannot follow a single rule or a precedent in any given situation because each and every situation is unique. Like all philosophy, Levinas is bound by the work preceding him. This is manifested most plainly when he writes: "Language, which does not touch the other, even tangentially, reaches the other by calling upon [them] or by commanding [them] or by obeying [them], with all the straightforwardness of these [ethical] relations" (p. 62). While Levinas' theory of ethics centers on the face of the other to reach "the idea of infinity," that infinity extends outwards, beyond the realm of the physical because language is not physical. Adrianna Cavarero complicates this notion.

Where Levinas believes ethics begins when we *see* the face of the Other, Cavarero believes we are called into Being by *hearing* the voice of the Other. Levinas argued that language does not touch—Cavarero (2005) says it does:

The sense of hearing that is privileged here . . . nonetheless transfers the perception of uniqueness from the corporal surface, *from the face*, to the internal body. The sense of hearing, characterized as it is by organs that are internalized by highly sensitive passageways in the head, has its natural referent in a voice that also comes from internal passageways: the mouth, the throat, the network of

lungs. The play between vocal emission and acoustic perception necessarily involves the internal organs. It implicates a correspondence with the fleshy cavity that alludes to the deep body, the most bodily part of the body. The impalpability of sonorous vibrations, which is as colorless as the air, comes out of a wet mouth and arises from the red of the flesh. (p. 4, emphasis added)

This physical connection to the voice of the Other is where ethics begins for Cavarero, which she uses to shift the idea of infinity. For most of Western philosophy, just like Levinas, the idea of metaphysics and the infinite are something to seek out for “higher” meaning, which is the definition of theory according to Crowley and Hawhee (2004), “The English word *theory* derives from a Greek word (*theorein*) which literally means ‘to sit in the highest row of the arena,’ More freely translated, the term meant something like ‘to observe from afar’” (p. 57). For Cavarero, we come into being when we hear the voice of the Other and their echoing vibrations physically enter our bodies. Ethics and infinity cannot be found from afar; they must be found in close proximity to another person—another unrepeatable, unique individual. We create meaning when we hear their call, those sonorous sounds enter our physical bodies, reverberating within us, shaping us anew. That voice we hear is necessarily unique, and those sounds create a relationship before the meaning of the words are even processed (Cavarero, 2000, p. 3). With Levinas we are called in to Being when we see the face of the Other, and with Cavarero the relationship with the Other is a physical thing that happens before context is realized. For both Levinas and Cavarero, this call from the Other, whether visual or auditory, comes

from an unrepeatable and unique individual, and how we respond to the Other must also be unrepeatable and unique.

### **Am I obligated to do social justice research?**

In short, yes. According to Levinas and Cavarero, I exist *because* of the Other. I am called into Being by the Other. Therefore, I am indebted to the Other. The issue is *not* whether or not I respond to the Other by doing social justice research. All of our relationships are always-already ethical relationships. The question is what I do when I realize the always-already nature of my relationships as ethical. The question isn't "Why should I do social justice research?" The question is how can I do anything but social justice research? We are in fact called into Being by the Other, and each time we respond to an unrepeatable, unique individual, our response ought to be equally unrepeatable and unique. Each and every project should be approached personally, without regards for precedent, which is a nearly impossible task, which is why each project ought to be rethought systematically as *The Belmont Report* suggests (p. 8). Levinas and Cavarero's work doesn't merely encourage social justice research; being called forward to respond by the face and voice of the Other means that *all* research stems from the Other. All research is grounded in an ethical dilemma. Technical communication research is always-already intertwined with responding to the Other. It is always-already intertwined with the social justice, but whether our work promotes social justice or perpetuates social injustice is another question.

## **Identity**

### **Who am I? Does it matter?**

I remember sitting in a research workshop at a conference. Sitting in small desks arranged in a triangle, I listened to a fellow researcher as he deconstructed and reconstructed feminism using Deleuze and Guattari. When he finished, the female in the group asked about how he would respond to a critique that as a white male, it might not be the best idea to take such ownership of feminism. His response surprised me: “I’m not interested in talking about identity politics. Someone’s identity shouldn’t impact the theory.” For some time I felt the same, but I am also a white male. As I followed the footpaths of ethics, however, I see the inherent flaw of his thinking. As Miller (1979) noted, knowledge formation happens in a community. That community is made up of unrepeatable, unique individuals, and those individuals have identities. Philosophy and ethics are born in identity. If we are called into Being by the Other, then our existence is born from an individual. This, for Cavarero, is the real infinity—not looking up for a theory without identity, but looking backwards to the material conditions of our existence, born from Mothers, individuals born from individuals. Eternity, for Cavarero, is not in death. It’s in birth.

As an unrepeatable, unique individual born from a mother who is her own unrepeatable, unique individual, I have an identity. This identity has given me nearly every privilege, position, and power available (cf. Jones, Walton, & Moore, 2016): able-bodied, Christian-raised, cis male, educated, married, straight, upper-middle class, white academic. Much of Western philosophy has been shaped by this or a strikingly similar

demographic, and there is little need to even argue or cite this claim. But what are its implications? In his essay, “Critical Sub/Versions of the History of Philosophical Rhetoric,” Vitanza (1987) declares:

With Lyotard’s view in mind, then, I posit that it will not be “persuasion” or “identification” (“consubstantiality”) that will inform our “newer” histories of Rhetoric. . . . Instead, it will be a Rhetorical/critical attitude and practice known as “paralogism,” or what I have called “Sub/Version” (and possibly even sub-sub-versions) a kind of intellectual guerilla warfare conducted by [marginalized individuals], that will function as a de/stabilizing principle (through paradox or irony) or as a dis/placing principle (through oxymoronic metonymy) in the writing of our “newer” histories. (p. 52; emphasis added)

Vitanza calls for new rhetorics and new histories of our rhetorics, but those new histories and rhetorics will be “a kind of intellectual guerilla warfare conducted by [the marginalized].” I am not marginalized. I am the opposite of marginalized. The combination of identity markers I hold take up the entire page! As someone whose identity is what it is who wants to do “good” in the world through social justice research, I face two problems that make the wicked problem of social justice even more wicked: 1) taking up physical space for the physical voices of marginalized people to be heard, and 2) doing research setting out to do “good” but turning into a savior complex or further exploitation by another able-bodied, Christian-raised, cis male, educated, married, straight, upper-middle class, white academic.

My identity enables me to say things in certain circles where others may not be as well received, but my identity is not a pass into all circles. The likely reader of this article is someone who already has some sort of interest or commitment to social justice research. It's probably the reason you've made it to this point in the article, and it's likely that I don't need to convince you of all of this. It is imperative to acknowledge our identities as we navigate the wicked problems of the world. For the purposes of this article, I focus on my privilege, position, and power of being an academic, not unlike Spivak (1989) speaking of her teaching experience in India, "which persuaded [her] that the indigenous elite must come to terms with its unacknowledged complicity with the culture of imperialism" (p. 206). It is apparent that there are wicked problems in the world, and technical communication scholars are in a position to address them. Who addresses those problems, though, *matters*. If I (whose identity is what it is) do an in-depth study of the communication practices of marginalized women in Brazil, my study will always be seen through my eyes and my identity. Of course we can reduce these biases—that's the purpose of *The Belmont Report* and the IRB process in the first place—but we cannot eliminate it. This does not mean my hope of doing social justice research is lost. Spivak writes: "I felt I must reckon with the legacy of patriarchy which, like the culture of imperialism, is a dubious gift that we can only transform if we acknowledge it" (p. 208). I, too, must acknowledge the dubious gifts of my privilege in an effort to transform—not only myself as a researcher but the institutions that built those privileges. We cannot eliminate the impact of dominant narratives, but we can disrupt them (Jones, Walton, & Moore, 2016). In an interview with Stephen Colbert in response to a question

about what to do about privilege, DeRay Mckesson (2016) said it better than I can now: “What you can do is extend that privilege to dismantle it. . . . You can use your resources to create space for people.”

## **Research**

### **How do scholars justify the exploitation inherent in human-subject research?**

A researcher cannot view themselves as *just* an individual when it comes to research. The act of research is an act *on behalf of* or *to the benefit of* an institution: academia. In fact, it is a prerequisite for belonging in the institution of academia. In her essay, Longo (1998) describes institutions as cultural agents that impact practices and knowledge formation. Academia is such an institution! Longo asserts: “A view of culture that is limited within the walls of one organization does not allow researchers to question assumptions . . . because those practices are not placed in relationship to influences outside the organization under study” (p. 55). Just as Longo encourages researchers to view culture as part of their study, researchers ought to acknowledge the impact of our own culture on the study itself. When researchers do not see their own role within that larger institution, we run the risk of not seeing ourselves in those cultural contest for privilege, position, and power like tenure and funding. Every researcher is in a contest to be able to do more research. In doing social justice research for, and even with the Other, we are using their labor to our benefit, which is not always equally, mutually beneficial to those for whom we advocate. How do researchers justify this? Whether intentionally or unintentionally, social justice researchers employ a philosophy of war—the doctrine of double-effect.

Mangan (1949) writes: “Formulated, therefore, in its full modern dress, [the doctrine of double-effect] may be expressed as follows: A person may licitly perform an action that [they] foresee will produce a good and a bad effect provided that four conditions are verified at one and the same time” (p. 43). Table 1.1 adapts the doctrine of double-effect to an example of its use during a wartime conflict and a research study that might stem from a research area of interest to myself as a researchers—prisons.

During military conflict, there may be a target housing an enemy of the state. Those in charge have been informed that a well-placed drone strike would kill this enemy of the state. This drone strike, however, will also kill civilians. Killing the enemy of the state is the primary goal, which they believe is a good, moral goal. Killing the civilians is not the intention, and if the leaders could avoid it, they would. The targets and the civilians will be killed in the same action, and killing the target would potentially save millions of lives. The leaders order the drone strike, killing the targets and the civilians. This is the doctrine of double-effect: a morally good goal that harms innocent people. Social justice researchers face a similar dilemma. During a study, a researcher may be trying to improve understanding of the conditions of inmates in a prison. This study follows all IRB protocols, perhaps even to the point of guaranteeing no harm or risk will come to any of the inmates as research participants. By the time the study has concluded, the researcher may have published two articles and started a book project, which might lead to tenure, a better salary, and more funding. While the inmates have not been explicitly harmed, their labor has been exploited for a form of profit to the researcher. This exploitation and profit was not the goal of the researcher, and the intention was to



actually improve conditions for inmates. A researcher allows for this exploitation through the use of the doctrine of double-effect. The bad was not intended, but it was unavoidable for a better good.

Table 1.1 The doctrine of double-effect (adapted from Mangan, 1949, p. 43)		
Doctrine of Double-Effect*	Military Conflict	Academic Research
The action in itself from its very object be good or at least indifferent.	Leaders uncover the location of an enemy of the state with plans to destroy critical resources, and they decide to kill the target with a drone strike. The drone strike, however, has the high potential of killing civilians.	A researcher conducts a study of incarcerated individuals in an effort to improve prison living conditions.
The good effect and not the evil effect be intended.	Killing civilians is certainly not what the leaders intend to do--it is collateral damage.	The researcher does not <i>intend</i> to continuously exploit the unpaid labor of inmates, but this is unavoidable when doing human-subject research.
The good effect be not produced by means of the evil effect.	Killing the enemy of the state did not happen <i>because</i> of the civilian deaths.	The researcher did not choose to study inmates as human-subjects <i>in order</i> to exploit them.
There be a proportionately grave reason for permitting the evil act.	Killing the enemy of the state will save millions of lives.	Prison conditions are poor, and this study may help improve those conditions should the researcher present the information to a party outside of academia.

### Who is impacted by human-subject research?

As discussed above, a researcher's identity and the academic institution complicates research. As an able-bodied, Christian-raised, cis male, educated, married, straight, upper-middle class, white academic, I ought to tread carefully in regards to

research. If I create a study in an effort to advocate for trans women of color in the technical communication workplace, I inevitably benefit from their labor. On top of that, I run the high risk of projecting my own ideologies onto the study—not even as I write the results of the study, but in the *design* of the study itself. As a researcher, *not* projecting one’s own ideology is impossible. Any argument that says identity politics do not matter is probably coming from someone high on the spectrum of privilege, position, and power. My identity makes this more problematic as white males have a historically proven record of exploiting others for profit, and although my conducting such a study would have good intentions, the ethical implications of a man exploiting women is inevitable. In fact, in doing my research for this paper a friend shared a blog from a trans woman, Anne Tagonist (2009), writing to academic researchers. It comes as no surprise that the entire premise of this article was articulated by a marginalized individual nearly a decade ago:

Dear Mr. or Ms. Grad Student, I am sorry to report that I will not participate in your study as a data point. I don't understand what you're trying to accomplish. I don't trust you. I don't like you. I don't care if you succeed. In fact, I kind of think you suck. . . . I'm sure you have self-serving justifications. Everybody has self-serving justifications, it's how us humans get through the day. . . . What trans people need is to get through a day without being inspected, not by the guy making change at the WaWa and not by the hipster with an academic stipend. We need data, ideas, plans and strategies, but we need to see them coming from people like us, people who don't, right now, seem to make it into your little

position of power. We don't need your study, we don't need your thesis, and we really don't need you to graduate and 'do good work.' And you? You don't need us either. You are pretty much guaranteed a good life with or without my participation in your little project so please—stuff it up your ass.

This is the intellectual guerilla warfare conducted by marginalized people of which Vitanza speaks! What is the difference between her blog and this article? Privilege. Position. Power. I add my voice to hers and to Vitanza's. We need studies for marginalized people conducted by marginalized people.

**What are some current TPC best practices for ethical research that reduces exploitation?**

The doctrine of double-effect does not mean research should not happen. Five best practices of research have spurred from this uncomfortable awareness of the doctrine of double-effect (though never voiced in such phrases).

*The Three P's.* First, in their recent award-winning article, Jones, Moore, and Walton outline a method of disrupting dominant narratives in order to make room for other voices generally marginalized, which has been referenced throughout this article. They begin by reframing what technical communication research has been to what it should be: a focus on clarity and conciseness to a focus on listening to and with the subjects for whom we advocate. They offer a heuristic for such disruption, which they call the 3P's: position (identity markers), privilege (unearned advantages due to those identity markers), and power (having more influence than others). By understanding these questions as they pertain not only to the design of a research project but to the

researcher themselves as well, this heuristic has the ability to draw attention to who is speaking, who is being spoken to, and who is listening—all goals of social justice research.

*Symbolic-analytic workers.* Second, Johnson-Eilola's (1996) concept of becoming a symbolic-analytic worker:

Symbolic-analytic workers rely on skills in abstraction, experimentation, collaboration, and system thinking to work with information across a variety of disciplines and markets. Importantly, symbolic-analytic work mediates between the functional necessities of usability and efficiency while not losing sight of the larger rhetorical and social contexts in which users work and live. (p. 245)

Johnson-Eilola's work calls for technical communicators to balance what is functional while not losing sight of larger social, cultural, and rhetorical contexts in which they find themselves.

*Culture.* Third, specifically in regards to international and intercultural research, Agboka (2012) argues for a cultural approach to research in an effort to understand the subtleties of subcultures caught up in larger culture studies: "Not only do these ['large culture'] ideologies fail to account for cultural practices and values within less comprehensive groups within culture, but they do not accommodate the inputs individuals make in specific communication contexts" (p. 159). Agboka's concerns hint towards the doctrine of double-effect in that in an effort to do good, ethical research on a culture, the intricacies of the subcultures are assimilated into the large culture. To remedy this, he suggests researchers look to "culture" as an adjective rather than a noun: "Instead

of focusing on the noun, *culture*, which denotes a ‘thing,’ we [should] shift attention to the adjective, the *cultural*, which makes room for and recognizes group and individual agency, multiple contextual factors, shifting identities, and differences and similarities alike” (p. 170). This would, of course, reduce the “bad effect” as described in the doctrine of double-effect, but not eliminate it. To push Agboka’s argument—with a theoretical foundation of Levinas and Cavarero—even a study paying attention to the intricacies of a subculture risks assimilating the *individual*.

***Participatory localization.*** Fourth, in another article by Agboka (2013), he advocates for what he calls participatory localization. In technical communication, many have called for participatory methods before, which includes users in the design of a product in order to better meet their needs. Participatory localization takes this one step further in regards to international and intercultural research (though this principle can surely be applied in studies that do not identify as such). Rather than a user as an individual, which would risk the projection of that user’s ideology—coupled with the ideologies of the researcher—onto the product design, Agboka pushes for a user-in-community: “Simply, in this approach, users may determine what is relevant for them and work with a developer to create a product and documentation that are mutually beneficial to both of them” (p. 43). A user-in-community would understand issues of translation, local knowledge systems, and economic impacts, which situates the user in those larger social, cultural, and rhetorical contexts described by Johnson-Eilola.

***De-centering humans for human advocacy.*** Finally, Rose and Walton (2015) offer another best practice for research. As an initial critique of the goal of object oriented

ontological practices in technical communication research, Rose and Walton argue “that extreme positions that put all matter on equal footing can distract from and dilute a focus on human experience of inequalities” (p. 2), which is *the* goal of social justice research—advocating for disenfranchised humans. Simultaneously, they acknowledge that “[it] may seem counterproductive that theories which de-center humans could richly convey human experience, but an always-connected, always-collective, always-contextualized, and always-in-process perspective of humanity can lend complex insights into designing for humans, including those who are marginalized and otherwise oppressed” (p. 3). While not completely ignoring non-humans and not sacrificing the goals of social justice research, a de-centered human research practice can actually advocate for humans. Essentially, by complicating the wicked problem even further, these three best practices offer researchers with a better path for a more ethical practice of social justice research.

These heuristics do a great deal in reducing the negative impacts of the doctrine of double-effect, but they does not eliminate it. When working with human subjects, there is little to no way to eliminate exploitation—when working with human subjects.

**How do I conduct ethical research that advocates for humans without exploiting them?**

The answer, for me, is simple—don’t do human-subject research. *The Belmont Report* explains in its summary that the report “is a statement of basic ethical principles and guidelines that should assist in resolving the ethical problems that surround the conduct of research with human subjects” (p. 1). As we know, however, “resolving” ethical problems with human subjects is only possible in that researchers justify the

exploitation through the doctrine of double-effect. How TPC resolves these ethical dilemmas is not the concern for this article. Where *The Belmont Report* and other ethical commentaries in TPC focus on what the field should do, doing so risks top-down assimilation into larger cultures cautioned by Agboka (2012; 2013). As mentioned earlier, these questions are personal with which each scholar ought to engage, so I can only offer what I feel is the most ethical way to conduct social justice research for me. As an able-bodied, Christian-raised, cis male, educated, married, straight, upper-middle class, white academic, I embrace social justice by disrupting the dominant narratives of which I am apart. For me, I cannot and should not do social justice research that should be done by those who are most impacted by it. Those spaces—rightfully so—are being and should continue to be filled by those who should be filling them, calling back to Tagonist’s (2009) open letter to academics.

We speak of institutionalized oppression, but when these circumstances are uncovered, those arguments—especially those coming from marginalized researchers—are often dismissed as having agendas, being oversensitive, or committing an anecdotal fallacies. Rather than use TPC as a venue for conducting human-subject research advocating for oppressed people, I use TPC to interrogate institutional genres of oppression using computational content analysis, providing data-driven results that reveal the patterns of oppression in language. Rickert’s (2013) *Ambient Rhetoric* addresses the things that happen in the background that impact how we live our lives—Rivers and Weber (2011) call these “mundane artifacts.” It is these ambient, mundane artifacts that help build, sustain, and replicate institutional oppression (cf. Althusser; Foucault). As

stated by countless marginalized individuals, drawing attention to institutional oppression for those who are oppressing is unfair labor. Carson (2015) articulates it this way:

Our society claims that there is a need to address racism while disavowing that almost anyone is a racist. But conversations about race will never be effective if white people keep looking to me (or someone like me) to be the equivalent of an expert witness to testify that they as individuals and we as a collective or institution were not, are not and potentially have never been racist, intentionally or otherwise. . . . And I'm tired of them expecting me to tell other people that on their behalf.

The work of dismantling those oppressive institutions ought to come from those who benefit from it the most. People like me. And so, I will do the work that I feel I ought to do, and this work serves two primary purposes: 1) to dismantle institutional oppression through computational content analysis of genres, and 2) provide open access to all assets of the study (datasets, algorithms, results, etc), providing data-driven research that will enable and support the intellectual guerilla warfare conducted by marginalized individuals. I may not be able to dismantle entire institutions, but I can crack them open for deep interrogations. Interrogations that carry more authenticity coming from marginalized, unrepeatable, unique individuals.

### **Concluding Thoughts**

Several years ago, I watched an episode of *Last Week Tonight with John Oliver* (2014) that focused on America's prisons. After the twenty-minute report, the episode ended with a Sesame Street parody. True to the genre of public broadcasting, John Oliver



and the Sesame puppets concluded with, “America’s prison epidemic is brought to you by decades of neglect, the GEO group, and viewers like you.” Those last words pierced me. I was called forth by the Other; I saw their faces, and I heard their voices. I knew then that I wanted to do social justice research. When I read Lanham’s “Q” Question, I faced the problem of what it meant to *do* that research. These incarcerated individuals are already being exploited, and my own research would further exploit them. Coming from a position of privilege and power, that inevitable exploitation cut deeper. I knew I had to do social justice research, but I just didn’t know how to do it in the most ethical way possible. According to Levinas and Cavarero, however, not only am I pulled into existence by the face and the voice of the Other, but my dissertation project is as well. It is/was always-already about social justice, which answers the first part of the revised “Q” Question. But how do I conduct that research without exploiting those for whom I wish to advocate? How can I reduce the negative impacts of the doctrine of double-effect? As a researcher, I am unrepeatable and unique, and I wish to advocate for unrepeatable, unique individuals incarcerated in America’s prison system not by studying prison cultures but by studying the mundane. I focus my attention on the authors of the dominant narratives I wish to disrupt. While this study design may not focus on the unrepeatable, unique voices of those who are and have been marginalized, it will peel back the layers of position, privilege, and power dominating the practice of incarceration in prisons, providing a space where others can speak and where we can all listen.

As you are undoubtedly aware, I do not go into the specifics of what a computational content analysis of institutional genres looks like. This is intentional. This

is not an article about methodologies. It's an article about wrestling with the tensions of wanting to do "good" work, wanting to help people, wanting to understand other people, wanting to work in a way that doesn't work against those people, wanting to dismantle my privileges and not reproduce them. I don't recommend every scholar approach these questions the way I did, but we need to approach them. The "we" here is not what we need to do as a field, it's what we need to do as unrepeatable, unique individuals who want to do good work.

## References

- Agboka, G. (2012). Liberating intercultural technical communication from “large culture” ideologies: Constructing culture discursively. *Journal of Technical Writing and Communication*, 42(2), 159–181. <https://doi.org/10.2190/TW.42.2.e>
- Agboka, G. (2013). Participatory localization: A social justice approach to navigating unenfranchised/disenfranchised cultural sites. *Technical Communication Quarterly*, 22(1), 28–49. <https://doi.org/10.1080/10572252.2013.730966>
- Althusser, L. (1971). Ideology and ideological state apparatuses. In B. Brewster (Trans.), *Lenin and philosophy, and other essays* (pp. 127–186). New York: Monthly Review Press.
- Carson, A. D. (2015, September 18). It’s not my job to absolve white friends of racism, but it can seem that way. Retrieved March 18, 2018, from <http://www.theguardian.com/commentisfree/2015/sep/18/not-my-job-to-absolve-friends-of-racism>
- Colton, J. S., Holmes, S., & Walwema, J. (2017). From NoobGuides to #OpKKK: Ethics of Anonymous’ tactical technical communication. *Technical Communication Quarterly*, 26(1), 59–75. <https://doi.org/10.1080/10572252.2016.1257743>
- Connors, R. J. (1999). The rise of technical writing in America. In Kynell & M. G. Moran (Eds.), *Three Keys to the Past: The History of Technical Communication*. Greenwood Publishing Group.
- Elliot, N. (2016). A theory of ethics for writing assessment. *The Journal of Writing Assessment*, 9(1). Retrieved from <http://journalofwritingassessment.org/article.php?article=98>
- Foucault, M. (1977). *Discipline and punish: The birth of the prison* (1st American ed). New York: Pantheon Books.
- Hopton, S. B. (2013). If not me, who?\*: Encouraging critical and ethical praxis in technical communication. *Connexions: International Professional Communication Journal*, 1(1), 65–68.
- Johnson-Eilola, J. (1996). Relocating the value of work: Technical communication in a post-industrial age. *Technical Communication Quarterly*, 5(3), 245–270. [https://doi.org/10.1207/s15427625tcq0503\\_1](https://doi.org/10.1207/s15427625tcq0503_1)
- Jones, N. N., Moore, K. R., & Walton, R. (2016). Disrupting the past to disrupt the future: An antenarrative of technical communication. *Technical Communication Quarterly*, 25(4), 211–229. <https://doi.org/10.1080/10572252.2016.1224655>

- Katz, S. B. (1992). The ethic of expediency: Classical rhetoric, technology, and the Holocaust. *College English*, 54(3), 255–275. <https://doi.org/10.2307/378062>
- Lanham, R. A. (2010). The “Q” question. In *The electronic word: Democracy, technology, and the arts* (pp. 154–194). University of Chicago Press.
- Lévinas, E. (1978). *Totality and infinity: An essay on exteriority*. Hebrew University Magnes Press.
- Longo, B. (1998). An approach for applying cultural study theory to technical writing research. *Technical Communication Quarterly*, 7(1), 53–73. <https://doi.org/10.1080/10572259809364617>
- MacIntyre, A. C. (1988). *Whose justice? Which rationality?* Notre Dame, Ind: University of Notre Dame Press.
- Mangan, J. T. (1949). An historical analysis of the principle of double effect. *Theological Studies*, 10(1), 41–61.
- Mckesson, D. (2016, January 19). DeRay Mckesson helps Stephen address his privilege [YouTube]. Retrieved from <https://www.youtube.com/watch?v=qffCO1b-7Js>
- Miller, C. R. (1979). A humanistic rationale for technical writing. *College English*, 40(6), 610–617. <https://doi.org/10.2307/375964>
- Prison: Last Week Tonight with John Oliver*. (2014). HBO. Retrieved from [https://www.youtube.com/watch?v=\\_Pz3syET3DY](https://www.youtube.com/watch?v=_Pz3syET3DY)
- Rickert, T. J. (2013). *Ambient rhetoric: The attunements of rhetorical being*. Pittsburgh, Pa: University of Pittsburgh Press.
- Rittel, H. W. J., & Webber, M. M. (1973). Dilemmas in a general theory of planning. *Policy Sciences*, 4, 155–169.
- Rivers, N. A., & Weber, R. P. (2011). Ecological, pedagogical, public rhetoric. *College Composition and Communication*, 63(2), 187–218.
- Rose, E. J., & Walton, R. (2015). Factors to actors: Implications of posthumanism for social justice work. In *Proceedings of the 33rd Annual International Conference on the Design of Communication* (p. 33:1–33:10). New York, NY, USA: ACM. <https://doi.org/10.1145/2775441.2775464>

- Ryan, K. J., Brady, J. V., Cooke, R. E., Height, D. I., Jonsen, A. R., King, P., ... Turtle, R. H. (1979). *The Belmont Report: Ethical Principles and Guidelines for the Protection of Human Subjects of Research*. The National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research. Retrieved from [https://www.hhs.gov/ohrp/sites/default/files/the-belmont-report-508c\\_FINAL.pdf](https://www.hhs.gov/ohrp/sites/default/files/the-belmont-report-508c_FINAL.pdf)
- Salvo, M. J. (2001). Ethics of engagement: User-centered design and rhetorical methodology. *Technical Communication Quarterly*, 10(3), 273–290. [https://doi.org/10.1207/s15427625tcq1003\\_3](https://doi.org/10.1207/s15427625tcq1003_3)
- Sapp, D. A., Savage, G., & Mattson, K. (2013). After the International Bill of Human Rights (IBHR): Introduction to special issue on human rights and professional communication. *5756219854*, 4(1), 1–12.
- Spivak, G. C. (1989). Feminism and deconstruction, again: Negotiating with unacknowledged masculism. In *Between Feminism and Psychoanalysis* (pp. 206–223). New York, N.Y: Routledge.
- Tagonist, A. (2009, December 10). Fuck you and fuck your fucking thesis: Why I will not participate in trans studies. Retrieved March 15, 2018, from <https://tagonist.livejournal.com/199563.html>
- Vitanza, V. J. (1987). Critical sub/versions of the history of philosophical rhetoric. *Rhetoric Review*, 6(1), 41–66.
- Walton, R. (2016). Supporting human dignity and human rights: A call to adopt the first principle of human-centered design. *Journal of Technical Writing and Communication*, 46(4), 402–426. <https://doi.org/10.1177/0047281616653496>

## ARTICLE 2

### ALGORITHMIC AURAS AND INSTITUTIONAL GENRE ANALYSIS: A CASE STUDY FOR WORKING ETHICALLY WITH BIG DATA

#### **Introduction**

In building his argument centered on user design, Albers (2003) refers to three levels of situation awareness: knowing the data, knowing the interrelations within the data, and being able to make predictions or relate information to the larger picture (pp. 264-265). He cites these levels as a foundation for engaging with content analysis as a means to address complex problems, but *data* meant something different in 2003 than it does today. According to a 2013 SINTEF study, “90% of all the data in the world has been generated over the last two years,” and IBM estimates that 2.5 quintillion bytes of data are created daily—another stat from 2013, and 2013 was a long time ago in the digital realm. Since then, data has grown exponentially, moving from Albers’ conception of a complex problem to what Rittel and Webber (1979) would call a wicked problem: problems do ill-defined there may be no solution to be had. While the sheer amount of data may be astounding, what is done with the data becomes even more problematic. This article is not necessarily a critique of how corporate giants and dominant governments use and manipulate that data. It’s about how academic researchers ought to approach big data and the algorithms that construct them. Using the theoretical work of Walter Benjamin and Adrianna Cavarero, I first re/establish the concept of identity reframed into what I call “algorithmic aura”; second, I situate big data methodologies as ideological frameworks along with the messy ethical implications of user-generated data; and third, I offer examples from my own research about how to ethically engage with big data

through a method I call “institutional genre analysis,” which allows for research centered on the preservation of human dignity. Ultimately, this theoretical framework and best practices may be used as a heuristic for scholars looking to repurpose big data methodologies in an effort to make room for what Vitanza (1987) calls “intellectual guerilla warfare conducted by [marginalized individuals]” (p. 52).

This concept must be stressed. There are a myriad of different ways to do big data research, and researchers must be aware of the consequences of adopting those methodologies: “You adopt a methodology, you adopt its flaws” (Stephens, 2017). What makes institutional genre analysis significant is *intention*. Several of the concerns about the emerging methodology of big data discussed in this article center on the impacts these methods have on individuals, but this method focuses on empowering individuals—specifically those marginalized individuals who seem to constantly be a source of data and not the ones conducting the studies themselves. In his discussions on new rhetorical histories, Vitanza (1987) calls for this intellectual guerilla warfare conducted by marginalized individuals. More than three decades later that call must be heard even louder today. Institutional genre analysis has its foundations in Blyler’s (1995) critical interpretive ideology, which focuses on uncovering large oppressive social patterns. This methodology, I contend, is part of the fundamental shift towards ethics and social action in the various fields of writing and composition (see Inoue, 2015; Jones, Moore, & Walton, 2016; and Colton, Holmes, & Walwema, 2017). The research we do as scholars has impact outside of our own conception of the field, and often the harm researchers do to enact social action is unseen or ignored. Institutional genre analysis is designed to

advocate for individual users by studying an institution's own writing they produce, allowing machine learning to make connections that a traditional reading simply could not identify. I am not a marginalized individual, but I want to participate in this intellectual guerilla warfare. The theoretical foundation weaved from Benjamin and Cavarero enables researchers to understand the deep, ontological issues with big data, enabling researchers to actively preserve human dignity and advocacy without replicating algorithmic auras.

### **Replicating the Unrepeatable, Unique Individual**

#### **Ethical Problems with Big Data**

Before delving into the deeper problems of big data, it will be helpful to have a brief review of literature regarding the concerns raised about big data analyses from a variety of fields looking to apply its methods. In doing so, I move forward with the same assumption voiced by Halavais (2015): "There seem to be more people with opinions about big data than there are studies utilizing large social data sets" (p. 583). In their oft-cited article on problems with big data, boyd [sic] and Crawford (2012) clarify what it is: "Big Data is less about data that is big than it is about a capacity to search, aggregate, and cross-reference large data sets" (p. 663). They draw our attention to some of the sticky, messy problems of big data; some of which include how big data shapes knowledge, claims of objectivity and accuracy, taking data out of context, accessing big data despite privacy concerns, and access to the methodology of big data itself. They argue that big data has ushered in a profound shift in how we think about research regarding epistemology and ethics. According to O'Neil's (2016) *Weapons of Math Destruction*,



these algorithms and data shape our lives in ways we had not realized, often drawing conclusions on data we as users willingly provided for nearly nothing. Big data changes how we engage with knowledge, and while some researchers have the capacity and understanding to ask critical questions, most are met with black boxed algorithms and proprietary processes.

From the field of journalism, Fairfield and Shtein (2014) contend that “social scientists are undergoing a fundamental shift in the ethical structure that has defined the moral use of these techniques” (p. 38). This regrounding forced upon researchers brings questions of morality and ethics back into the lime light. Much of the user-generated data and its metadata is collected without informed consent from millions of users, a basic tenant of *The Belmont Report* (1979) and the International Review Board (IRB) approval process. Users do provide data “free” to the world if they do not have their setting set to private—even then corporations have incredible access to a user’s data. While there may be an argument to be had about a user’s responsibility in understanding the terms and conditions of whatever application they are using on their phones, the questions and responses of expectations of privacy and user responsibility reflect this fundamental shift in ethics referred to by Fairfield and Shtein.

This shift, however, focuses on user-generated data. In his article, Halavais (2015) makes a similar distinction by referring to “big social data,” datasets using information from social networks. “The real danger,” he argues, “is allowing some combination of availability, methods, marketing, and scholarly fashion to bend and shape social research rather than being guided by a deeper sense of inquiry” (p. 583). This is not a suggestion

to abandon the method, but he emphasizes the importance of creating a study informed by established theories and ideologies. Lewis and Westlund (2015) discuss this shift as more than a shift in knowledge creation but “a sociological phenomenon with cultural, economic, and political origins and implications; it is, indeed, a mythology as much as a science or business” (p. 2). What big data can and cannot do remains a mystery to most (if not all), and in that mystery people seem to fall into acceptance rather than a critical questioning. One problem of this shift in thinking and acceptance, according to Mahrt and Scharkow (2013), is the temptation for researchers to use data for research that was not specifically designed for research. If a researcher finds a dataset and does not know how the raw data came to be a completed database, then the researcher adopts the unacknowledged biases and ideologies of the people who created the database itself.

### **Algorithmic Auras and Human Dignity**

What big data can reveal about society is one of the newest and most original set of questions and methods in recent research methodological history. Yes, there are questions of privacy and bias, but the temptation of big data as some sort of omnipotent methodology is just that—tempting. What, then, does the mythology conceal? Alluded to by Fairfield and Shtein (2014) is the cost of dignity and personhood of individuals:

Consider the problem of informational harms. On a cost-benefit analysis, leaking someone’s data as part of a big dataset may not be catastrophic. In dollar figures, the cost to a consumer of being part of a data spill may be low. But the cost to the dignity and personhood of an individual whose entire search history has been

exposed to the world can be significant, if not easily measured in cost-benefit terms. (pp. 42-43).

While Fairfield and Shtein appeal to human dignity, the problem goes deeper still. In their critique, they limit the scope of dignity to that of the consequences of a user's data that may be accidentally exposed in the process of research. Even before the user's data has the opportunity to be exposed, their dignity as a human is already at stake. An obvious problem of big data analyses is how it shapes society; as Foucault (1977) might argue, it is not a question of *if* but of *how*. The risk of generalized results applied to whole cultures would be a hallmark use of Foucault's Panopticon and issues of position, privilege, and power. Big data is a cultural phenomenon (boyd and Crawford) *because* it impacts and normalizes behavior and society.

Walter Benjamin's concept of the aura—coupled with the work of Adrianna Cavarero—complicates a Foucauldian analysis of big data as a cultural and sociological phenomenon. Benjamin (1936) applies aura to art:

[T]hat which withers in the age of mechanical reproduction is the aura of the work of art. . . . To pry an object from its shell, to destroy its aura, is the mark of a perception whose "sense of the universal equality of things" has increased to such a degree that it extracts it even from a unique object by means of reproduction.

He applies his critique to a live performance of *Macbeth* to a filming of the play:

For aura is tied to his presence; there can be no replica of it. The aura which, on the stage, emanates from *Macbeth*, cannot be separated for the spectators from that of the actor. However, the singularity of the shot in the studio is that the

camera is substituted for the public. Consequently, the aura that envelops the actor vanishes, and with it the aura of the figure he portrays.

While new materialist and object-oriented ontology scholars look to Benjamin to flatten the plane between objects and subjects, I use Adrianna Cavarero's work for an obverse reading of Benjamin's aura in an effort to *focus on the subject*—what Cavarero calls the unique, unrepeatable individual (2000, p. 2). For Benjamin, the filming of an actor is the replication of the aura he portrays and gives to the audience. If a filming of a play causes the play as an object to lose its aura, what of the replication of a person?

Cavarero's (1995) critique of a male-centered philosophical tradition surely applies to the philosophical work of Benjamin (and especially Foucault's phallogocentric Panopticon). She writes: "Here, a male subject claiming to be neutral/universal declares his central position, disseminating a sense of the world cut to his own measure and revealed in his own mythic figures" (p. 2). Rather than use Benjamin's concept of the aura to engage with the metaphysical, an attempt to make a universal claim, I read aura as a person's *identity*. If an object's replication causes that aura to diminish, then it follows that a person's aura must also diminish if reproduced. In 1936, the idea of replicating an identity was science fiction and has been until recently. With the continuing emergence of this digital data-driven era at an exponential rate, a person's aura can be replicated to a remarkable degree of accuracy. Cheney-Lippold (2011) calls this a "new algorithmic identity":

The networked infrastructure of the internet, with its technological capacity to track user movements across different websites and servers, has given rise to an

industry of web analytics firms that are actively amassing information on individuals and fine-tuning computer algorithms to make sense of that data. The product of many of these firms is a “new algorithmic identity,” an identity formation that works through mathematical algorithms to infer categories of identity on otherwise anonymous beings. (p. 165).

What Cheney-Lippold calls categories of identity, Cavarero (2016) would call inclinations:

To incline is to bend, to lean down, to lower. . . . Not all the phenomena that language ascribes to the term *inclination* interests philosophers; indeed many possible meanings remain consistently marginalized to speculative turbulence and receive little attention from philosophers. . . . Besides posing a moral problem for the modern conception of the self, inclination is a matter of structural equilibrium and thus, in the end, becomes an ontological question as well. (pp. 3, 5, & 6)

These categories of identity show the ontological inclinations of unrepeatable, unique individuals whose auras are replicated and sold over and over again. Most critique of big data as a methodology will cite concerns for privacy or informed consent, and those are problems with which every scholar ought to engage when approaching big data, but the real ethical dilemma is an ontological one. Fairfield and Shtein come so close to the critical problem of big data as it pertains to a person’s dignity and personhood, but the risk is not having their search histories exposed. The risk is in the replication of a person’s identity—their algorithmic aura.

## Big Data Methodologies as Ideologies

If one were to apply Cavarero's critique of inclinations to big data, the first mythology to fall would be the idea that big data provides *the* answer—the methodology void of rhetoric and bias. This is simply not true. While big data may eliminate some bias, it introduces others. boyd and Crawford (2012) argue that big data should be defined “as a cultural, technological, and scholarly phenomenon that rests on the interplay of [technology, analysis, and mythology]” (p. 663). While Google and the NSA have entire facilities dedicated to housing data, an academic researcher might be able to fit all of their data on a flash drive. More than the amount of data is what is done with the data—how it is collected, how it is aggregated, and how it is searched. In an effort to redirect the ethical use of big data analysis, I make an important separation in types of data: user-generated data versus genre-generated data. Before making this distinction, however, it is important to understand that *all* methodologies are ideologies based on theories, tools, technologies, archives, datasets, and other ideological tendencies. Put into other words:

Research techniques are sometimes seen as atoms or essential building blocks of research projects: invariant, inviolable steps that are applied the same way, no matter what the socioeconomic characteristics of the environments in which they are deployed. That is, they are often seen as arhetorical, and rhetorical choice and agency play a role in how they are arranged and implemented. (Spinuzzi, 2005, p. 411)

From this understanding, I offer not so much a critique but an alternative use of big data methodologies grounded in the ambient, mundane artifacts of cultural institutions in an effort to protect human dignity.

### **All Methodologies Are Ideologies**

**Theories.** While it is tempting to think of big data analysis as a search for the objective reality, it simply falls short, just as every other methodology that makes such claims. In his essay on using theory in technical communication research, Porter (2013) suggests that researchers turn to conceptual maps of the field as a starting place, a heuristic for deeper understanding. In discussing theorizing as an activity, Porter writes, “theorizing, or reflecting critically about the strengths and weaknesses of various theories, can expand the way we think by challenging our existing frameworks and giving us new ways of seeing” (p. 128). However, if we rely too heavily on what “our” field has done and some of the directions it can go, then we might blind ourselves. Those conceptual mappings orient researchers into seeing in one particular way of doing things. Porter encourages researchers to turn to theory in an effort to unground the way we see the field and the way we think and asking questions grounded in a variety of theories forces us to revise what we see. An outstanding example of such a practice is Jones, Moore, and Walton’s work, “Disrupting the Past to Disrupt the Future,” where they critique the conceptual mapping of technical communication as a field and use a theoretical heuristic (the 3Ps: position, privilege, and power) in order to reorder the dominant narrative, making room for marginalized voices to speak and to be heard. Even before a methodology is applied, theories and ideologies are employed, shaping the

questions we even ask (sometimes without our even realizing it). An awareness of those theories and ideologies don't eliminate that bias, but it can reduce it.

**Tools and technologies.** In addition to the theories that inform methodologies, the tools for data collection and analysis are just as important with just as many implications. As Kranzberg's work suggests: "technologies is neither good nor bad; nor is it neutral" (qtd. in boyd & Crawford, 2012, p. 662). In his essay, Swarts (2013) discusses this non-neutrality of technology, suggesting that the tool one uses shapes the social interactions of research as much as the practice of research itself: "A tool shapes both the practice of technical communication and the social interactions that technical documents foster. . . . A more common way to describe how tools shape and organize activities is to say that they 'mediate' those activities by imposing a structure on them" (p. 149). As research methodologies continue to rely more and more heavily on the mediation of tools, the importance of learning *which* tools to use for which projects becomes more important. A scholar may turn to a traditional rhetorical reading of a translated document (Katz, 1992) or to something more creative like collage work in Rwanda (Walton, Zraly, & Mugengana, 2015). Whichever tool is used to conduct the method, every tool has its own history and ideology, and Swarts (2013) argues that "[t]ools have histories, and by reading a tool's history one can understand how that tool has shaped an activity over time and how those mediating influences persist in the accumulated design" (p. 149). What's important to recognize here is the idea that tools and methodologies have impact—they are not arhetorical as Spinuzzi reiterates. Feenberg (2002) goes so far as to say that the



tool and technology we use not only shapes the field and our research but shapes *us* as well.

**Archives and datasets.** In a critique of historical research, Meng (2016) deconstructed the objectivity of historical artifacts by drawing on two of the above factors: theory and tools. He argued that while some (if not most) historians claim objectivity in their artifacts, they cannot claim such because their own ideologies and tools already shape what is and is not included as an artifact: “The current model of scholarly engagement tends to involve one scholar correcting the views of past scholars. The model is driven by correctness. For historians, the impulse to correct issues in an obsession with evidence, especially unpublished evidence conserved in an archive.” Meng’s critique of the historical archive as a source for truth is not unlike big data’s appeal to objective truth: just as a historian decides which archive she decides to research, a big data methodology is already subjected to ideology by the source of the data itself. Was it collected on Facebook? Twitter? The U.S. Census Bureau? Which dataset a researcher chooses further removes the concept of objectivity in big data methodologies. More importantly than which dataset a researcher uses is how the results of that dataset are applied to the larger picture. In big data analysis, many jump to apply the results to larger cultures. Fairfield and Shtein (2014) emphasize that a database has the potential to “capture entire communities. This can be invaluable in researching a particular subject, but the difficulty is in respecting the rights of other community members who may not be the subject or research or who have not given consent” (p. 45). There may be other members of such a community that do not participate in social media

platform from which the results of a study might have a significant impact on their lives. As Agboka (2012) cautions, a “large culture” study risks alienating and assimilating subcultures—and when his thinking is pushed further—these studies assimilate individuals as well.

**Two Ideological Tendencies.** Ideological biases are manifested in the theories, the tools, and the datasets a researcher turns to for their projects. According to Blyler (1995), there are two prominent ideological tendencies or leanings that each researcher faces: functionalism vs. critical interpretations. A functionalist ideology maintains that reality is external, and research is used to discover what that reality may or may not be. Many corporate proponents of big data methodologies echo this functionalist perspective, suggesting that its big data analysis simply observes that which is already happening—an external reality. The damage of such a perspective is generalizability (as voiced in O’Neil’s *Weapons of Math Destruction*). If a claim founded from big data measures an external reality, then those results must also be applicable to a larger group of people. This carries a damning risk of assimilation into a larger culture cautioned by Agboka (2012) where big data does its most harm. O’Neil (2016) discusses several different examples of a functionalist analysis with devastating results to individual people: the housing crisis of 2009, the college ranking system from US News, test score evaluations to determine teacher effectiveness, and crime statistics used to deploy police officers in certain areas. A critical interpretive ideology, on the other hand, believes that reality is constructed. Each perspective seeks to understand the larger whole, but the main difference is intent. A functionalist seeks to describe and prescribe, but “the goal, then, of

critical interpretive research is . . . the ability to link experience to larger, oppressive social patterns” (p. 304). With that critical interpretation, a researcher will offer alternatives rather than prescriptions. According to boyd and Crawford (2012), big data methodologies preach functionalism, but it must be taken up with a critical interpretation in order to avoid generalization and assimilation of subcultures and individuals.

### **Two Types of Data**

In her article about service courses, Gulbrandsen (2012) recognizes a new economy, which is “an economy in which knowledge production is no longer contained within localized economic structures, but is vast and diffuse” (p. 247). While big data is the “buzzword *du jour*” as described by Lewis and Westlund (2015), Gulbrandsen calls for researchers to be able to interpret big data analyses, arguing that there is “a shortage of the analytical and managerial talent needed to interpret and use big data as well as to recognize and manage its value” (p. 247). One of the first steps in critically interpreting analyses and creating studies utilizing big data methodologies is to question the source of data. Calling back to Meng (2016), the archive matters.

**User-generated data.** When most people think of big data, they think of user-generated data. An example from Facebook. An individual user sees a quiz that will identify which Hogwarts house, which Game of Thrones house, or which Disney princess they are. The user clicks the quiz, which is essentially a personality test, and answers the questions. This is data entered directly by the user. The host of the quiz will also take data *about* the user-data, which is metadata or trace data. This could be a timestamp, a GPS location, which websites the user came from and where they went afterwards, how

long the user took to answer questions, other information about the user accessed through action of clicking (friends of the user, likes of the user, etc). This metadata and user-data are then repurposed and sold repeatedly to political campaigns, advertising companies, and academic researchers (Wakefield, 2015).

**Genre-generated data.** User-generated data is rife with ethical implications that can be used to generalize and assimilate (discussed below). Genre-generated data is different (although researchers will still collect metadata or trace data about what was collected). Where user-generated data comes from individual users, genre-generated data treats text as data, specifically from institutions rather than individuals—institutional genre analysis. I take the term “genre,” a common operationalized concept in technical communication, while mindful of Miller’s (1984) “Genre as Social Action,” Douglas’s (1986) *How Institutions Think*, Spinuzzi’s (2003) *Tracing Genres through Organizations*, and Graham et al.’s (2015) “Statistical Genre Analysis.” While treating text as data is an emerging practice of big data analysis, this distinction helps to establish the intent of such an analysis. Institutional genre-generated analyses begin with the premise of Blyler’s (1995) ideological tendency to critically interpret the data in order to uncover oppressive social patterns.

### **Institutional Genre Analysis**

There are deep problems with big data if the application and sources of the data go unquestioned. This is not to say that the tool of big data is itself good or bad, but as Kranzberg cautions, a tool is not neutral. Revealing the loss of human dignity through the mechanical reproduction of algorithmic aura is not meant to add to the list of critique of

big data, but to situate an alternative source of data that holds special interest for various writing scholars. Genre-generated data removes the privacy concerns and replicated auras because it does not collect data from individual users. Instead, it collects data from an institution in the form of text in any given genre. Immediately, one might realize that this will take quite a bit more work than tracking a hashtag. Good. The process of using text as data is as arduous as it is tedious. Few databases of genre-defined data exist, meaning researchers must curate their own datasets, which includes defining the genre, finding hundreds (if not thousands) of examples of the genre, formatting the documents, parsing out the language, and designing the algorithms to use. According to Longo (1998), an institution is a cultural agent that shapes society, and if an institution is seen as producing objective truths (like people's perceptions of science and big data), then their use of power and dominance goes unquestioned. A user-generated dataset risks reproducing the power and dominance of institutions, but a genre-generated dataset moves from a functionalist perspective to a critical interpretive perspective. No longer is the researcher trying to make sense of society by how impacts are manifested by users; instead, a researcher can ask more critical questions of the cultural institutions themselves.

### **Establishing Best Practices for Big Data Methodologies**

Before a study is even considered, a researcher has ideologies and biases, impacting how they approach a problem—even the selection of the problem reflects ideological biases. Once a topic is decided upon, a researcher must also determine which tools they will use to conduct their study. Whether one chooses ethnography, close content analysis, surveys, or any number of methodologies, that decision reflects

ideological biases and is a rhetorical decision. The decision about the archive or dataset a researcher chooses is often an economic one where money is not the only currency but time and access as well. A researcher is limited by each of these factors when choosing the dataset, and the dataset can shift the questions being asked. Finally, concerning functionalist vs. critical interpretive ideologies, a researcher makes a rhetorical decision (sometimes unconsciously) about how this research will be used: the heavy risk of generalization and exponential growth or the disruption of dominant narratives. Every un/conscious decision shapes the study, the results, and what is done with the results

Using texts as genre-generated data still has its issues. When turning to a dataset comprised of text, Grimmer and Stewart (2013) encourage every researcher to begin with the same assumption: “We emphasize that the complexity of language implies that automated content analysis methods will never replace careful and close readings of texts” (p. 268). In fact, they argue that the complexity of language and contexts means all methods of content analysis using big data are necessarily *incorrect*, and researchers should consider these methods “as *amplifying* and *augmenting* careful reading and thoughtful analysis” (p. 268). Not only is language complex, but the language used to train machine learning is limited to a single genre. This is important, although possibly frustrating, concept. The implication here is that an analysis used on one genre from one institution may not work across other genres and/or institutions. Big data isn’t *the* answer to anything, but it can enhance a researcher’s understanding of oppressive social patterns manifested in the mundane artifacts of institutional genres. When turning to big data and institutional genre analysis, researchers should consider the following best practices.

Table 3.1 below outlines these best practices, offer an interpretation, and a concrete example of its application as a case study from my own research regarding prisons where I collected 347 inmate handbooks, totaling nearly 16,000 pages and more than 425,000 unique n-grams.

Table 2.1 Establishing best practices for big data methodologies

	<b>What is the best practice?</b>	<b>What does it mean?</b>	<b>What does it look like?</b>	<b>How is does it support the intellectual guerilla warfare conducted by marginalized individuals?</b>
<b>O'Neil's (2016) Weapons of Math Destruction</b>	"Even if the participant is aware of being modeled, or what the model is used for, is the model opaque, or even visible?" (p. 28)	Have you made your assets available for replication and/or critique?	In order to enable other researchers to engage with my data, every handbook, line of code, and the method of content analysis data (i.e. the codebook) will be made available.	By making all assets available, not only can the study be replicated, but the results can be interrogated by those who are impacted most by the institution.
	"Does the model work against the subject's interest? In short, is it unfair? Does it damage or destroy lives?" (p. 29)	Can this analysis and database be used or co-opted to further oppress marginalized individuals?	Rather than turn to user-data that may be used against individuals, my analysis approaches the institution in order to dismantle it.	In this case, inmate advocates would be able to use this study to point to data-driven examples of oppression across time and space manifested in this particular genre of inmate handbooks.
	Does the "model have the capacity to grow exponentially?" (pp. 29-30).	Do the results have potential impact outside of its immediate sphere of influence?	By centering the analysis on the heuristic of genre, it reemphasizes the contextual nature of the data, meaning the analysis necessarily should be able to scale to another genre of communication.	The mundane artifacts (see Rivers and Weber, 2011) that shape our institutions can be used to identify oppression and subvert it.



Nelson's (2017) "Computational Grounded Theory"	"Pattern design using human-centered computational exploratory analysis." (p. 9)	This method of unsupervised learning analyses the corpus of text in an effort to see patterns in language that an un-aided reading would produce. Unsupervised learning allows the data to speak to the researcher without the researcher's ideological biases impacting the questions they ask.	In this first step of analysis, the computer interrogates the text for me. The computer doesn't know what the word "property" means, but based on its proximity to other words a latent meaning emerges.	Building an archive and providing the assists for the study enable the machine learning to uncover patterns in the language that my own ideological tendencies may block.
	"Hypothesis refinement using human-centered interpretation. Grounded theory involves moving back and forth between the results of the analysis and the data" (p. 23).	Once the unsupervised learning uncovers language patterns, the researcher can study see those patterns and begin asking pointed questions of single documents comprising the corpus. Many quantitative researchers might be opposed to this order, but with big data analysis like this, hypotheses are often formed, modified, or scrapped based on the insights of the learning model.	With patterns established, I can begin asking questions of the data to build a supervised learning model, which is when a researcher explicitly imposes an ideological framework on the data. In the case of "property," I can ask a question of a single document like: "What latent meanings exist between 'property' and the concept of rehabilitation? Is property something <i>given</i> to inmates? Or is it something <i>taken away</i> as a punishment?"	While the machine's unsupervised learning uncovered patterns the my own ideology may have blocked, opening the data to inmate advocates would allow for individuals with other ideologies to ask even more questions about individual documents blocked by my own ideology.
	"[Pattern confirmation] tests whether the patterns identified in the first two steps are generalizable to the	Note that this step asks if the patterns hold in the entire corpus, not the entire genre. After using the unsupervised learning to engage in a close	A close reading of a document is subject to a critique of interpretation; however, if I have established a pattern through my close reading, I can test it against	This is, perhaps, the strongest tactic for intellectual guerilla warfare. A close reading of a single text coming from a marginalized individual could

	entire corpus" (pp. 28-29).	supervised learning of individual texts, researchers can test the results against the entire corpus.	the entire corpus. This moves from a shaky reading of a single text to data-driven evidence of latent meaning based on machine learning.	be dismissed as a clouded reading based on identity politics. Being able to show systematically how these oppressive patterns exist across time and space adds an incredible amount of <i>ethos</i> to their argument.
--	-----------------------------	------------------------------------------------------------------------------------------------------	------------------------------------------------------------------------------------------------------------------------------------------	------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

## **A Note on Training Coders**

Once researchers move from unsupervised learning to supervised learning, they deliberately frame their analysis in an ideology. It is important to understand that in the training of these models, the accuracy is only as good as the coding, and the models *will* adopt ideological inclinations. As researchers prepare their studies and train their coders, I recommend they turn to the literature of writing assessment. I suggest this because using text as data with human coders draws from principles of writing assessment in the classroom. The processes are similar: researchers are asking coders to make judgements about writing just as a writing instructor may do with a student's essay. In his article, Elliot (2016) reasons through a detailed analysis of ethical philosophies that all writing assessment must be founded on the principle of fairness. He maintains that the only way writing assessment can be ethical is if it is fair, and the only way for writing assessment to be fair is to ask the assessors to understand the philosophical analyses associated with fairness. He claims that the aim of such an approach is to reduce harm to unrepeatable, unique individuals, which is a primary goal of everything discussed up to this point. Elliot's proposed method of moving forward is not to necessarily ask writing instructors to articulate philosophical theories, but they ought to have concepts like ethics, fairness, and compassion at the core of their writing assessment—or in this case coding content.

While Elliot encourages these conceptual awarenesses, Inoue (2015) takes them one step further in his book, *Antiracist Writing Assessment Ecologies*. Inoue makes a compelling argument regarding the embedded biases of racism, sexism, genderism, and all other forms of marginalization that exist in the practices of writing assessment. While

his entire book is dedicated to acknowledging and reducing these racial biases, researchers using institutional genre analysis might take a page or two from his work to articulate and deconstruct those biases. In terms of a big data analysis, the same racist, sexist, and weapons of marginalization are manifested in the algorithms and data collections (as emphasized by O’Neil). Inoue calls for teachers to continuously theorize and practice writing assessment simultaneously, calling back on Elliot’s argument for an informed theoretical and philosophical understanding of fairness. The take-away from these two works is the necessity for a recursive practice of training coders. Using trial and error to refine the coding instructions until a fair and ethical standard is set forth for the particular study, and an understanding that the “final” result of that coder training is unique to that study—not one to be used and applied to another one without extensive revision situated in context.

### **Conclusion**

Big data itself has been hailed as *the* method to tackle the wicked problems of the world simply because it has the capacity to look at *everything*. As researchers begin to adopt these methods, they ought to be aware of the theories, tools, technologies, archives, datasets, ideologies, and types of data that shape the study before results are even generated. While being aware of these issues will improve the quality of the study, researchers ought to consider the deeper, ontological issues revealed from a Cavarerian reading of Benjamin’s aura. It is this fracturing of the algorithmic identity that fuels the intention of an institutional genre analysis—an intention meant to preserve human dignity and support intellectual guerilla warfare conducted by marginalized individuals (Vitanza,

1987). We are surrounded by institutions with thousands of genres shaping our society and our lives; institutional genre analysis disrupts that shaping, making room for marginalized voices to speak and be heard.

## References

- Agboka, G. (2012). Liberating intercultural technical communication from “large culture” ideologies: Constructing culture discursively. *Journal of Technical Writing and Communication*, 42(2), 159–181. <https://doi.org/10.2190/TW.42.2.e>
- Albers, M. J. (2014). Complex problem solving and content analysis. In M. J. Albers & M. B. Mazur (Eds.), *Content and complexity: Information design in technical communication* (pp. 263–283). Routledge.
- Benjamin, W. (1936). The work of art in the age of mechanical reproduction. Retrieved March 18, 2018, from <https://www.marxists.org/reference/subject/philosophy/works/ge/benjamin.htm>
- Blyler, N. R. (1995). Research as ideology in professional communication. *Technical Communication Quarterly*, 4(3), 285–313. <https://doi.org/10.1080/10572259509364602>
- boyd, danah, & Crawford, K. (2012). Critical questions for big data. *Information, Communication & Society*, 15(5), 662–679. <https://doi.org/10.1080/1369118X.2012.678878>
- Cavarero, A. (1995). *In spite of Plato: A feminist rewriting of ancient philosophy*. Taylor & Francis.
- Cavarero, A. (2000). *Relating narratives: Storytelling and selfhood*. London ; New York: Routledge.
- Cavarero, A. (2016). *Inclinations: A critique of rectitude*. (A. Minervini & A. Sitze, Trans.). Stanford, California: Stanford University Press.
- Colton, J. S., Holmes, S., & Walwema, J. (2017). From NoobGuides to #OpKKK: Ethics of Anonymous’ tactical technical communication. *Technical Communication Quarterly*, 26(1), 59–75. <https://doi.org/10.1080/10572252.2016.1257743>
- Douglas, M. (1986). *How institutions think*. Syracuse University Press.
- Elliot, N. (2016). A theory of ethics for writing assessment. *The Journal of Writing Assessment*, 9(1). Retrieved from <http://journalofwritingassessment.org/article.php?article=98>
- Fairfield, J., & Shtein, H. (2014). Big data, big problems: Emerging issues in the ethics of data science and journalism. *Journal of Mass Media Ethics*, 29(1), 38–51. <https://doi.org/10.1080/08900523.2014.863126>

- Feenberg, A. (2002). *Transforming technology: A critical theory revisited*. New York, N.Y: Oxford University Press. Retrieved from <http://libproxy.clemson.edu/login?url=http://site.ebrary.com/lib/clemson/Doc?id=10269075>
- Foucault, M. (1977). *Discipline and punish: The birth of the prison* (1st American ed). New York: Pantheon Books.
- Halavais, A. (2015). Bigger sociological imaginations: Framing big social data theory and methods. *Information, Communication & Society*, 18(5), 583–594. <https://doi.org/10.1080/1369118X.2015.1008543>
- Inoue, A. B. (2015). *Antiracist writing assessment ecologies: Teaching and assessing writing for a socially just future*. Fort Collins, Colorado : Anderson, South Carolina: The WAC Clearinghouse ; Parlor Press.
- Jacobsen, R. (2013, April 24). 2.5 quintillion bytes of data created every day. Retrieved March 18, 2018, from <https://www.ibm.com/blogs/insights-on-business/consumer-products/2-5-quintillion-bytes-of-data-created-every-day-how-does-cpg-retail-manage-it/>
- Jones, N. N., Moore, K. R., & Walton, R. (2016). Disrupting the past to disrupt the future: An antenarrative of technical communication. *Technical Communication Quarterly*, 25(4), 211–229. <https://doi.org/10.1080/10572252.2016.1224655>
- Lewis, S. C., & Westlund, O. (2015). Big data and journalism: Epistemology, expertise, economics, and ethics. *Digital Journalism*, 3(3), 447–466. <https://doi.org/10.1080/21670811.2014.976418>
- Longo, B. (1998). An approach for applying cultural study theory to technical writing research. *Technical Communication Quarterly*, 7(1), 53–73. <https://doi.org/10.1080/10572259809364617>
- Mahrt, M., & Scharkow, M. (2013). The value of big data in digital media research. *Journal of Broadcasting & Electronic Media*, 57(1), 20–33. <https://doi.org/10.1080/08838151.2012.761700>
- Meng, M. (2016, October). *The imperialism of historie*. Presented at the Research Forum, Clemson University.
- Miller, C. R. (1984). Genre as social action. *Quarterly Journal of Speech*, 70(2), 151–167. <https://doi.org/10.1080/00335638409383686>

- Nelson, L. K. (2017). Computational grounded theory: A methodological framework. *Sociological Methods & Research*, 49(124117729703).  
<https://doi.org/10.1177/0049124117729703>
- O'Neil, C. (2016). *Weapons of math destruction: how big data increases inequality and threatens democracy* (First edition). New York: Crown.
- Porter, J. E. (2013). How can rhetoric theory inform the practice of technical communication? In J. Johnson-Eilola & S. A. Selber (Eds.), *Solving problems in technical communication* (pp. 125–145). Chicago, IL: University of Chicago Press.
- Rittel, H. W. J., & Webber, M. M. (1973). Dilemmas in a general theory of planning. *Policy Sciences*, 4, 155–169.
- Rivers, N. A., & Weber, R. P. (2011). Ecological, pedagogical, public rhetoric. *College Composition and Communication*, 63(2), 187–218.
- Ryan, K. J., Brady, J. V., Cooke, R. E., Height, D. I., Jonsen, A. R., King, P., Turtle, R. H. (1979). *The Belmont report: Ethical principles and guidelines for the protection of human subjects of research*. The National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research. Retrieved from [https://www.hhs.gov/ohrp/sites/default/files/the-belmont-report-508c\\_FINAL.pdf](https://www.hhs.gov/ohrp/sites/default/files/the-belmont-report-508c_FINAL.pdf)
- SINTEF. (n.d.). Big data, for better or worse: 90% of world's data generated over last two years. *ScienceDaily*. Retrieved from <https://www.sciencedaily.com/releases/2013/05/130522085217.htm>
- Spinuzzi, C. (2003). *Tracing genres through organizations: A sociocultural approach to information design*. MIT Press.
- Spinuzzi, C. (2005). Lost in the translation: Shifting claims in the migration of a research technique. *Technical Communication Quarterly*, 14(4), 411–446.  
[https://doi.org/10.1207/s15427625tcq1404\\_3](https://doi.org/10.1207/s15427625tcq1404_3)
- Stephens, E. J. (2017). Doing big data: Considering the consequences of writing analytics. *Journal of Writing Analytics*, 1(0). Retrieved from <https://journals.colostate.edu/analytics/article/view/110>
- Swarts, J. (2013). How can work tools shape and organize technical communication? In J. Johnson-Eilola & S. A. Selber (Eds.), *Solving problems in technical communication* (pp. 146–164). Chicago, IL: University of Chicago Press.



- Vitanza, V. J. (1987). Critical sub/versions of the history of philosophical rhetoric. *Rhetoric Review*, 6(1), 41–66.
- Wakefield, J. (2015, November 26). The truth about Facebook quizzes. Retrieved March 18, 2018, from <http://www.bbc.com/news/technology-34922029>
- Walton, R., Zraly, M., & Mugengana, J. P. (2015). Values and validity: Navigating messiness in a community-based research project in Rwanda. *Technical Communication Quarterly*, 24(1), 45–69.  
<https://doi.org/10.1080/10572252.2015.975962>

## ARTICLE 3

### RHETORICAL CONVERSATIONS AND RADICAL COLLABORATIONS: BUILDING PORTABLE, DURABLE RESEARCH IN TECHNICAL AND PROFESSIONAL COMMUNICATION

#### **Introduction**

Technical and professional communication (TPC) has long been a field of practice (Connors, 1999; Moran, 1985; O’Hara, 2001; Tebeaux, 1999; Brockman, 1998; Rutter, 1991) but only recently become a field of study (Pringle & Williams, 2005; Longo, 2000). Today, TPC is a thriving field with several journals, academic and professional conferences, thousands of active researchers, and countless practitioners across the globe. As a field, TPC faces a troubling problem manifested by the editors’ of this particular special issue: “Our journals are replete with insights about effective approaches to technical and scientific communication practice. Little of this research, however, seems to have affected public discourse” (Graham & St. Amant, 2017). In their call for papers, Graham and St. Amant suggest this leads to a crucial question: “How can our research more effectively engage (in) these broader societal conversations?” Before answering this question, however, we must understand *why* this question exists. The depressing answer rests in Virilio’s (2007) claim: “To invent the sailing ship or the steamer is to invent the shipwreck. To invent the train is to invent the . . . derailment. To invent the [car] is to produce the pile-up on the highway” (p. 10). In the act of building and establishing TPC as a field, we have isolated ourselves from those broader societal conversations. To reiterate, we have isolated ourselves from those *conversations*. Do we engage with those societal concerns as a field? Absolutely. Are we engaging with society

about those concerns? In the classroom, yes, but on the research or public discourse levels? We're not so sure. This is the greatest irony of this special issue. As a field, we claim to know the nuances of communication practices in technical, professional, and scientific discourses. We preach rhetorical literacies, emphasizing the importance of working with and for our audiences. But we can't get people outside of our discipline to listen? Everything in this article stems from this important premise: To be durable and to be portable, our research must be rhetorical in its most classic, traditional sense—speaker, message, and audience.

In 1988 at the Conference on College Composition and Communication, eight of rhetorics' top scholars gathered together for a panel focused on the directions and perspectives of their field. This panel, known as the first of three Octalogs, set out to establish rhetoric and composition as a legitimate field worthy of attention, scholarship, and (hopefully) funding—not unlike TPC. This is academic politics. In the conversation, however, Vitanza makes a bold proclamation: “I don't think that we have to validate what we're doing. We are [2,500 years] old or more. . . . We are not a discipline. We are a meta-discipline. . . . We inform all the other disciplines” (p. 31). Technical and professional communication is such a meta-discipline! Research, generally speaking, has three prongs: theory, methods, and artifacts. Nearly every study in the field of the humanities and social sciences shares theories and methods, but it is the artifacts that separate disciplines—and the artifacts of interest to TPC scholars reside *in other disciplines*. Our sites of study bleed from every field *because* we are a meta-discipline. If TPC scholars wish to engage in those societal conversations, then researchers must

actually *converse*. To build this durable, portable research grounded on the premise of rhetorical conversation, TPC researchers might turn to radical collaboration, which is parsed into four types: interdisciplinary work, augmented literature reviews, crowdsourced coding, and publication practices. After defining the characteristics of durable, portable research as rhetorical conversation, we outline the four types of radical collaboration, each with a brief review of literature and examples from various fields of study. We conclude with what a single study might look like if TPC researchers enact all four types of radical collaboration. The source of our artifacts? America's prisons.

### **Defining Durable, Portable Research as Rhetorical Conversations**

At a recent conference round-table on preparing graduate students for research, a table of scholars were discussing how some of their graduate students seem to be intimidated by research methods, even with several course offerings. To ease their anxieties, one professor suggested that students ask a familiar question as they consider methodology: "What do I want to learn from this project?" This is an important question, especially for students beginning to develop their own academic identities, but it's lacking in context. If we consider this question as part of a rhetorical conversation, we see that it focuses on a message as it helps a scholar to define themselves as a speaker, but there is little attention being paid to the audience. Established scholars and new students would both benefit from asking, "What do I want to do with what I learn from this project?" Only by considering what we want our research to do will we be able to begin building durable, portable research. Rather than take the time to define what we mean by "durable" and "portable" as individual terms, we define them in tandem as they

pertain to two aspects of the rhetorical conversation: the speaker and the audience. The article (and the special issue) as a whole engages with what it means to have a durable, portable message, which is why we only focus on the speaker and audience in this section.

### **Speaker**

In their article, Grant-Davie, Matheson, and Stephens (2017) describe several ways how graduate students and early-career academics might articulate who they are as scholars. One particular suggestion is worth repeating here: finding a vocation with TPC, which “means both work and a personal calling” (p. 160). Part of discovering who we are as academics means asking ourselves why we want to do this work in the first place. Finding—and sometimes having to remember—our passions is worth the time it takes to articulate. To be frank, the field of TPC has no need to justify itself or its work to others because it’s already everywhere. Individual scholars will be the ones to bring the work of TPC to the attention of other disciplines and public discourses, and that means those individual scholars ought to know who they are, what they do, and why they do it. If an individual speaker cannot answer for themselves why they do what they do, how can they expect to explain to someone else why they should listen to them?

Before scholars consider how to make their own research more durable or portable, their own identities ought to be durable and portable as well. According to Takayoshi, Tomlinson, & Castillo (2012), who we are as scholars shapes how we see problems or even which problems we see:

If researchers are to be in control of their research practice, it's crucial to explore (and understand) the roles our epistemological, political, and ideological assumptions and commitments, as well as our experiences and knowledge, play in the shaping of our problems and questions. (p. 98)

A researcher's identity and ideology are always-already intertwined, and owning this interconnectedness promotes durability and portability. Jones, Moore, & Walton (2016) discuss six narrative threads in TPC as they relate to identity: feminism, race and ethnicity, international/intercultural professional communication, community and public engagement, user advocacy, and disability and accessibility. In today's global academic environment, a researcher ought to slow down and consider their own narratives and whether they implicitly or explicitly promote a dominant narrative that marginalizes individuals or whether they promote those antenarratives discussed by Jones, Moore, & Walton. As scholars discover how they fit into the larger conversations (knowing when and even if they should speak), they will be able to navigate those conversations more adeptly, providing the portability that research needs to be in order to be heard. This is self-reflective process should be practiced *constantly*.

### **Audience**

In their CFP for this special issue, Graham and St. Amant (2017) discuss the "epistemological anxieties" that TPC has concerning "constructs like validity, reliability, and replicability." They go onto write, "Interestingly, research from our own disciplines indicates such constructs result in findings that carry greater caché." This, of course, seems obvious—a basic tenant of audience awareness even. If a researcher's goal is to

communicate their findings to a group of people, then the researcher should do what they can to appeal to their audience. If your audience gives weight to storytelling, tell a story. If they respect those constructs like validity, reliability, and replicability, then it makes sense that the studies who embody those constructs will be better received than others. Essentially, if a researcher wants to do durable, portable work, then they must first consider where and with whom they want this work to be durable and portable, then they should conduct an audience analysis and figure out how their audience defines durable and portable research, and then *do* that.

In their seminal article on audience awareness in composition theory and pedagogy, Ede and Lunsford (1984) critique the two categories of audience: “addressed” and “invoked.” They write: “The ‘addressed’ audience refers to those actual or real-life people who read a discourse, while the ‘invoked’ audience refers to the audience called up or imagined by the writer” (p. 156). These principles go beyond pedagogy. We teach these things to our students with an expectation that they will be able to understand the needs and wants of their audiences. Shouldn’t we enact the practice in our own writing? Ede and Lunsford believe we do: “Writers who wish to be read must often adapt their discourse to meet the needs and expectations of an addressed audience” (p. 165). In the work we do as a field, we can certainly draw attention to the constructedness of validity, reliability, and replicability, but we need not abandon them. If that is what an audience expects, then a researcher who meets those expectations is much more likely to be included in the conversation. In the words of Ede and Lunsford: “To ignore or devalue such a central function [of audience] is to risk distorting the writing process as a whole. . .

. [W]riters create readers and readers create writers. In the meeting of these two lies meaning, lies communication” (p. 169), to which we add, “lies conversation, lies durability, and lies portability.”

### **Defining Radical Collaboration**

Collaboration is nothing new to TPC. The nature of TPC work, which is workplace communication, is necessarily collaborative. TPC researchers enter workplace environments, talk with people, study their work, and then write about it. Collaboration exists at nearly every level of research from idea brainstorming, to research design, to conducting the studies, and writing results and findings. As a field, TPC embraces collaboration. Burnett, Cooper, and Welhausen (2013) define it “as an intentional, sustained interaction toward a common goal” (p. 454), which this article builds upon here. We use “radical” in two senses of the word: first, it is different or unexpected, and second, it promotes change. As mentioned previously, TPC is a meta-discipline that informs every other discipline and industry, and this positioning enables researchers to insert themselves into the conversation. For many disciplines, getting into the conversation is the hardest part. As indicated by this special issue, TPC’s issue is staying in the conversation. By reconceptualizing collaboration, researchers will be more prepared to stay engaged in those societal conversations. In this section, we outline the four types of radical collaboration, answering three questions for each type: 1) What is it?; 2) How does it promote durability and portability?; and 3) What does it look like?



## **Interdisciplinary Work**

**What is it?** Given the placement of the artifacts of interest to TPC, the field is and has been always-already interdisciplinary. While TPC has embraced collaboration, doing interdisciplinary work means more than having collaborators. It involves re-seeing our own roles in research and discourse. In making his proclamation about rhetorics as a meta-discipline, Vitanza says, “We inform all the other disciplines. They do not inform us” (Octalog, 1988, p. 31). Being a meta-discipline is not, however, omnipotence. TPC informs all the other disciplines, but it is also informed by them. In a study on how rhetoric of science work was received by practicing science studies scholars, Ceccarelli (2005) writes: “rhetoricians of science should acknowledge the rhetorical contributions of non-rhetoricians and negotiate a shared space rather than attempt to fill perceived lacunae in the literature” (p. 257). We should approach these research problems and questions in a way to improve mutual understanding for both TPC and the disciplines with which we hope to engage. Being a TPC scholar doing research in another discipline isn’t necessarily doing interdisciplinary work—to do interdisciplinary work means working with scholars from other disciplines. Cagle and Tillery (2015) contend that doing this type of interdisciplinary work leads to three distinct advantages: “(a) a broader context for our existing research; (b) new avenues for our future research; and (c) awareness of close overlaps between other fields’ scholarship and our own” (p. 147).

**How does it promote durability and portability?** The core question for this special issue is not, “Why don’t TPC scholars read other TPC scholarship?” We read each other’s work, attend each other’s conference panels, and encourage each other in a

myriad of different ways. We are invested in each other. This is durability. To do interdisciplinary work is to include other fields with that same passion; what better way to do so than to include *individuals*. Don't engage with the field in the abstract sense, engage with individuals from that field. Talk with them. Write with them. Research with them. Publish with them. This is portability. Doing this interdisciplinary work situates the TPC researcher as a guest rather than a trespasser, avoiding what Ceccarelli calls "disciplinary politics" (p. 258).

Ceccarelli, quite boldly, reprinted some of the critique from those scientists, who wrote that her book was a "flagrant violation of etiquette," a "hotch-potch," a "hatchet-job," an "uncomfortable example," and this gem: "Latecomers, such as literature of science and rhetoric of science, have been regarded, understandably, as interlopers rather than as contributors to the conversation" (p. 258). Interdisciplinary work steeped in respect may help TPC researchers to avoid receiving a similar reception as Ceccarelli's book.

**What does it look like?** For a prime example of interdisciplinary work, see Walton, Zraly, and Mengengana's (2015), "Value and Validity: Navigating Messiness in a Community-Based Research Project in Rwanda."

### **Augmented Literature Review**

**What is it?** Texts have always been a site of study for TPC scholars, and with the rise of big data in nearly every field, it is only a matter of time before using large-scale texts as data will take its place firmly as a TPC methodology as well (see Graham, Kim DeVasto, and Keith, 2013). According to Grimmer (2015), "big data provides the

opportunity to learn about quantities that were infeasible only a few years ago” (p. 80). What quantities exactly? In 1992, Katz performed a close rhetorical reading of a memo written by a Nazi German officer by the name of Just. This is possible with only one memo, perhaps a small stack even. A few years ago, if a scholar came across a database of 5,000 Nazi memos it would be nearly impossible to identify any common thread let alone read all of them. With text as data, this becomes a possibility: “the systematic analysis of large-scale text collections without massive funding support” (Grimmer and Stewart, 2013, p. 268). In her article, “Computational Grounded Theory,” Nelson (2017) develops a three-step process to using text as data:

The first, pattern detection step, involves inductive computational exploration of text, using techniques such as unsupervised machine learning and word scores to help researchers to see novel patterns in their data. The second, pattern refinement step, returns to an interpretive engagement with the data through qualitative deep reading or further exploration of the data. The third, pattern confirmation step, assesses the inductively identified patterns using further computational and natural language processing techniques. The result is an efficient, rigorous, and fully reproducible computational grounded theory. (p. 1)

What makes this radical collaboration? The source of the data. In a traditional research study, authors will engage with the field by performing a literature review. Often this is how TPC scholars will establish themselves in the conversation. But rather than engage with 5, or 10, or even 15 articles and their authors, why not engage with all of them simultaneously with machine learning?

Nelson's first step of pattern detection can be used on an entire journal—every article they've ever published that has been digitized and optimized. This step will reveal common themes of the field as perceived by its journals, providing TPC researchers with statistical representations of what the field has said and how it has developed across time and space. In doing so, TPC researchers can situate their work in the other field using the four *topoi* described by Grant-Davie, Matheson, and Stephens (2017): a problem that needs correcting, identifying a knowledge gap, presenting a new perspective, or the assertion of a new complication (p. 156). The idea of big data is already compelling, providing some of the most unique set of research problems and questions in recent methodological history, and while big data has its flaws, it can still be used as a way to draw attention to the *topoi* with which TPC scholars want to engage. If rhetoric can be defined “as the faculty of observing in any given case the available means of persuasion” (Aristotle, 2004), then why not turn to machine learning to help make those observations? The academic database JSTOR has a division called Data for Research (DfR) where researchers may submit a request for these journals: “Datasets are produced at no cost to researchers and may include data for up to 25,000 documents.” All you need to do is make the request.

**How does it promote durability and portability?** Building research that will be involved in other disciplines and public discourses begins with engaging as much as possible with those disciplines and discourses. While interdisciplinary work focuses on working with the individuals with the field, this method provides researchers with the means to engage with the field as a whole. Grimmer and Stewart (2013) caution that

language is complex and using machine learning to interact with the text cannot and should not replace close readings. “Rather,” they continue, “the methods that we profile here are best thought of as *amplifying* or *augmenting* careful reading and thoughtful analysis” (p. 268; original emphasis). This augmented literature review gives the home discipline the weight a researcher thinks their traditional literature review gives. Instead of building bridges to a few scholars, an augmented literature review builds bridges with all of them. Once those patterns have been identified, TPC researchers can use those insights to inform their study, suggesting to the audience that the research questions were not designed as a critique of the field but grew from their own work. An augmented literature reviews allows a researcher to give as much respect and homage as possible before engaging. Think Rogerian argumentation.

**What does it look like?** To see the inspiration for this type of radical collaboration, see chapter 2 of Grimmer’s (2013) *Representational Style in Congress: What Legislators Say and Why It Matters*.

### **Crowd-sourced Coding**

**What is it?** In step three of “Computational Grounded Theory,” Nelson (2017) suggests researchers use those themes to begin conducting the pattern confirmation step, which is testing a theory. This is essentially content analysis on a large-scale. Content analysis, however, can be time consuming for a small dataset, and that time is exponentially magnified if your dataset is in the tens of thousands of pages. Here it may be helpful to describe the process of machine learning:

1. Build the database and parse the language.

2. Develop research questions.
3. Create instructions for what to code for the content analysis.
4. Find some coders
5. Instruct coders using the codebook.
6. Coders then read/review a dataset, marking each unit with a value.
7. Determine how well the coders agreed on the value of units, which is known as inter-coder reliability.
8. Feed coded content in the learning model.
9. The machine then attempts to replicate the results from the coder, but the machine is about half as “smart.”
10. As the machine iterates through “epochs,” its coding reliability increases significantly.
11. Once the machine reaches a pre-defined cutoff, the researcher “tests” the machine on a set of data that the machine has never seen before. These results can be compared to those of a coder to assess the algorithm’s efficacy.
12. If the researcher is satisfied with the results, the algorithm is complete. If not, the researcher can “tune” the parameters of the model to improve results and repeat the training and cross-validation until the machine is as effective as necessary.

One immediate problem that TPC researchers will be quick to notice is that the machine will also learn and replicate the ideologies of their human coders, which isn’t necessarily bad if acknowledged, but is still a problem, calling back to those epistemological anxieties. Rather than turn to friends or research assistants for help, we recommend

researchers turn to crowd-sourced coding. The most popular of which is Amazon's Mechanical Turk, or MTurk. According to a Pew case study:

The Mechanical Turk website was the idea of Amazon chief executive Jeff Bezos, who believed a platform could be created to exploit the fact that humans can easily perform certain tasks that were difficult for computers. He predicted there was a business to be built around connecting those who wanted research done with those who were willing to do it. By creating the Mechanical Turk marketplace, Bezos tried to create a phenomenon he called "artificial artificial intelligence." (Hitlin, 2016)

Some of those tasks performed better by humans than machines? Content analysis. The process is the similar to a traditional coding: researchers give instructions for a Human Intelligence Task (HIT) that asks the "Turker" to code the sentence. That's it. They can just do one or they can do hundreds. How well the HIT is designed and how much it pays will dictate the rate of completion and the quality of the data. While some may be hesitant to the effectiveness of such an approach, Lind, Gruber, and Boomgaarden (2017) argue that crowd-sourced content analysis is a viable option with benefits that greatly outweigh the disadvantages: "Crowdsourcing is a very efficient and cost-effective tool for the production of quantitative data, in particular with regards to contents that are harder to pre-define in dictionaries, i.e., less manifest, protectively latent constructs" (p. 2). While coders cannot be individually selected, this actually removes a risk of finding individual coders who might be too similar, which would leave room for the reproduction of ideological biases.

**How does it promote durability and portability?** According to Benoit, Conway, Lauderdale, Laver, and Mikhaylov (2016), it has become a best practice for journals to ask authors to provide their data and the codebook in an open access venue. They reason that “this allows other scholars to replicate and extend published results by reanalyzing the data, rerunning and modifying the code. Replication of an *analysis*, however, sets a far weaker standard than reproducibility of the *data*” (p. 278). A common critique from quantitative researchers about qualitative research is this idea that an analysis is difficult to replicate, and knowing that our potential audiences value the ability to replicate the *analysis* may ease those tensions (even if no one ever really does replicate it since there isn’t a market for replicated studies). With crowd-sourced content analysis, those results will actually be easier to replicate as described by Benoit, Conway, Lauderdale, Laver, and Mikhaylov (2016): “With just hours from deployment to dataset, and for very little cost, crowdsourcing enabled us to generate externally valid and reproducible data related to our precise research question” (p. 290).

**What does it look like?** As cited above, we recommend you turn to Benoit, Conway, Lauderdale, Laver, and Mikhaylov’s (2016) “Crowd-sourced Text Analysis: Reproducible and Agile Production of Political Data.”

### **Publication Practices**

Rather than divide this section into the three questions as the previous types of collaboration, we want to take this small section to get real for a quick moment. While augmented literature reviews and crowd-sourced coding sound fancier and more radical than publication practices, it may not be the case. If we want to engage in other



disciplines and public discourse, then we ought to publish there, and sometimes that means *not* publishing in our own journals. At a recent job interview, Eric was asked where he'd like his work to be published. Being an interdisciplinary scholar, he listed off several journals from computational methodologies, prison studies, and popular culture. The search committee asked what he wants his "spine" of work to be, implying that it should be technical and professional communication. Why do we ascribe so much of our identity as scholars to the journals we publish in and not the content we write? Publishing our research in the journals of other disciplines carries an enormous advantage: it shows we want to have those conversations with those disciplines and public discourses. It shows us reaching out to our audiences! Why should we expect other disciplines and the public discourse to come to us?! Are we that arrogant? If we truly wish to be involved in those rhetorical, societal conversations, then we ought to walk over and introduce ourselves.

### **TPC in Prisons: An Example of Radical Collaboration in Action**

With the concepts of rhetorical conversations and radical collaborations established, what would a study look like that enacts all four types of radical collaboration? For this portion of the article, we shift into a narrative-form from Eric's perspective: a TPC scholar who identifies as an interdisciplinary researcher looking to enter the fields of data science and prison studies.

According to the ACLU, the United States houses 25% of the world's prison population, a nearly 700% growth since 1970. This epidemic is what Rittel and Webber (1973) would call a "wicked problem": one so messy and ill-defined that no single, linear

solution exists (or any solution at all). As a TPC scholar, I designed this project using big data methodologies to understand correctional practices in America's prison system manifested through genres of technical and professional communication. Treating text as a data source, the project focuses on a single genre—inmate handbooks. I collaborate with industry data scientists, Ben and Katie, to understand the conventions of this particular genre, its themes, and its impacts on “correction.” These facilities claim to correct or rehabilitate individuals to what is “normal,” but what is normal is (and should be) contested. Through machine learning and computational content analysis, this study creates a continuum of punishment and reward at the sentence level; doing so adds nuanced understanding to Foucault's theory of normalization. For example, are family visits framed as a reward for good behavior? or are they taken away for bad behavior? What follows is how we approached the study as a rhetorical conversation using all four types of radical collaboration.

### **The Rhetorical Conversation**

This project was born from a societal conversation. As I watched John Oliver's (2014) report on the state of America's prisons on *Last Week Tonight*, I realized that this was a conversation that I could be a part of as a TPC scholar. We often hear things like “institutional racism” or “institutional sexism,” but pointing to individual stories isn't the best way to enter the conversation if you hope to change people's minds. Someone's story, while tragic, can easily be dismissed as an anecdotal fallacy. It was just “one bad apple.” I knew I wanted my work to matter, to impact change, so I couldn't do a close analysis of just one document or series of documents from the local prison. If I performed

a close analysis of the Perry Correctional Institution in Pelzer, SC, my analysis could easily be dismissed by a policy maker in Oregon. They might say, “Well, that’s the South, not Portland.” To prevent the accusation of an anecdotal fallacy, I turned to big data to enter the conversation in a way that would be heard.

### **Interdisciplinary Work**

Knowing that I wanted to engage with the public discourse and policy makers drove my decision to turn to big data, but it didn’t mean I suddenly knew how to do it. I talked to my friend Ben who had been working as a data scientist for NLP Logix in Jacksonville, FL. He thought the project sounded interesting and decided to help, and when he told his co-worker, Katie, about it, she joined as well. Thank goodness. Working with data scientists has been an invaluable experience, and it allows me to be confident in the actual process and coding. The last thing I wanted was to go to a conference and have a “math person” ask me a math question I don’t know the math answer to. If they asked today, I still couldn’t answer it, but at least I can say, “I’ll turn that question to Ben and Katie.” Having someone from the field of big data building the algorithms and code to answer those questions gives my audience more reason to listen.

The first step was to build the archives: one for the inmate handbooks and one for the journals from prison studies. Many of these handbooks are available at an individual prison’s website, but I found the majority of them through a targeted Google search. I collected 347 handbooks from across the country, totaling 15,719 pages and over 425,000 individual n-grams. The python code written by Ben and Katie, which we have made available at the end of the dissertation, prepared the text for the machine learning. With a

single command, the code separates each page of the article into its own PDF, puts the single PDF files into a folder for the article, converts those PDF files into TXT files, scrubs the text, separates each sentence, gives each sentence a random ID, parses the sentences into n-grams, and pulls a random sampling for coding. While we are able to make to code and handbook data public, publishing the data from the journals violates copyright law.

### **Augmented Literature Review**

With the help of a few friends, I also downloaded every article from the last 20 years published by two of prison studies top journals: *The Prison Journal* and *Punishment and Society*. (The articles were not available through JSTOR's DfR, so each article had to be individually downloaded.) In total, there are 936 articles, totaling 17,953 pages. In order to produce the images below, we completed the following steps:

1. Download the documents
2. Convert files from PDF to TXT
3. Parse documents (splitting up individual sentences)
4. Remove non-alphanumeric characters, stop words, etc.
5. Create n-grams (n=1-8) of series of words that are frequently used together in context
6. Create a CSV file of all sentences
7. Train a machine learning model to give a numerical output for each word in the context of the rest of the sentence
8. Use linear algebra to assess similarities among words within the documents

What comes of this process is a visualization of the journals based on themes. These visualizations, coupled with a close reading of specific articles, provides a thorough investigation of the journals' themes. Doing so accomplishes two goals for building durable, portable research: first, it provides the code and dataset to others who are able to replicate the analysis if desired, and 2) it engages with an entire field of study rather than the small percentage of close readings a traditional literature review provides.

Below are screenshots of two dashboards created in Tableau 10.5. After reading through a small sample of articles, I determined two themes to investigate (while giving a nod to Foucault): “discipline” and “punishment.” Machine learning does not know the meanings of words based on a dictionary definition; instead, the learning model searches the parsed n-grams and finds which ones are similar or not similar based on the frequencies and relationships they have together inside the corpus. The learning model can identify synonyms for an n-gram *based on the genre itself*. Each figure is a side-by-side comparison of a theme, which is labeled as “token.” The higher the word is on the list, the stronger the statistical similarity; if the bar is greyed it fell below the .85 average cosine similarity threshold. The numbers to the right of the bars are the n-gram's frequency in the corpus. Looking at the lists of words can help researchers begin to interpret why some of these synonyms exist, informing a much closer reading of the journals themselves while guiding the investigation of the inmate handbooks.

Figure 3.1 Machine learning generated synonyms for “discipline” in two penal studies journals

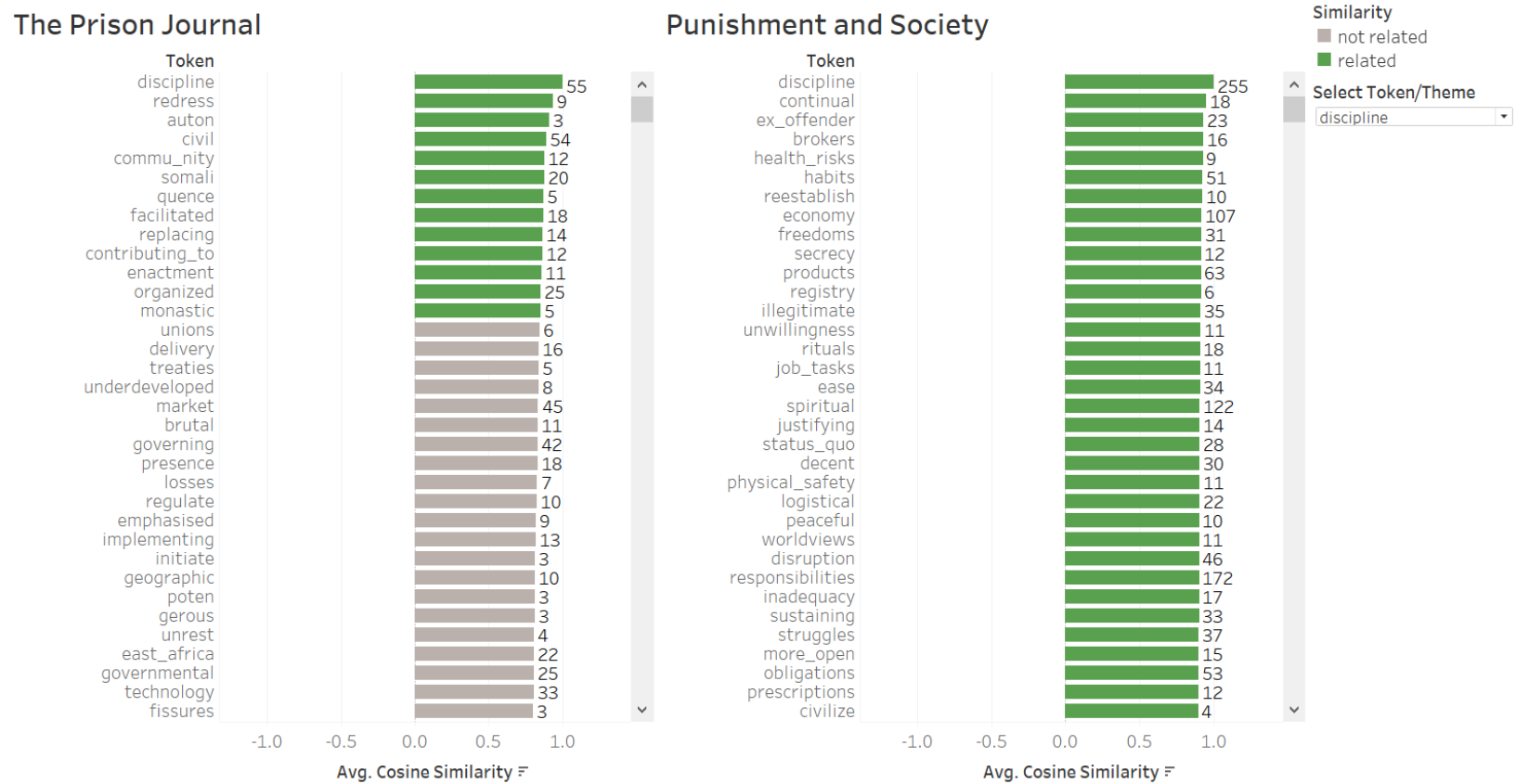
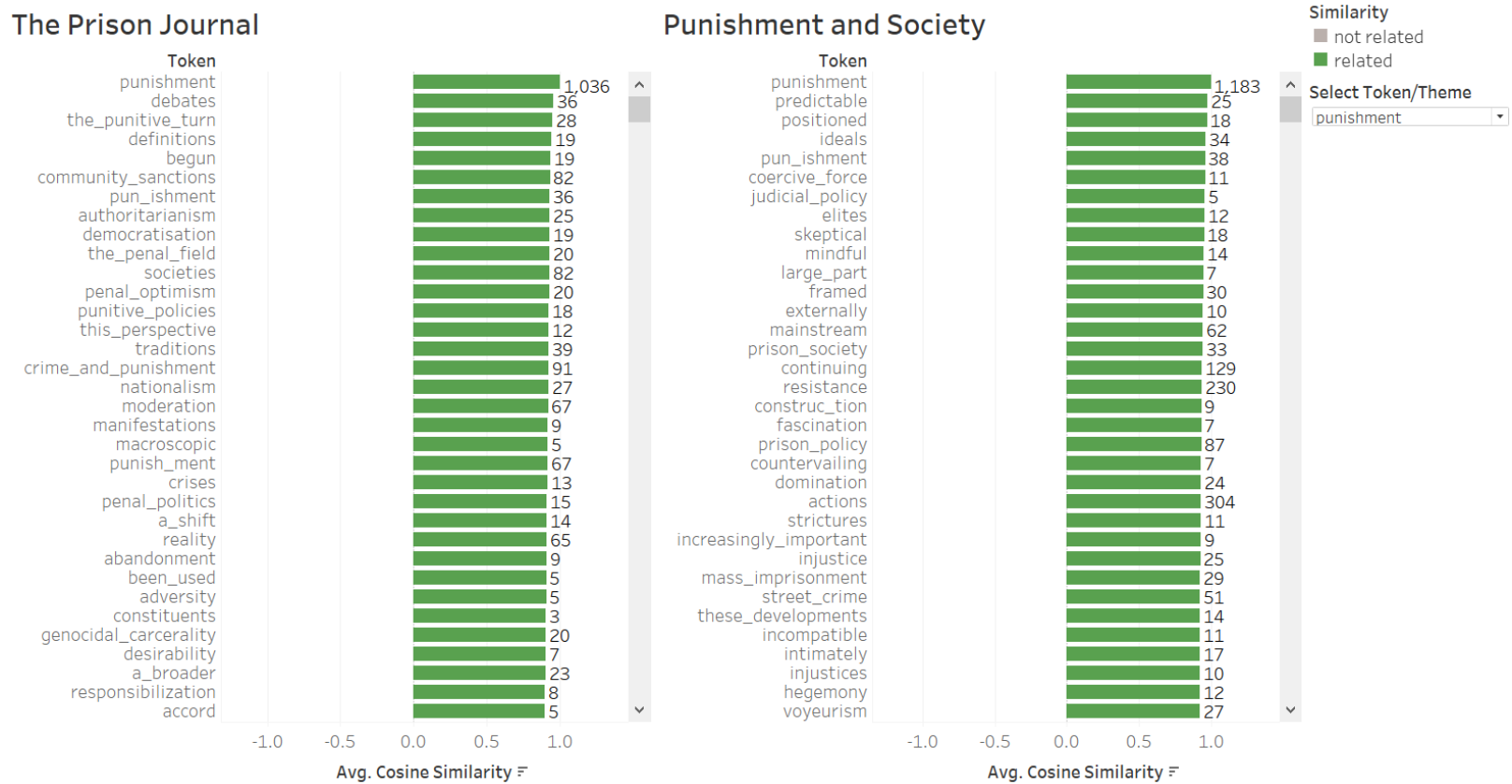


Figure 3.2 Machine learning generated synonyms for “punishment” in two penal studies journals



## Crowd-Sourced Coding

For this stage of the project, I turn to Foucault’s concepts of discipline and punishment. Essentially, we have a single, driving question for this stage of the project: Do inmate handbooks have a statistically significant difference in language that rewards behavior or punishes it? To answer this question, we plan to code the random sampling of sentence units provided by the computer on a scale of 0 to 1. Each unit will be coded with three categories:

Category	Explanation
Punishment	Is there a positive consequence for an action?
Reward	Is there a negative consequence for an action?
Administrative	There are no consequences/actions mentioned

For each sentence unit, coders will measure it against these three categories, a 1 for “yes” and a 0 for “no,” and each unit will be coded by three separate coders, allowing us to measure inter-coder reliability. Below is a small set of sentences as an example.

Punishment	Reward	Administrative	Sentence
1	0	0	Inmates found guilty of violating jail rules and ordered confined in disciplinary confinement will automatically lose all privileges for the duration of said confinement.
0	1	0	Regular inmates (non-trustee) may earn up to a total of eight (8) days per month.
0	0	1	At the point of entry into the facility, you will be asked to share basic personal information.
0	1	0	You have the right to a safe, clean and healthy environment.
0	0	1	The deposit must be in the form of a money order made out to the inmate's full



			committed name and complete eight-digit register number.
1	0	0	It is important for you to know that violations of rules and regulations will be dealt with swiftly and decisively.

As you may recall, there are over 425,000 individual n-grams in the 347 inmate handbooks. To even code 10% of the data means coding 42,500 sentences for three categories, three times each, providing us with 382,500 data points. From that data, the computer will be able to identify how much each sentence leans towards punishment and/or reward.

### **Publication Practices**

The driving question here: “What do I want to do with what I learn from this project?” Where we decide to publish this work will be determined by what we want to do with the work. If we want to establish a new approach to doing research in TPC, then a TPC journal is the perfect fit. If we want to submit this work for the data science side of the interdisciplinary work, then we’ll write up and publish the results in a data science journal. If we want this discussion to enter prison studies, then we’ll submit it to one of the journals from the augmented literature review. If we want to engage with the public discourse, then we can talk with Vox.com or someplace similar. Where we publish should have less to do with how we identify as scholars and more to do with what impact we want our work to have.

### **Concluding Thoughts**

It’s a good time to be a technical and professional communication scholar. Our sites of study are rich with data, rich with implications, and rich with information and

perspectives that can have a positive impact on so many of those societal conversations. With such a strong swing towards taking action as a field in the past two years (see Jones, Moore, and Walton, 2016; Colton, Holmes, and Walwema, 2017), we ought to prepare our work in a way that we can converse with other disciplines and public discourses. Radical collaboration is such a way. Doing work with the *people* from those disciplines and discourses with which we wish to engage. Using machine learning technology to give our audiences the respect they deserve. Being open about ideological influences in our work and exposing it to as many other ideologies as possible to produce data *and* analysis that can be reproduced. Publishing our work in the places our intended audience already reads. This special issue asks one simple, penetrating question: Are you persuasive? Technical and professional communication *is* persuasion. Persuasion *is* rhetoric. We're rhetoricians! Let's be rhetorical.

## References

- American Civil Liberties Union. (2015). *Overcrowding and overuse of imprisonment in the United States* (p. 23). ACLU.
- Aristotle. (2004). *Rhetoric: Book I*. (W. R. Roberts, Trans.). Dover Thrift.
- Benoit, K., Conway, D., Lauderdale, B. E., Laver, M., & Mikhaylov, S. (2016). Crowd-sourced text analysis: Reproducible and agile production of political data. *American Political Science Review*, *110*(2), 278–295. <https://doi.org/10.1017/S0003055416000058>
- Brockmann, R. J. (1998). *From millwrights to shipwrights to the twenty-first century: Explorations in a history of technical communication in the United States*. Hampton Press.
- Cagle, L. E., & Tillery, D. (2015). Climate change research across disciplines: The value and uses of multidisciplinary research reviews for technical communication. *Technical Communication Quarterly*, *24*(2), 147–163. <https://doi.org/10.1080/10572252.2015.1001296>
- Ceccarelli, L. (2005). A hard look at ourselves: A reception study of rhetoric of science. *Technical Communication Quarterly*, *14*(3), 257–265. [https://doi.org/10.1207/s15427625tcq1403\\_3](https://doi.org/10.1207/s15427625tcq1403_3)
- Colton, J. S., Holmes, S., & Walwema, J. (2017). From NoobGuides to #OpKKK: Ethics of Anonymous' tactical technical communication. *Technical Communication Quarterly*, *26*(1), 59–75. <https://doi.org/10.1080/10572252.2016.1257743>
- Connors, R. J. (1999). The rise of technical writing in America. In Kynell & M. G. Moran (Eds.), *Three Keys to the Past: The History of Technical Communication*. Greenwood Publishing Group.
- Data for research (DfR). (2018). Retrieved March 24, 2018, from <https://www.jstor.org/dfr/>
- Ede, L., & Lunsford, A. (1984). Audience addressed/audience invoked: The role of audience in composition theory and pedagogy. *College Composition and Communication*, *35*(2), 155–171. <https://doi.org/10.2307/358093>
- Foucault, M. (1977). *Discipline and punish: The birth of the prison* (1st American ed). New York: Pantheon Books.

- Graham, S. S., Kim, S.-Y., DeVasto, D. M., & Keith, W. (2015). Statistical genre analysis: Toward big data methodologies in technical communication. *Technical Communication Quarterly*, 24(1), 70–104. <https://doi.org/10.1080/10572252.2015.975955>
- Graham, S. S., & St. Amant, K. (2017). CFP: TCQ special issue on durable and portable research in technical and scientific communication. Retrieved March 23, 2018, from <http://sscottgraham.com/tcq-cfp>
- Grant-Davie, K., Matheson, B., & Stephens, E. J. (2017). Helping doctoral students establish long-term identities as technical communication scholars. *Journal of Technical Writing and Communication*, 47(2), 151–171. <https://doi.org/10.1177/0047281617692071>
- Grimmer, J. (2013). *Representational style in congress: What legislators say and why it matters*. Cambridge University Press.
- Grimmer, J. (2015). We are all social scientists now: How big data, machine learning, and causal inference work together. *PS: Political Science & Politics*, 48(1), 80–83. <https://doi.org/10.1017/S1049096514001784>
- Grimmer, J., & Stewart, B. M. (2013). Text as data: The promise and pitfalls of automatic content analysis methods for political texts. *Political Analysis*, 21(3), 267–297.
- Hitlin, P. (2016, July 11). Research in the crowdsourcing age, a case study. Retrieved March 25, 2018, from <http://www.pewinternet.org/2016/07/11/research-in-the-crowdsourcing-age-a-case-study/>
- Jones, N. N., Moore, K. R., & Walton, R. (2016). Disrupting the past to disrupt the future: An antenarrative of technical communication. *Technical Communication Quarterly*, 25(4), 211–229. <https://doi.org/10.1080/10572252.2016.1224655>
- Katz, S. B. (1992). The ethic of expediency: Classical rhetoric, technology, and the Holocaust. *College English*, 54(3), 255–275. <https://doi.org/10.2307/378062>
- Lind, F., Gruber, M., & Boomgaarden, H. G. (2017). Content analysis by the crowd: Assessing the usability of crowdsourcing for coding latent constructs. *Communication Methods and Measures*, 0(0), 1–19. <https://doi.org/10.1080/19312458.2017.1317338>
- Longo, B. (2000). *Spurious coin: A history of science, management, and technical writing*. SUNY Press.

- Moran, M. G. (1985). The history of technical and scientific writing. In *Research in Technical Communication. A Bibliographic Sourcebook*. (pp. 25–38). Greenwood Press, 88 Post Road West, Box 5007, Westport, CT 06881 (\$57.50). Retrieved from <https://eric.ed.gov/?id=ED258201>
- Nelson, L. K. (2017). Computational grounded theory: A methodological framework. *Sociological Methods & Research*, 49(124117729703). <https://doi.org/10.1177/0049124117729703>
- Octalog. (1988). The politics of historiography. *Rhetoric Review*, 7(1), 5–49.
- O’Hara, F. M. (n.d.). A brief history of technical communication. In *STC Annual Conference proceeding* (pp. 500–504). Arlington, VA.
- Pringle, K., & Williams, S. D. (2005). The future is the past: Has technical communication arrived as a profession? *Technical Communication*, 52(3), 361–370.
- Prison: Last Week Tonight with John Oliver*. (2014). HBO. Retrieved from [https://www.youtube.com/watch?v=\\_Pz3syET3DY](https://www.youtube.com/watch?v=_Pz3syET3DY)
- Rutter, R. (1991). History, rhetoric, and humanism: Toward a more comprehensive definition of technical communication. *Journal of Technical Writing and Communication*, 21(2), 133–153. <https://doi.org/10.2190/7BBK-BJYK-AQGB-28GP>
- Takayoshi, P., Tomlinson, E., & Castillo, J. (2012). The construction of research problems and methods. In K. M. Powell & P. Takayoshi (Eds.), *Practicing research in writing studies: Reflexive and ethically responsible research* (pp. 97–121). Hampton Press.
- Tebeaux, E. (1999). Technical writing in seventeenth-century England: The flowering of a tradition. *Journal of Technical Writing and Communication*, 29(3), 209–253. <https://doi.org/10.2190/0ET6-4V6N-KWLE-XJE1>
- Virilio, P. (2007). *The original accident*. (J. Rose, Trans.). Polity.
- Walton, R., Zraly, M., & Mugengana, J. P. (2015). Values and validity: Navigating messiness in a community-based research project in Rwanda. *Technical Communication Quarterly*, 24(1), 45–69. <https://doi.org/10.1080/10572252.2015.975962>

## ARTICLE 4

### FROM AN ETHICS OF CARE TO EXPEDIENCY IN AMERICA'S PRISONS: USING INSTITUTIONAL GENRE ANALYSIS TO UNDERSTAND THE LANGUAGE OF CORRECTIONS IN ADULT INMATE HANDBOOKS

#### **Introduction**

In 1972, for every 100,000 Americans, 161 were incarcerated. In less than 35 years, “that rate had more than quintupled to a peak of 767 per 100,000,” bringing the nation’s jail population to more than 2.23 million, making up 25% of the world’s incarcerated population (National Research Council, 2014, p. 33). As a country, the United States has a problem with prisons. The reasoning for this drastic increase has been attributed to race (Alexander, 2010; Coates, 2015), poverty (Wacquant, 2009), and the shifts in the purposes of prisons themselves (Foucault, 1977). These and other works thoroughly address what has been called by others as the “pipeline to prison” (Wald and Losen, 2003), and while this article does not address *why* people end up in prison, it readily admits there is a problem. Rittel and Webber (1973) would call it a “wicked problem”—one so messy and ill-defined that there is no easy solution if there is even a solution at all. Instead, this article addresses how prisons as an institution see their own roles in the rehabilitation of inmates. These are, after all, “correctional facilities” being run by “correctional officers,” so what are these inmates being corrected *to*? As early as 1958, scholars identified a shift in the purpose of these correctional facilities from rehabilitation to incapacitation of inmates (see Craig, 2004). Some even went as far as to say that “rehabilitation as the primary goal of criminal justice was declared dead, or at least on its way to the grave (Lynch, 2000, p. 40).

If the goal of the correctional facility is no longer rehabilitation but incapacitation, how has this change been manifested? In his work *Ambient Rhetoric*, Rickert (2013) describes the embedded nature of rhetoric as ambient: “Rhetoric has a material dimension, and it is an embodied and embedded practice. Rhetoric is an emergent results of environmentally situated and interactive engagements. . . . To be situated means that one’s emplacement is inseparable from the rhetorical interactions taking place” (p. 34). Which material, embodied, embedded rhetorical practices? In their article, Steen and Bandy (2007) contend that “[t]he rise of retribution as a guiding principle for punishment contributed to the development of, support for, and ultimately passage of state policies” (p. 6). According to Rivers and Weber (2011), one of these ambient, rhetorical practices are the “mundane” artifacts that surround us: “While these mundane documents are not always as exciting or visible as the rhetorical frameworks of more obvious public displays, supporting documents are no less necessary for the creation and re-creation of publics” (p. 188). Which publics? Even though they often aren’t in the public’s eyes, prisons are (mostly) public institutions (at least for now). Which mundane artifacts? Inmate handbooks. Of these documents, Bosworth (2007) writes: “Once translated into the mundane and the banal, they have, in turn, become the values inscribed in, and underpinning, the experience of prison itself” (p. 68). The ambient, mundane artifacts surrounding prisons can be a rich source of data in understanding how the prison institution sees their own role in the rehabilitation and care for inmates—or the rather the lack thereof.

In her analysis of orientation and admission handbooks in the United States at the federal level, Bosworth (2007) argues that “[a]lthough the admission documents cannot show us how penal philosophy and policies are actually implemented, they do represent an official articulation of rules, regulations and *ethos*” (p. 71). Technical and professional communication (TPC) research is a prime location for studying how these mundane artifacts function rhetorically in the construction of the prison industrial complex, but what *ethos* is being communicated in these documents? Drawing on literature from criminal justice and penal studies, I establish prisons as maintaining what Colton, Holmes, and Walwema (2017) call an ethic of care. In the recent turn to retribution rather than rehabilitation, however, prisons have turned to what Katz (1992) calls an ethic of expediency. In order to make this conclusion, this study operationalizes Bosworth’s (2007) close readings of several federal orientation handbooks and uses computational content analysis to understand the systematic nature of these oppressive ideologies in prisons. Rather than a close reading on a select number of inmate handbooks, I collected and organized 347 handbooks from local, county, state, and federal facilities, totaling 15,719 pages and more than 425,000 unique n-grams. Drawing on Nelson’s (2017) computational grounded theory and Stephens’s (2018) institutional genre analysis, this study seeks to perform a similar rhetorical analysis as Katz (1992) in “The Ethic of Expediency”—not on a single document but an entire genre.

## **Genres**

To situate the use of genres to understand what correction has become, I turn to Spinuzzi’s (2003) *Tracing Genres* where he discusses activity theory, which “posits that



in every sphere of activity, collaborators use instruments to transform a particular object with a particular outcome in mind” (p. 37). If we approach the act of correction as an activity, we can begin to see the role of genre in corrections. Spinuzzi describes the *instrument* of an artifact as a *mediating* artifact: “[P]eople use these external instruments to reach some goal, and in the process, the people themselves are psychologically transformed: as they use these external means to regulate themselves, they begin to think and act differently” (p. 38). This study, however, cannot and does not attempt to determine how or even if these genres impact how inmates think or act. Instead, I read them, as Bosworth states above, as “an official articulation of rules, regulations and *ethos*” (p. 71). Bodker (1997) describes artifacts as “crystallized knowledge” (p. 150), which constitutes a genre. According to Spinuzzi (2003), genres are a type of tradition, “not simply text types.” He goes on to write, “they are culturally and historically grounded ways of seeing and conceptualizing reality” (p. 41). Essentially, these inmate handbooks are crystallized practices that reveal the cultural and historical ways prisons perform the activity of correction as a whole institution.

### **An Ethic of Care**

In their article, Colton, Holmes, and Walwema (2017) describe an ethic of care based on the philosophical work of Adrianna Cavarero: “an ethics of care recognizes moral value in the reciprocal and singular relations of caring between individuals that ensures one another’s well-being” (p. 60). At its core, an ethic of care rests on the unrepeatable, unique nature of each individual (Cavarero, 2000, p. 2). Colton, Holmes, and Walwema (2017) go on to explain:

Each and every living human being is always in a vulnerable relation to others, but the degree to which that relation is one of caring or wounding depends upon localized and concrete factors, such as materiality, age, time, space, and power. For Cavarero, wounding and caring do not correspond to a basic binary (e.g., wounding = bad, caring = good). Rather, these terms offer a set of fluid ratios to allow us to characterize the totality of relations. . . . [In] some cases, [those relations] will involve wounding certain individuals to help ensure our collective ability to ensure an ethics of care for the most vulnerable. (p. 60)

Seeing inmates as a vulnerable population may be hard for some. They are, by definition, convicted criminals. Here, it is important to define “vulnerable.” Cavarero argues that all humans are vulnerable in the sense that each of us “is irremediably open to relations of caring or wounding” (Colton, Holmes, and Walema, 2017, p. 60). In this case, “vulnerable” and “marginalized” carry similar meaning, but may be better seen on a continuum as many communities are made more vulnerable or subject to wounding than others. According to Bosworth (2007), inmates may have legal rights to goods and services, “but they have little means of ensuring their delivery. In contrast, the institution has an array of sanctions it may implement when inmates are found not to have upheld their end of the bargain” (p. 73). By definition, inmates are more vulnerable than the non-incarcerated individual. This ethic of care for the individual, at one time, was the driving purpose behind prisons in America.

In her historical perspective on the development of prisons as an institution, Craig (2004) cites the preamble from the Philadelphia Society for Alleviating Miseries in

Public Prisons:

When we reflect upon the miseries [seen in prisons] . . . it becomes us to extend our compassion to that part of mankind, who are the subjects of these miseries. By the aids of humanity, their undue and illegal sufferings may be prevented . . . and such degrees and modes of punishment may be discovered and suggested, as may, instead of continuing habits of vice, become the means of restoring our fellow creatures to virtue and happiness. (Vaux, 1826, as cited in Craig, 2004, p. 93S)

In fact, Philadelphia is home to America's first penitentiary, the Walnut Street Jail.

According to the *Oxford English Dictionary*, the word "penitentiary" literally means "house of penance," and "penance" is the "performance of some act of self-mortification or the undergoing of some penalty as an expression of sorrow for sin or wrongdoing."

This concept of penance has transformed into "rehabilitation," whose definition I take from Lynch (2000): any language or action that indicates an aim to reform the inmate, either psychologically, interpersonally and situationally, more structurally, or some combination (p. 45). Lynch continues, rehabilitation includes "any discourse or practices that speak to transforming or normalizing the criminal into a socially defined non-deviant citizen" (p. 45). With America's Judeo-Christian roots, a penal system based in concepts of religion makes sense. The United States has its foundations in the concept of a penitentiary as a prison with the "twin ideas that imprisonment should serve as a more humane form of punishment while rehabilitating the offender" being a part of America's

prisons “since its inception” (Craig, 2004, p. 93S). With Quaker influences in the late 18<sup>th</sup> century, “the Walnut Street Jail was used almost exclusively for the ‘correction’ and rehabilitation of convicted felons,” which came later to be known as “the Pennsylvania system” (Roth, 2005, p. 89). With this heavy influence from the Quakers and other Protestant fundamentalist influences (see Grasmick, Davenport, Chamlin, and Bursik, 1992), this system of incarceration had an ethic of care at its core.

Despite this ethic of care at its core, prison administrators also had to be creative with the ways they funded their prisons. According to Craig (2004), as prison administrators turned to prison labor to supplement the cost of maintenance, they introduced what has been called the “New York system,” where “inmates worked at hard labor in groups during the day and were confined to individual cells at night” (p. 94S). Together, these two systems “laid the groundwork for the future of corrections in the United States” (p. 94S). With the ever-increasing rise of prison populations, however, “punishment in practice became less concerned with implementing methods for reforming criminals, and began leaning more toward simply incapacitating offenders as efficiently as feasible” (Lynch, 2000, p. 41). This shift from rehabilitation to incapacitation is a shift from an ethic of care to an ethic of expediency.

### **An Ethic of Expediency**

According to Steen and Bandy (2007), two U.S. Supreme Court cases marked the beginning of “national shifts in the philosophy of punishment away from rehabilitation and treatment toward just deserts and retribution” (p. 6): *Furman v. Georgia* in 1972 and *Gregg v. Georgia* in 1976. Although these cases dealt specifically with the use of capital

punishment, they broadly addressed “the role of retribution in punishment” as well. Through majority and dissenting opinions in each case, “the U.S. Supreme Court declared that public opinion, including the public’s presumed desire for retribution, can be a legitimate basis for penal policy” (Grasmick, Davenport, Chamlin, and Bursik, 1992, p. 21). Following another U.S. Supreme Court decision in 1975 with *O’Connor v. Donaldson*, which ruled that the state cannot forcibly hospitalize individuals with mental illnesses, the state began using prisons instead of hospitals, which contributed to the sharp increase in prison populations. As more people arrived in prisons, their sparse funding was stretched even further. From these national shifts occurring, Steen and Bandy (2007) identify three types of punishment (pp. 7-8):

- 1) Expressive goals: “From a Durkheimian perspective, punishment is first and foremost a mechanism through which moral values are taught and enforced.”
- 2) Utilitarian goals: “In contrast to the expressive mode of reasoning, utilitarian models recognize crime control as the paramount goal of punishment.”
- 3) Managerial goals: These “privilege the management of an offender population over all punishment goals.”

With increases in prison population and the combination of several U.S. Supreme Court cases, managerial goals began dominating the function of prisons in America, leaving prison administrators left trying to figure out how to manage their populations in the most cost effective way possible, paving the way for the privatization of aspects of prison operations to the privatization of entire prison facilities—the prison industrial complex.

As these shifts took their roots, prisons no longer had an ethic of care at their core; instead, according to Garland (2001), prison “serves as an expressive satisfaction of retribution sentiments and an instrumental mechanism for the management of risk and the confinement of danger” (p. 199). In his thorough rhetorical analysis of a memo from Nazi Germany, Katz (1992) details what he calls the ethic of expediency. He describes this ethic as one being too technical, too logical—so much so, that the “writer shows no concern that the purpose of his memo is the modification of vehicles not only to improve efficiency, but also to exterminate people” (p. 257). As prisons have shifted away from rehabilitation, they “began leaning toward simply incapacitating offenders as *efficiently* as feasible” (Lynch, 2000, p. 41; emphasis added). This move towards efficiency is manifested in these handbooks. Katz argues, “It is well known that to perform well in a professional organization, writers must adopt the *ethos* of that organization” (p. 257). The same is true of prisons and inmate handbooks. An *ethos* has been adopted by the prison industrial complex. Katz goes on to suggest “that it is the ethic of expediency that enables deliberative rhetoric and gives impulse to most of our actions in technological capitalism” (p. 258). The prison industrial complex has embraced deliberative rhetoric and pushed prisons into a tool for control and capitalism. As Garland writes:

Crime control and criminal justice have come to be disconnected from the broader themes of social justice and social reconstruction. Their social function is now the more reactionary, less ambitious one of re-imposing control on those who fall outside the world of consumerist freedom. (p. 199)

What, then, is the *ethos* that has been adopted by the prison industrial complex?

According to Bosworth (2007), “The choice of rehabilitation or reform has become the individual prisoner’s sole responsibility. The prison is merely expected to provide the arena for such personal decisions while warehousing inmates securely” (p. 68). Bosworth argues that this is the “primary means of creating accountable and thus governable and obedient citizens” (p. 68).

This shift from an ethic of care to an ethic of expediency involves a fundamental shift of the subject: “Under this model, the driving ideal was that the punishment should fit the crime rather than the criminal” (Steen and Bandy, 2007, p. 9). An ethic of care, according to Cavarero, is inextricably connected to the unique and unrepeatable individual. An ethic of care connected a punishment to the person. An ethic of expediency leaves those unrepeatable, unique individuals to fend for themselves while administrations worry about the most effective way of housing inmates. An ethic of expediency connects punishment to the cost of housing inmates, not rehabilitating them. In Bosworth’s (2007) study of federal admission handbooks, she describes these shifts in language:

Whereas booklets from the 1960s and 1970s promise individualized care and attention in preparing inmates for release, the recent manuals are characterized by mission statements and promises of inmate satisfaction that seek less to help prisoners realize their potential, but rather to motivate them into becoming willing actors, working towards the goals of the institution and, increasingly, of the wider, globalized, society and marketplace. (p. 69)

Essentially, prisons have moved from sites of rehabilitation to sites of production. The production of goods run by the prison industrial complex and the production of willing actors to become inmate-workers producing goods and services to sustain their incapacitation. According to Craig (2004), “prison management is also personnel management, albeit the management of less-than-willing inmate-workers” (p. 97S). In the shift from rehabilitation to retribution, inmates are little more than a reluctant source of labor, and these handbooks reflect this change. Craig posits that prisons have turned to Theory X management philosophy developed by McGregor, an organizational theorist: “The hallmark of Theory X is the use of centralized control strategies to manage inherently untrustworthy workers” (p. 97S). According to Theory X, workers “are considered essentially lazy and motivated primarily by money,” and Theory X advocates for a mixture of detailed task specifications, functional specializations, rigid department boundaries, and bureaucratic hierarchies “designed to prevent the exercise of employee initiative” (pp. 97S-98S). Under this ethic of expediency and Theory X management, the needs of unique, unrepeatable inmates are lost. Lynch (2000) argues that under this new model “is the notion that the individual no longer needs to be known” (p. 41), which leads to what Trammel and Rundle (2015) call a “culture of disrespect,” where inmates are treated as nonpersons. In their study of correctional officer confrontation in front of inmate, Trammel and Rundle write that “the nonperson almost serves as an inanimate object that goes unrecognized because they have little power to discredit or disrupt the performance of those acting before them” (p. 473). Rather than individualized treatment for rehabilitation, inmates are treated as lazy, reluctant workers—if recognized at all.



## Methods

Using text as a data source is an emerging technique in several different fields (see Grimmer, 2015), and institutional genre analysis has heavy influence from computational grounded theory (Nelson, 2017), computational content analysis (Grimmer and Stewart, 2013), statistical genre analysis (Graham, Kim, DeVasto, and Keith, 2015), and qualitative content analysis (Hsieh and Shannon, 2005; Boettger and Palmer, 2010; Lewis, Zamith, and Hermida, 2013). Prisons are rich sources of data for research, but this has led to an unintended consequence described by Franko Aas (2004): “if social actors [inmates] need to be disciplined in a way that makes their behavior amenable to data collection, what are the qualities that distinguish (these types of) data from previous forms of knowledge” (p. 380)? Essentially, as researchers turn to a database to collect information about individuals, then what can be measured in a database takes precedence and redefines/reprioritizes certain behaviors. Franko Aas’s argument is that databases ought to be used as cultural innovations: “As a cultural expression, the database represents the world as a collection of items on which a user can perform various operations” (p. 383). Generally, these databases used in prisons are used to study how the institution impacts inmates as manifested by inmate behavior that has been cataloged. This project builds a database, but not from user-generated data; I build the database from institutional genres (see Stephens, 2018). I turn to genre not only for its value in revealing the activity of correction, but for the value in revealing the activity *of the institution itself*. In their article, Graham, Kim, DeVasto, and Keith (2015) show the value of reading an entire genre using statistical analysis: “It is a method designed to move technical

communication toward the ability to offer encompassing conclusions about larger data sets without losing the craft character of rhetorical inquiry” (p. 72). This method enables a rhetorical reading of an entire genre, and with genres as sites of social action (Miller, 1984), it ought to lead to action. With the intention rooted in Blyler’s (1995) critical interpretation and Jones, Moore, and Walton’s (2016) antenarratives and disruption of dominant narratives, an institutional genre analysis seeks to “[make room] for marginalized voices to speak and be heard” (Stephens, 2018). In this case, inmates.

### **Describing the Genre**

Many of the federal handbooks described by Bosworth (2007) have similar characteristics: information about the discipline system, listing offences and associated punishments, the organizational structure of the prison, regulations and opportunities for work, and visits and education programs (p. 71). Bosworth’s study in 2007 had a sample of federal handbooks that averaged about 70 pages per handbook. Of the 347 handbooks I collected, 116 of them are federal handbooks ranging from 2008 to 2018, averaging 63.5 pages per handbook. The formatting between all of the federal handbooks are quite similar, as detailed by Bosworth. Many of the same components are found in handbooks at the local, county, and state levels, but the level of detail and complexity varies wildly. Marion County Jail in Knoxville, IA has a handbook only three pages long; the Texas State Department of Corrections uses one handbook for multiple facilities, which is 146 pages long. Some of the handbooks contain welcome messages from the warden, and others do not. Nearly all of the handbooks contain the logo or emblem of the governing body on the first pages, only a small portion in color. More than 90% of the handbooks

were written and designed for a single complex with an average length of 45.2 pages. The other 10% cover more than one complex with an average of 46.3 pages per handbook. Only nine (2.6%) of the handbooks identify a specific gendered audience: two handbooks for females and seven handbooks for males. The other 97.4% of the handbooks either do not specify a gendered audience or explicitly identify both female and male inmates. Only three of the handbooks were created before 2000: the Maryland Division of Corrections from 1984, the Bucks County Correctional Facility from 1988, and the Halawa Correctional Facility from 1999. The rest of the handbooks range from 2004 to 2018 with 89.6% created since 2010.

### **Building the Database**

Handbooks were collected from the local, county, state, and federal levels from all 50 states and one federal medical detention center in Guaynabo, Puerto Rico. In total, there are 347 handbooks with 15,719 pages and more than 425,000 unique n-grams. Many of the facilities from which these handbooks were mined have them available on their individual websites. Often these handbooks are made available digitally so family and friends can access the same information, implying secondary audiences for the content. Rather than visit and comb each facilities websites, I performed an advanced Google search using variations on the phrase, “inmate handbooks filetype:pdf.” Often I would add a particular state in the search as I exhausted the results of the initial search. Like the handbooks themselves at the federal level, to access the federal handbooks was systematic, making the mining much easier. Figures 1 and 2 show a screenshots of the Bureau of Prisons website.

Figure 4.1 Screenshot of the listed locations of federally run prisons

The screenshot shows the Federal Bureau of Prisons website. At the top, there is a navigation bar with links for Home, About Us, Inmates, Locations, Jobs, Business, Resources, and Contact Us. Below this is a section titled "Our Locations" with sub-links for "List of our facilities", "Map of our locations", and "Search for a facility". There are several filter buttons: "All Regions", "All States", "All Facility Types", "All Security Levels", "All Prison Types", and "Uncategorized". The main content area displays a grid of 60 facility names, such as Adams County CI, Chicago MCC, Gilmer FCI, Marion USP, Oklahoma City FTC, and Schuylkill FCI.

Adams County CI	Chicago MCC	Gilmer FCI	Marion USP	Oklahoma City FTC	Schuylkill FCI
Alderson FPC	Chicago RRM	Glynco	McCreary USP	Orlando RRM	Seagoville FCI
Aliceville FCI	Cincinnati RRM	Grand Prairie	McDowell FCI	Otisville FCI	SeaTac FDC
Allenwood FCC	Coleman FCC	Great Plains CI	McKean FCI	Oxford FCI	Seattle RRM
Ashland FCI	Cumberland FCI	Greenville FCI	McRae CI	Pekin FCI	Sheridan FCI
Atlanta RRM	D. Ray James CI	Guaynabo MDC	Memphis FCI	Pensacola FPC	South Central RO
Atlanta USP	Dallas RRM	Hazleton FCC	Mendota FCI	Petersburg FCC	Southeast RO
Atwater USP	Danbury FCI	Herlong FCI	Miami FCI	Philadelphia FDC	Springfield MCFP
Baltimore RRM	Detroit RRM	Honolulu FDC	Miami FDC	Philadelphia RRM	St Louis RRM
Bastrop FCI	Devens FMC	Houston FDC	Miami RRM	Phoenix FCI	Taft CI
Beaumont FCC	District of Columbia RRM	Houston RRM	Mid-Atlantic RO	Phoenix RRM	Talladega FCI
Beckley FCI	Dublin FCI	Jesup FCI	Milan FCI	Pittsburgh RRM	Tallahassee FCI
Bennettsville FCI	Duluth FPC	Kansas City RRM	Minneapolis RRM	Pollock FCC	Terminal Island FCI
Berlin FCI	Edgefield FCI	La Tuna FCI	Montgomery FPC	Raleigh RRM	Terre Haute FCC
Big Sandy USP	El Paso RRM	Leavenworth USP	Montgomery RRM	Ray Brook FCI	Texarkana FCI
Big Spring (Flightline) CI	El Reno FCI	Lee USP	Morgantown FCI	Reeves III CI	Thomson AUSP
Big Spring CI	Elkton FCI	Lewisburg USP	Moshannon Valley CI	Rivers CI	Three Rivers FCI
Big Spring FCI	Englewood FCI	Lexington FMC	MSTC	Rochester FMC	Tucson FCC
Bostom RRM	Estill FCI	Lompoc FCC	Nashville RRM	Sacramento RRM	Victorville FCC
Brooklyn MDC	Fairton FCI	Long Beach RRM	New York MCC	Safford FCI	Waseca FCI
Bryan FPC	Florence FCC	Loretto FCI	New York RRM	Salt Lake City RRM	Western RO
Butner FCC	Forrest City FCC	Los Angeles MDC	North Central RO	San Antonio RRM	Williamsburg FCI
Canaan USP	Fort Dix FCI	Manchester FCI	Northeast RO	San Diego MCC	Yankton FPC
Carswell FMC	Fort Worth FMC	Marianna FCI	Oakdale FCC	Sandstone FCI	Yazoo City FCC
Central Office HQ	Giles W. Dalby CI				

Once a user clicks on a location, like “Leavenworth USP” for example, they can click on “Resources for sentenced inmates,” and then view/download the current handbook in circulation as shown in Figure 4.2.

Figure 4.2 Screenshot of USP Leavenworth's landing page



1300 METROPOLITAN  
LEAVENWORTH, KS 66048

Email: [LVN/ExecAssistant@bop.gov](mailto:LVN/ExecAssistant@bop.gov)  
Phone: 913-682-8700  
Fax: 913-578-1010

Inmate Gender: Male Offenders  
Population: **1,930 Total Inmates**  
1,488 Inmates at the USP  
442 Inmates at the Camp  
Judicial District: Kansas  
County: LEAVENWORTH  
BOP Region: [North Central Region](#)

[Visiting Information](#) ▾  
[How to send things here](#) ▾  
[Resources for sentenced inmates](#) ▾  
[Driving Directions](#) 📍  
[Job Vacancies](#) 📄

## Resources for sentenced inmates

### Admissions and Orientation (A&O) Handbook

This document provides you with general information about the institution, programs, rules, and regulations that you will encounter during your confinement. Familiarizing yourself with this information and knowing your responsibilities will help you to adjustment to institution life.

[USP Leavenworth Admissions & Orientation Handbook](#) ▶

*Also available in Spanish: [USP Leavenworth Guía de Admisión y Orientación](#)*

As I collected the handbooks, I also made note of the location of each facility and their longitude and latitude coordinates. Figure 4.3 shows the locations of each facility while figure 4.4 shows the sizes of unique n-grams by operational level.

Figure 4.3 Locations of each handbook

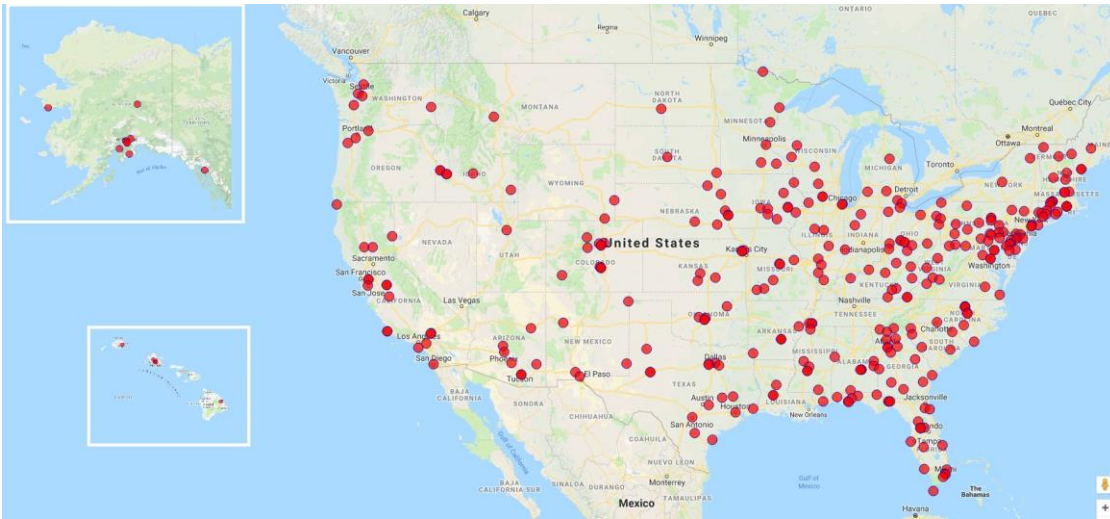
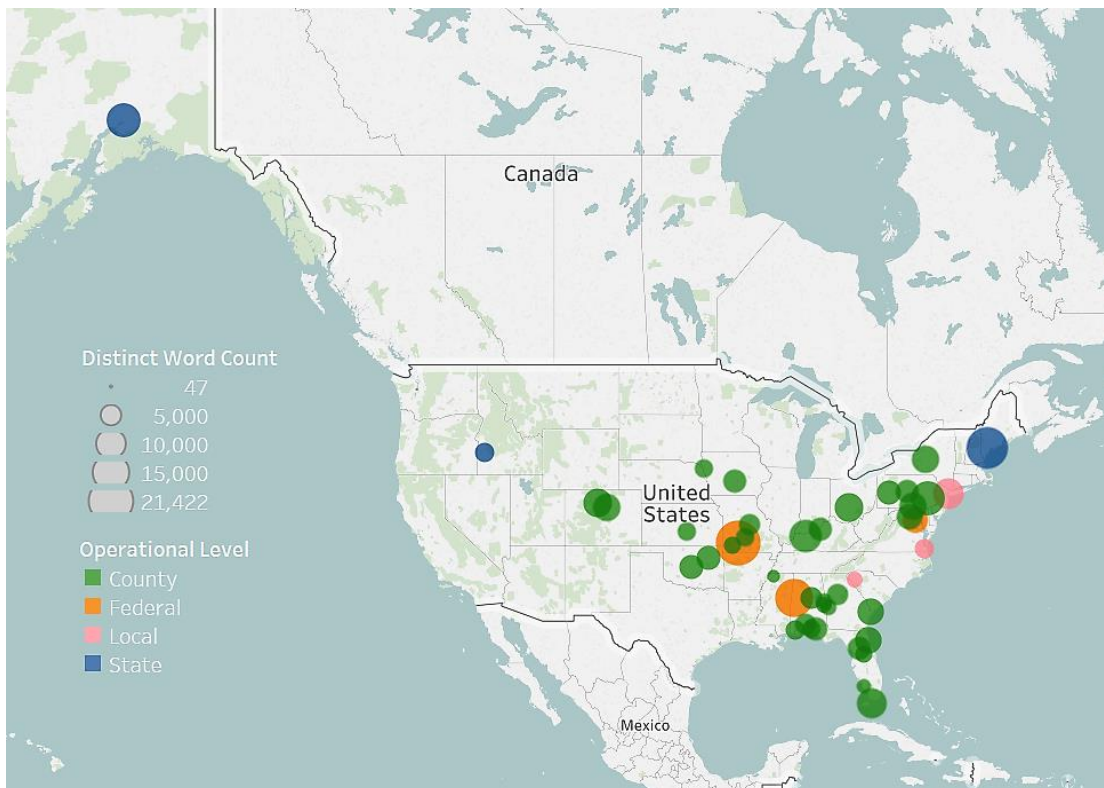


Figure 4.4 Map showing the sizes of unique n-grams by operational level



## Writing the Code

The first step in writing the code is to find someone to help you write the code if you don't know how. I turned to two professional data scientists, Ben Webster and Katie Bakewell. The task for the python code is to take the 347 PDF files and convert them into two types of deliverables: 1) a single TXT file for each individual page of each individual handbook, and 2) a CSV file with each sentence in its own cell with its own unique ID.

To do so, the code must accomplish several things:

- 1) It must separate each handbook into individual PDF files, which are then put into a folder for each handbook.
- 2) If needed, each individual page is optimized and decrypted.
- 3) Each individual page is converted from PDF to TXT.
- 4) Each TXT file is scrubbed of header and footer information, stop words, and unnecessary periods and abbreviations.
- 5) Each scrubbed TXT file is parsed into its individual sentences and put into a CSV file.
- 6) Each sentence is given a unique ID.
- 7) If needed, a CSV file of randomly mined sentences is created for content coding.

Rather than recreating this code, we have made it available as an appendix at the end of the dissertation.

Separating out each individual sentence has its advantages for future research and content analysis. Rather than giving content coders several sentences from individual handbooks, they code random sentences out of context. This enables the machine

learning models to treat each unique sentence not as a part of a particular handbook, but a sentence in a particular genre, allowing the model to read the entire corpus as a single document rather than 347 individual documents. As this study relies on unsupervised machine learning, it does not require the random sampling for content coding.

## **Results**

### **Looking for Themes**

The first step in Nelson's (2017) computational grounded theory is "pattern detection using human-centered computational exploratory analysis" (p. 9). She goes onto describe the step:

One of the principle ways computer-assisted text analysis techniques can help [researchers] explore text is by reducing complicated, messy text into simpler, more interpretable lists or networks of words. When compared to one another or when their frequencies are measured across texts, the lists or networks of words can suggest relevant patterns with then text, which can lead to extracting meanings embedded in the text. While output must still be interpreted by humans, computational exploratory analysis can suggest categories relevant to the text that researchers had not preconceived notions about, or the complexity of, the text. (pp. 9-10)

In other words, the unsupervised machine learning has the ability to identify complex patterns across a corpus of text—able to accurately identify synonyms and antonyms based on the genre itself—but humans must identify and interpret those patterns. In order to build on existing literature, I chose to look for themes identified by Bosworth's (2007)



of federal handbooks to see if the same patterns exist in more current handbooks at all levels of facility operations. As I read Bosworth (2007), I searched for two themes that might operate on two ends of a continuum, along with a third theme acting as the dependent variable: “rehabilitation,” “punishment,” and “work.” Figure 4.4, 4.5, and 4.6 show the tokens, the similarity cosine threshold of .80, a list of synonyms (that occurred at least 25 times) generated by the learning model, and the percentage of words in the handbooks related to the themes organized by region and operational level.

Figure 4.5 Breaking down “rehabilitation” in 347 inmate handbooks

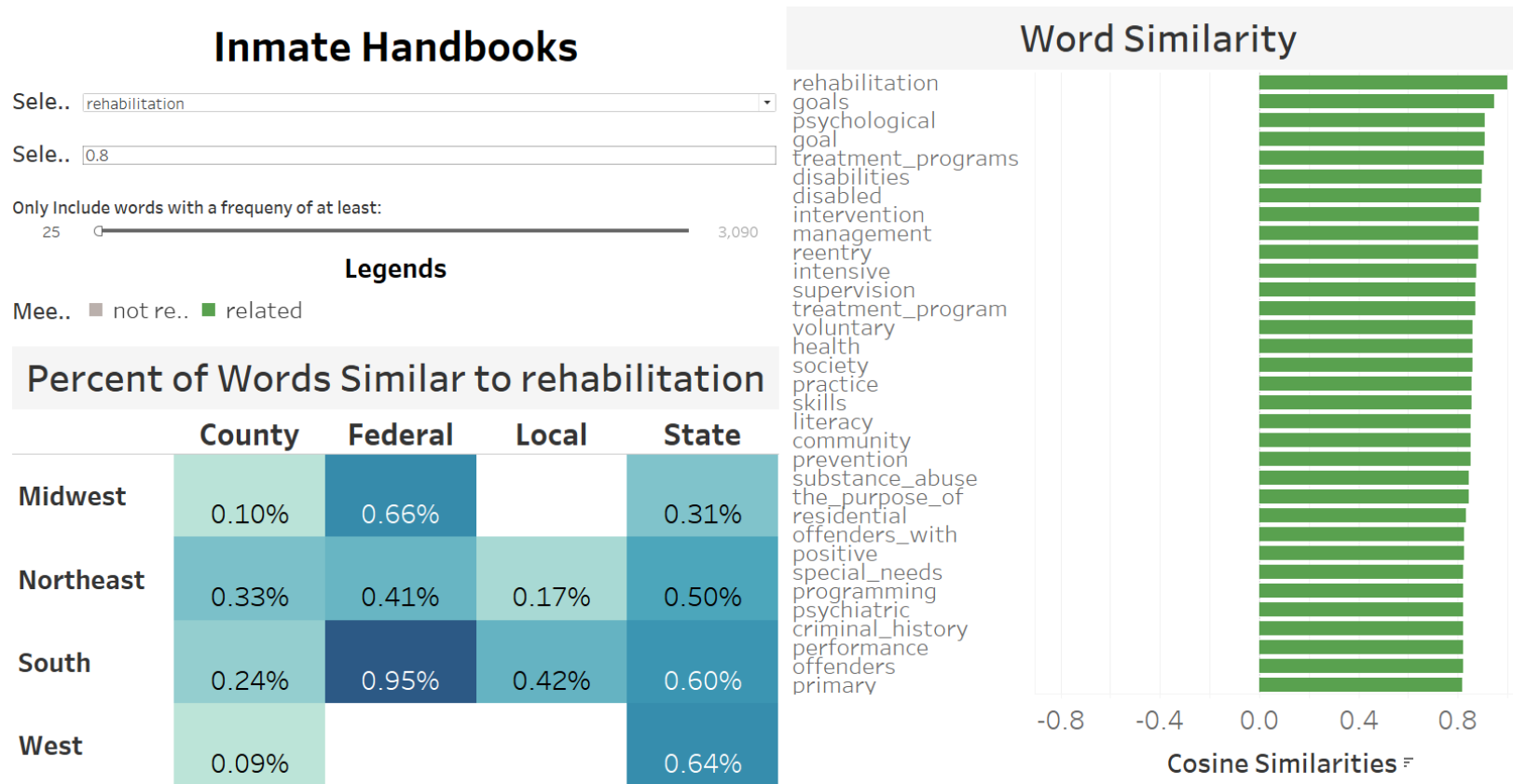


Figure 4.6 Breaking down “punishment” in 347 inmate handbooks

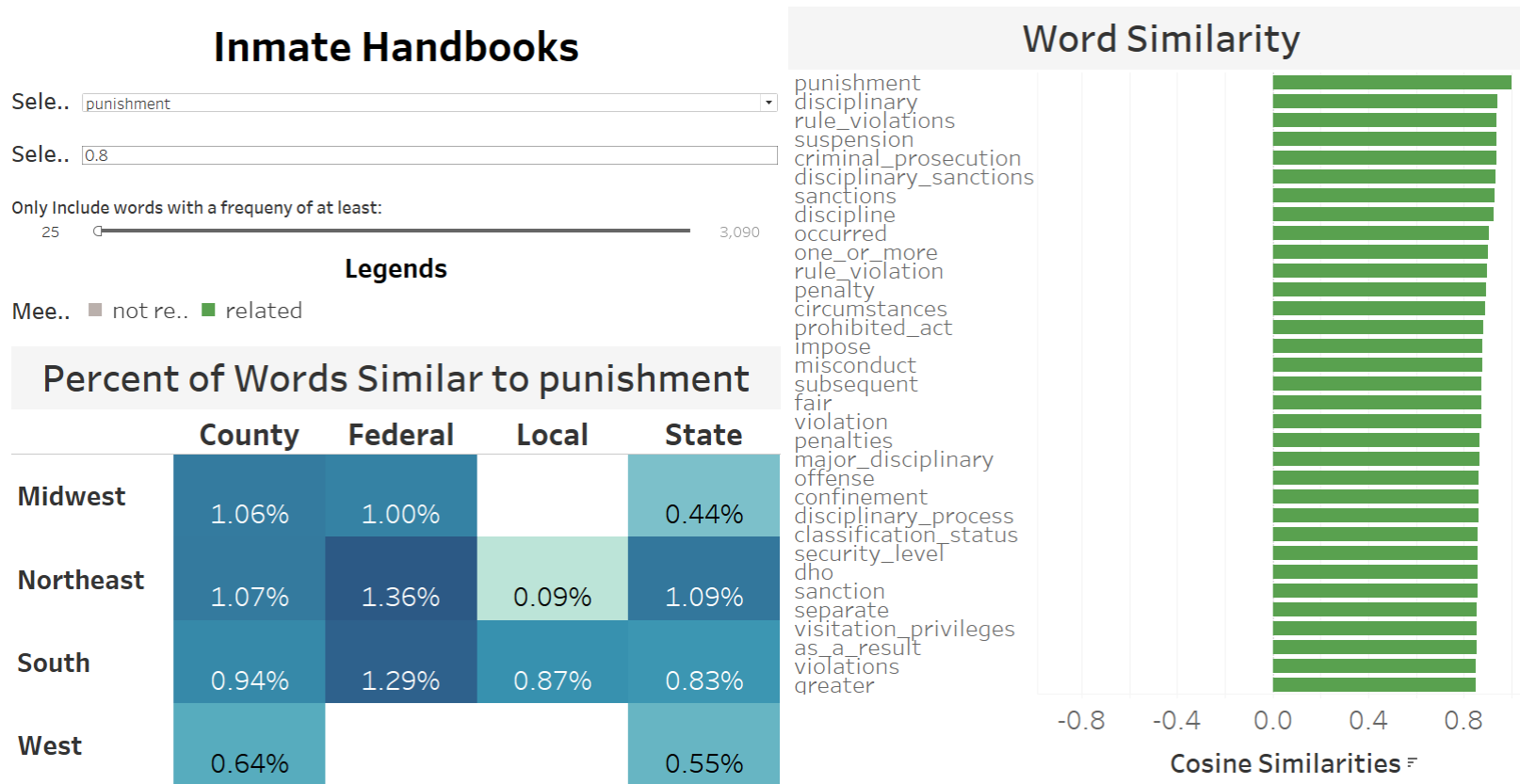
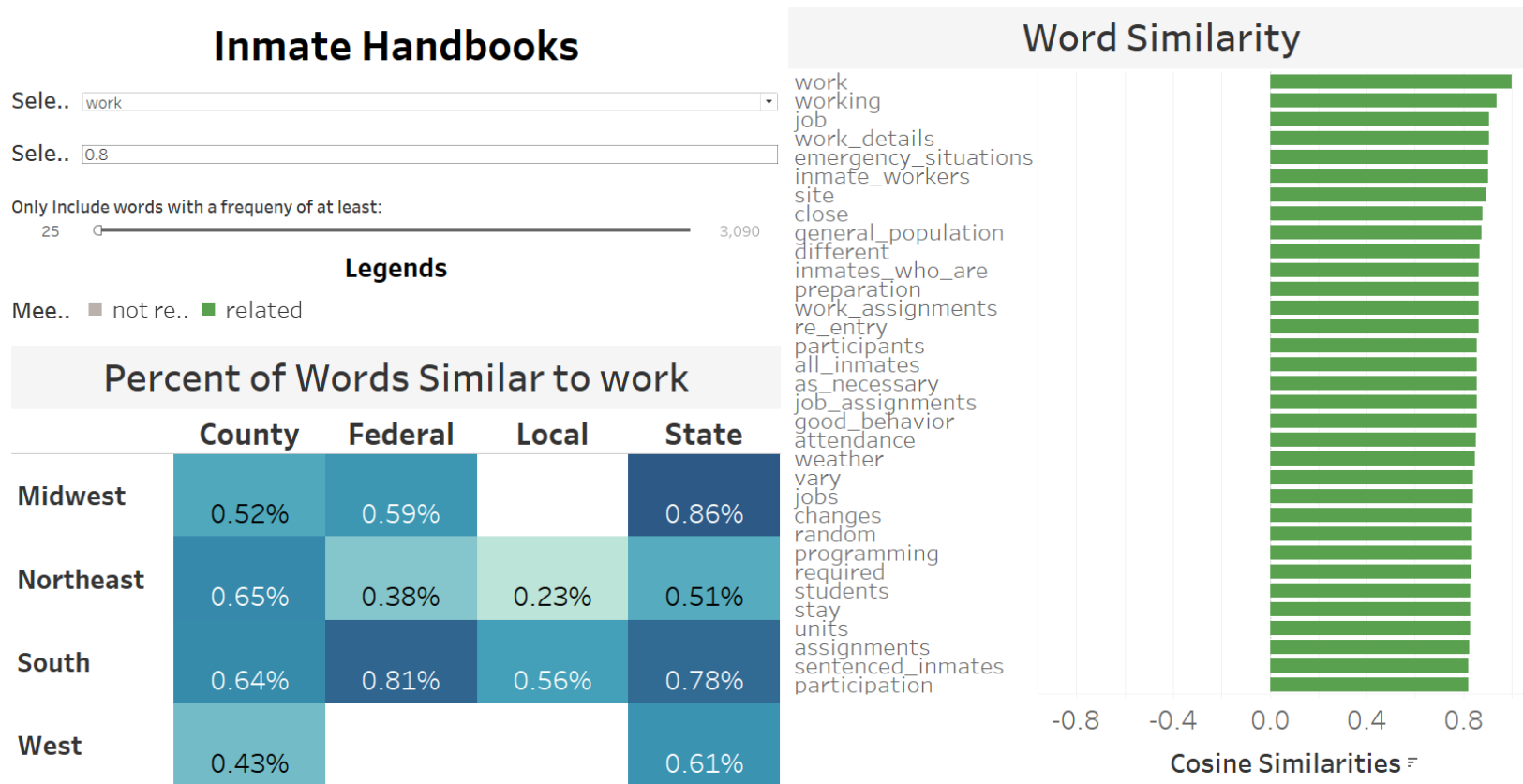


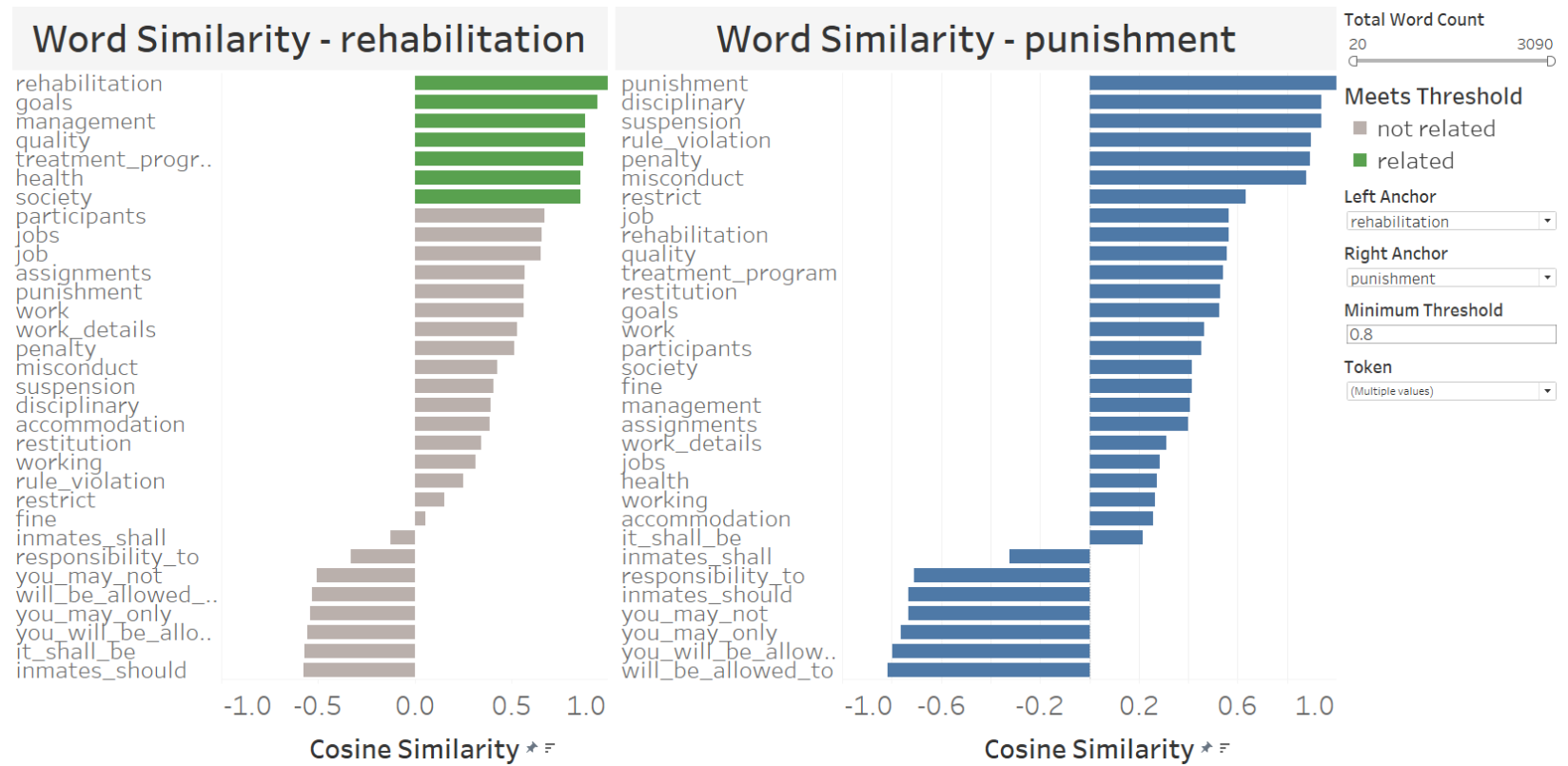
Figure 4.7 Breaking down “work” in 347 inmate handbooks



## **Testing for Themes**

With the two themes and a set of dependent variables selected, I can begin to triangulate the data. Figure 4.8 shows the two themes, “rehabilitation” on the left and “punishment” on the right. In addition to the token “work” as a dependent variable, I also selected what might be considered commands, but consist of subject-verb n-grams where inmates occupy most of the subject positions. The bars show dis/similarity levels, with synonyms moving to the left of 0.0 and antonyms moving to the right of 0.0.

Figure 4.8 Triangulating “rehabilitation,” “punishment,” and “work,” along with subject-verb n-grams



## **Analysis and Discussion**

In her analysis of federal handbooks written before 2007, Bosworth (2007) discusses the structure of the genre, “[containing] the same information about the discipline system, listing offenses, the punishment they will provoke and the procedures that inmates who violate good order and discipline will undergo” (p. 71). She goes on to suggest that the handbooks address “individuals capable of exercising reason” (p. 72). Craig (2004) argues that this shift from perceiving inmates as individuals needing rehabilitation to individuals capable of exercising reason has resulted in the use of Theory X, a management system that treats inmates as unmotivated workers.

According to Figure 4.5, many of the machine learning generated synonyms suggest that “rehabilitation” is most associated with making goals, psychological treatment programs, intervention, health, literacy, community, and reentry—all n-grams that one would already associate with rehabilitation. The same is true for “punishment” as seen in Figure 4.6. It is clear that both federal and state run institutions discuss rehabilitation more than county or local facilities, which are used more often as holding facilities for trial or short-term sentences. Interestingly, it appears that facilities in the South are more likely to discuss rehabilitation in their handbooks. Figure 4.7 shows that some synonyms for “work” that a human coder might relate to “rehabilitation”: job, good\_behavior, and participation. Figure 4.6 shows that “punishment” is discussed much more often (or at a similar rate) across all institutions and operational levels, except for local facilities in the Northeast. This implies that while local and county facilities are not

as concerned with rehabilitation, they devote more time to punishment in their handbooks.

According to the heat map in Figure 4.7, state facilities in the Midwest discuss “work” more than any other area or operational level. It appears that the South discusses work more than other regions, which supports Alexander’s (2010) argument in *The New Jim Crow* where she argues that governments use prisons as a source for cheap, racialized labor. The synonyms generated by the learning model suggest that work is designed to prepare individuals to reentry, but a close reading of the handbooks would show a more nuanced understanding of the relationship between “work” and Theory X.

Figure 4.8 shows the results with interesting implications. It appears that “work” and some of its synonyms are more closely related to “punishment” than “rehabilitation,” suggesting that work may be used more as something to be rewarded or taken away rather than a tool for re-entering society. This, of course, would support both Bosworth and Craig’s claims that prisons have transitioned into personnel management than rehabilitation. What it perhaps the most compelling piece of information from the learning model’s output is the relationship of punishment and rehabilitation with n-grams whose subjects are inmates: *you\_may\_not*, *you\_may\_only*, *you\_will\_be\_allowed*, *will\_be\_allowed\_to*, and *inmates\_should*. Each of these is considered an antonym to the two primary themes, which again suggests that language explicitly addressed to inmates has little to do with rehabilitation or punishment, implying that they are more associated with personnel management as Bosworth (2007) and Craig (2004) suggest.



## **Conclusion**

It is clear from these unsupervised learning models that genres have a great deal to teach researchers. Working with text as data as a reciprocal relationship. The learning models can produce myriads of information, and a human researcher can engage with it to begin building studies from the data using supervised learning techniques. Through the first step of computational grounded theory (Nelson, 2017), the models found several patterns from the corpus, and those findings have been shown to support the concept that prisons have in fact moved beyond rehabilitation to personnel management. Further study is recommended, however, in order to hone those ideological perspectives. This study shows is that TPC researchers can engage with data, and that the data can speak to the researcher. Ideally, this data will speak to and speak with marginalized individuals and those who advocate for them. Institutional genre analysis reveals oppressive patterns that have been crystallized into policy. In other words, the analysis of these handbooks empirically shows the current state of prisons acting not from an ethic of care but from an ethic of expediency.

## References

- Alexander, M. (2010). *The new Jim Crow: Mass incarceration in the age of colorblindness*. New York: New Press.
- Blyler, N. R. (1995). Research as ideology in professional communication. *Technical Communication Quarterly*, 4(3), 285–313.  
<https://doi.org/10.1080/10572259509364602>
- Bodker, S. (1997). Computers in mediated human activity. *Mind, Culture, and Activity*, 4(3), 149–158. [https://doi.org/10.1207/s15327884mca0403\\_2](https://doi.org/10.1207/s15327884mca0403_2)
- Boettger, R. K., & Palmer, L. A. (2010). Quantitative content analysis: Its use in technical communication. *IEEE Transactions on Professional Communication*, 53(4), 346–357. <https://doi.org/10.1109/TPC.2010.2077450>
- Bosworth, M. (2007). Creating the responsible prisoner: Federal admission and orientation packs. *Punishment & Society*, 9(1), 67–85.  
<https://doi.org/10.1177/1462474507070553>
- Cavarero, A. (2000). *Relating narratives: Storytelling and selfhood*. New York: Routledge.
- Coates, T.-N. (2015, October). The black family in the age of mass incarceration. *The Atlantic*, 60–84.
- Colton, J. S., Holmes, S., & Walwema, J. (2017). From NoobGuides to #OpKKK: Ethics of Anonymous' tactical technical communication. *Technical Communication Quarterly*, 26(1), 59–75. <https://doi.org/10.1080/10572252.2016.1257743>
- Craig, S. C. (2004). Rehabilitation versus control: An organizational theory of prison management. *The Prison Journal*, 84(4\_suppl), 92S–114S.  
<https://doi.org/10.1177/0032885504269394>
- Foucault, M. (1977). *Discipline and punish: The birth of the prison* (1st American ed). New York: Pantheon Books.
- Franko Aas, K. (2004). From narrative to database: Technological change and penal culture. *Punishment & Society*, 6(4), 379–393.  
<https://doi.org/10.1177/1462474504046119>
- Garland, D. (2012). *The culture of control: Crime and social order in contemporary society*. University of Chicago Press.

- Graham, S. S., Kim, S.-Y., DeVasto, D. M., & Keith, W. (2015). Statistical genre analysis: Toward big data methodologies in technical communication. *Technical Communication Quarterly*, 24(1), 70–104. <https://doi.org/10.1080/10572252.2015.975955>
- Grasmick, H. G., Davenport, E., Chamlin, M. B., & Bursik Jr., R. J. (1992). Protestant fundamentalism and the retributive doctrine of punishment. *Criminology*, 30(1), 21–45.
- Grimmer, J. (2015). We are all social scientists now: How big data, machine learning, and causal inference work together. *PS: Political Science & Politics*, 48(1), 80–83. <https://doi.org/10.1017/S1049096514001784>
- Grimmer, J., & Stewart, B. M. (2013). Text as data: The promise and pitfalls of automatic content analysis methods for political texts. *Political Analysis*, 21(3), 267–297.
- Hsieh, H.-F., & Shannon, S. E. (2005). Three approaches to qualitative content analysis. *Qualitative Health Research*, 15(9), 1277–1288. <https://doi.org/10.1177/1049732305276687>
- Jones, N. N., Moore, K. R., & Walton, R. (2016). Disrupting the past to disrupt the future: An antenarrative of technical communication. *Technical Communication Quarterly*, 25(4), 211–229. <https://doi.org/10.1080/10572252.2016.1224655>
- Katz, S. B. (1992). The ethic of expediency: Classical rhetoric, technology, and the Holocaust. *College English*, 54(3), 255–275. <https://doi.org/10.2307/378062>
- Lewis, S. C., Zamith, R., & Hermida, A. (2013). Content analysis in an era of big data: A hybrid approach to computational and manual methods. *Journal of Broadcasting & Electronic Media*, 57(1), 34–52. <https://doi.org/10.1080/08838151.2012.761702>
- Lynch, M. (2000). Rehabilitation as rhetoric: The ideal of reformation in contemporary parole discourse and practices. *Punishment & Society*, 2(1), 40–65. <https://doi.org/10.1177/14624740022227854>
- Miller, C. R. (1984). Genre as social action. *Quarterly Journal of Speech*, 70(2), 151–167. <https://doi.org/10.1080/00335638409383686>
- National Research Council. (2014). *The growth of incarceration in the United States: Exploring causes and consequences*. Retrieved from <https://www.nap.edu/catalog/18613/the-growth-of-incarceration-in-the-united-states-exploring-causes>

- Nelson, L. K. (2017). Computational grounded theory: A methodological framework. *Sociological Methods & Research*, 49(124117729703).  
<https://doi.org/10.1177/0049124117729703>
- Penance. (n.d.). *OED Online*. Oxford University Press. Retrieved from  
<http://www.oed.com/view/Entry/139991>
- Penitentiary. (n.d.). *OED Online*. Oxford University Press. Retrieved from  
<http://www.oed.com/view/Entry/263199>
- Rickert, T. J. (2013). *Ambient rhetoric: The attunements of rhetorical being*. Pittsburgh, Pa: University of Pittsburgh Press.
- Rittel, H. W. J., & Webber, M. M. (1973). Dilemmas in a general theory of planning. *Policy Sciences*, 4, 155–169.
- Rivers, N. A., & Weber, R. P. (2011). Ecological, pedagogical, public rhetoric. *College Composition and Communication*, 63(2), 187–218.
- Roth, M. P. (2010). *Crime and punishment: A history of the criminal justice system*. Cengage Learning.
- Spinuzzi, C. (2003). *Tracing genres through organizations: A sociocultural approach to information design*. MIT Press.
- Steen, S., & Bandy, R. (2007). When the policy becomes the problem: Criminal justice in the new millennium. *Punishment & Society*, 9(1), 5–26.  
<https://doi.org/10.1177/1462474507070548>
- Stephens, E. J. (2018). *Algorithmic auras and institutional genre analysis: A case study for working ethically with big data*. Dissertation chapter, Clemson University.
- Sykes, G. M. (1958). *The society of captives: A study of a maximum security prison*. Princeton University Press.
- Wacquant, L. (2009). *Punishing the poor: The neoliberal government of social insecurity*. Duke University Press.
- Wald, J., & Losen, D. J. (2003). Defining and redirecting a school-to-prison pipeline. *New Directions for Youth Development*, 2003(99), 9–15.  
<https://doi.org/10.1002/yd.51>

## ARTICLE 5

### THE WICKED AND THE MUNDANE: AN ALTERNATIVE PROJECT-BASED COURSE FOR THE MULTI-MAJOR TECHNICAL AND PROFESSIONAL COMMUNICATION CLASSROOM

#### **Introduction**

There are few other fields or practitioners whose professional identities are more fluid than that of technical and professional communication (TPC). While many scholars connect the practice of TPC to a variety of different historical eras, Pringles and Williams (2005) contend that the prime of the profession came about following the industrial revolution and during the advent of World War II when machines began replacing human labor. Technical communicators fulfil a wide variety of roles in the workplace: content managers and strategists, technical editors, information architects, user experience professionals, and project managers. In their introduction to *Solving Problems in Technical Communication*, Johnson-Eilola and Selber (2013) argue that being a technical communicator isn't just about helping people use technology; they need to work in tandem with technologies. They write: "technical communicators continually face both old and new problems [and] . . . problems vary by context and circumstance, sometimes dramatically" (p. 5). Given so many types of technologies and tools that exist in workplaces, few technical communicators share common job titles and nearly every profession practices TPC. Even defining the field of technical communication seems impossible with entire books and edited collections trying to articulate a cohesive definition. Selfe and Selfe (2013) use field maps to understand three ways to describe the field: its history, its research base, and its skillset, allowing students to situate themselves

in this growing field. I draw attention to the fluidity of TPC to suggest that technical communicators are well-equipped to address the “wicked problems” of the world (Rittel and Webber, 1979; Wickman, 2014).

Often in our classrooms, however, we aren’t always teaching students who want to become technical and professional communicators. Sometimes we’re teaching students from multiple majors in the same classroom, trying to convey abstract writing principles that might apply to individual writing contexts for students coming from and going to a variety of different professions and levels of writing needed in those professions. What follows is a TPC course for multi-majors that focuses on two abstract concepts: wicked problems and mundane artifacts. This course enacts a framework of design thinking (acceptance of ambiguity, productive failure, radical collaboration, and a bias towards action) and a strategic model of thinking popularized in the United Kingdom called PESTEL (political, economic, socio-cultural, technological, environmental, and legal). The combination of design thinking and PESTEL create a classroom environment where students can become subject matter experts (SME) as it pertains to their own fields of study, enabling students to understand the multi-faceted nature of problems while building the six layered literacies described by Cargile Cook (2002): basic, rhetorical, social, technological, ethical, and critical. This course seeks to encompass all of these things with a wicked problem that impacts everyone—a zombie apocalypse.

## **Wicked Problems and Mundane Artifacts**

### **Wicked Problems**

In his article, Wickman (2014) designed a course for his fall 2010 technical writing class that revolved around a single, massive problem: the Exxon Oil Spill. Not only was this a current event unfolding during the semester, but it was a problem with no easy cause and no easy solution. Wickman cites Rittel and Webber (1979), who developed the concept of a wicked problem, which is a problem so ill-defined that they demand our collective attention. Where many problems are complicated, like building an airplane that has millions of parts, wicked problems are complex with no linear model of definition and solution. Wickman (2014, pp. 28-29) summarizes their concept of wicked problems with ten characteristics:

- 1) There is no definite formulation of a wicked problem.
- 2) Wicked problems have no stopping rule.
- 3) Solutions to wicked problems are not true-or-false, but good-or-bad.
- 4) There is no immediate and no ultimate test of a solution to a wicked problem.
- 5) Every solution to a wicked problem is a “one-shot operation”; because there is no opportunity to learn by trial-and-error, every attempt counts significantly.
- 6) Wicked problems do not have an enumerable (or exhaustively describable) set of potential solutions, nor is there a well-described set of permissible operations that may be incorporated into the plan.
- 7) Every wicked problem is essentially unique.
- 8) Every wicked problem can be considered to be a symptom of another problem.

9) The causes of the wicked problem can be explained in numerous ways. The choice of explanation determines the nature of the problem's resolution.

10) The planner has no right to be wrong.

Mehlenbacher (2013) contends that these wicked problems have beginnings and endings that are difficult to identify with rules that are often obscured or incomplete (p. 191).

What might work in one situation won't work in another situation that may even seem to be identical. The volatility of wicked problems matches the fluidity of TPC.

Mehlenbacher (2013) claims that "technical communicators routinely generate documents in ill-structured domains, that is, in environments that are unstable, that demand flexibility and a creative ability to organize across similar but always different problems and to understand, argue, and evaluate both conceptually and pragmatically" (p. 190). TPC already operates within these domains while maintaining the interests of various stakeholders. For this reason, TPC has the prerequisite skills for not only engaging with wicked problems, but identifying them in the first place as well.

### **Mundane Artifacts**

While Rivers and Weber's (2011) article focus on a public rhetorics course, they draw attention to the idea of mundane artifacts. The problem they address is when students are asked to perform public rhetoric but are not taught about the bureaucracy involved in holding a political speech for example. For them, the end result of making rhetoric public should not be the only emphasis in a class but should also engage with mundane artifacts required to make that rhetorical argument public. They "want to move students beyond the idea that most public change happens through a single author writing



a single text for a single audience” (p. 189). For example, a political speech in a park might have a great impact, but the speech itself might not have happened without permits, advertising, or facilitation. In his discussion of big data methodologies, McNely (2012) calls attention to the ambient nature of these artifacts that operate in the background with powerful yet overlooked influence. “Approaching this challenge,” McNely writes, “means designing and conducting novel methods of *ambient research*—that is, exploring ambient data by making use of ambient strategies” (p. 28). What is ambient, according to Rickert (2013), is just as important in shaping the world and who we are as much as other rhetorical artifacts that spend more time in the lime light. In technical communication, Longo and Fountain (2013) claim that technical writing is full of these mundane artifacts: TPC “[creates] scientific, technical, and business documents that not only convey information but also create systems of order that influence routine practices” (p. 165). They go onto say that TPC, “then, [uses] documents to order knowledge, shape information, and make implicit and explicit arguments about what is to be valued” (p. 169). I read Longo and Fountain as an empowerment to students in the TPC classroom. The documents they make and will make actually matter, but with that empowerment comes accountability for those documents. This awareness implicates students, making them responsible for the words they contribute to these mundane artifacts.

### **Frameworks for TPC Pedagogy**

As technical communication developed as a field, so did the scholarship on how to best prepare students to enter the workforce. With Miller’s (1979) work, “A Humanistic Rationale for Technical Writing,” a shift seemed to occur in how TPC

approaches pedagogy. What I find amusing about Miller's discussion is how the circumstance and heuristic for writing the article was founded in a rather complex problem involving mundane artifacts. Students wanted to fulfil their humanities requirement by taking a technical writing course. The argument in Miller's case stemmed from a mundane artifact: course descriptions and department policies. What may seem like an arbitrary action item on an agenda ended up shaping the way TPC defines itself. With a new foundation for teaching technical communication through a humanities lens, scholars began to question the role of their current classroom practices. Instructors began seeing their courses as more than a "service" course embodying what Scott (2004) calls hyperpragmatism. Instructors shifted from the importance of efficacy and conciseness (as influenced by Katz, 1992) and moved their work towards a deeper understanding of the ethical impact these mundane artifacts have in the world. In one of his essays on the topic, Scott (1995) argued for a sophisticated approach to teaching in the classroom. With an understanding of sophisticated rhetorics and their teachings that language is an ambiguous thing that shapes language, knowledge, and society (p. 189), he establishes the idea of TPC as an ethical practice, enabling students in three ways: 1) recognizing versions of truth as culturally constructed, 2) considering the consequences of their words, and 3) inviting students to develop their own ethical codes (Scott, 1995, p. 194). Today, three types of TPC pedagogies have emerged that dominate teaching practices: case studies, client projects, and service courses.

## **Case Studies**

Due to resource limits, case studies are probably the oldest and most common form of TPC instruction where students and teachers read about the stories and the circumstances that may or may not have happened. While case studies certainly show a variety of situations, they are always pre-framed. Someone with a perspective wrote them in a way to teach a certain thing. According to a study by Rozumalski and Graves (1995), case studies proved helpful for inexperienced writers, but for experienced writers, the case studies “appeared to produce similar writing” as their control group (p. 91). Although these can be helpful, they shouldn’t be the primary heuristic for teaching technical communication. A problem that many instructors face no matter which pedagogy they use, each comes into conflict with what Spinuzzi (1996) calls pseudotransactionality: “writing that is patently designed by a student to meet teacher expectations rather than to perform the ‘real’ function the teacher has suggested” (p. 295). This conflict, however, appears to be more prevalent in case study pedagogy because there is little to no relationship being cultivated than teacher and student.

## **Client Projects**

One way to avoid pseudotransactionality is to give students actual clients. Blakeslee (2001) offers a thorough guide for anyone looking to see the advantages and disadvantages of a client-project based classroom: “Such assignments, which ask students to complete workplace projects provided by clients, potentially preserve more of the culture of the workplace, while also allowing students to address a variety of audiences” (p. 170). A client project, however, is rife with ethical implications of free labor that

primarily serve the interests of future employers. Bushneil (1999) provided an elaborate critique of using the classroom for corporate training grounds: “Not only have we generally accepted and worse, internalized corporate paradigms, but I’m also worried that even now we’re allowing them to subvert our own best revisionist and subversive insights” (p. 175). He goes onto argue:

Our job is not to be *like* the corporation; that does no one any good, despite short-term perceptions by some in industry and (apparently) a great many in academia. Let the corporations themselves fill that role. Our job is to challenge paradigms, *any* paradigms (not just those in business), and through that challenge to teach our students how to develop intellectually as human beings. (p. 184)

While a client project may be a step up from case studies in that they carry more weight and provide access to stakeholders, businesses are improving their situation (and ideally profits) with no monetary compensation for the students. Bushneil would say, as do I, that this is unethical.

### **Service Learning**

A service learning course removes the ethical implications of free labor by creating similar relationships with non-profit organizations; these courses still draw upon the recreation of a workplace activity network, which helps to avoid the issue of pseudotransactionality voiced by Spinuzzi. Service-learning courses give students the opportunities to engage with the community while enabling the community organizations to receive the assistance they need without monetary commitment. Both service-learning and client-based projects have a glaring problem—time and resources, which becomes

even more problematic for junior faculty who may not have the necessary connections in the community to establish such pedagogical practices (Stephens, 2016). Despite the opportunity to work towards civic engagement with these non-profits, yet another risks arises mentioned earlier—hyperpragmatism. In his essay, Scott (2004) cautions instructors turning to client based projects, whether those clients are corporations or non-profit organizations: “As an ideology and set of practices, hyperpragmatism is primarily concerned with helping students understand and successfully adapt to the writing processes, conventions, and values of disciplinary and workplace discourse communities” (p. 291). Hyperpragmatism, he argues, focuses so much on producing artifacts for clients that teachers and students miss the opportunity to critically engage with questions of knowledge production, power relationships, and limits the scope and potential of practices for students.

### **Other General Critiques of TPC Pedagogies**

As instructors draw on these different types of pedagogies, others have called for more radical approaches to the technical writing classroom. Herndl (1993) offered a critique of TPC pedagogy, claiming that traditional practices only lead students to conformity while reifying the social and cultural power structures that dominate the mindset of “preparing students for careers” (p. 360). He argues that the TPC classroom has much more potential to teach a rhetoric of dissention, echoing the work of Trimbur (1989) and Freire (1970). In Herndl’s radical pedagogy, he encourages instructors to teach students to recognize how their actions reproduce power relationships, hoping that this recognition will challenge students’ thinking that “that’s just how things are.” They

aren't just how things are. Technical communicators shape knowledge production, they don't just communicate it. For Herndl, if an instructor isn't teaching a radical pedagogy of rhetorical dissention, they are reinforcing those power dynamics: "There is no such thing as a neutral education" (p. 351). Bryan (1992) adds his critique in one simple line: "We cannot teach integrity, but we can teach the right questions" (p. 87). What questions we train students to ask depends on the stakeholders we center in our classroom. Schriver (2013) discusses the role of information design in technical communication by focusing on the relationships technical communicators have with their stakeholders: "A fundamental goal for information design is to enable and enhance relationships among stakeholders for an artifact--that is, among the variety of audiences, clients, critics, readers, listeners, users, and viewers who have a stake in the content" (p. 388). If the stakeholder is a corporate client, an instructor risks hyperpragmatism. If the stakeholder are those who are impacted by artifacts but are generally marginalized, then students can begin to critically question those power relationships alluded to by Herndl and Bryan.

In his discussion of the TPC classroom, Bushneil (1999) writes: "In the university, the mission is (or should be) to teach critical thinking, to encourage our students to be skeptics and to question what they perceive as 'authorities' (including us) both inside and outside the academy" (p. 184). A university's mission, especially in the humanities, is to help students understand the complexities of language and communication. Instead of leaving the university with knowledge of industry tools (which are changing constantly anyways), students should be, as Bushneil describes, "savvy and questioning thinkers rather than simply as efficient, problem-solving doers"

(p. 175). According to Johnson-Eilola (1996), one way that we can train students to become these savvy, questioning thinkers is by relocating the value of work in technical communication. He argues that we should develop what he calls symbolic-analytic workers:

Symbolic-analytic workers rely on skills in abstraction, experimentation, collaboration, and system thinking to work with information across a variety of disciplines and markets . . . while not losing sight of the larger rhetorical and social contexts in which users work and live. (p. 245)

This shifts the focus of value from creating those mundane artifacts to understanding how the information and symbolic nature of those mundane artifacts have real impact in a complex world. In other words, shifting the focus from the mundane to the wicked.

### **PESTEL and Design Thinking**

#### **PESTEL**

This form of situational analysis is a tool used to uncover some of the external forces facing organizations (Barrington, 2016). Put into TPC language, those external forces are stakeholders. In the classroom, PESTEL can be used as a heuristic for students to understand the complicated nature of wicked problems (adapted from the Oxford College of Marketing):

- Political Factor: How do governments and government policy impact the organization?
- Economic Factors: How do factors like interest rates, employment, or unemployment rates impact the organization?

- Social Factors: How does the organization impact and is impacted by socio-cultural issues like race, gender, and sexuality?
- Technological Factors: How does available technology (or technology not yet invented) impact the organization?
- Environmental Factors: How does an organization impact and is impacted by the environment?
- Legal Factors – What is an organization legally allowed to do/not do?

While the original focus of PESTEL is about the external forces on an organization, this approach also considers the organization's impact on these factors, acting as an external force itself. One particular reason for turning to PESTEL as a heuristic is as much a reminder to myself as an instructor as it is for my students. In the past two decades—and especially in the past five years—TPC has made a strategic shift towards social justice. This, I feel, is a good thing; however, to limit an analysis to *just* socio-cultural issues is a disservice to our students. Yes, these issues are paramount, but PESTEL enables students to understand just how complicated those social issues are when weaved together with other issues. Wicked problems are wicked because they impact and are impacted by all of these factors—not just one.

### **Design Thinking**

In a technical communication classroom, principles of design enable instructors to shift from the hyperpragmatic to the symbolic-analytic worker. Below are some of the principles of design thinking already manifested in existing technical communication scholarship (Pope-Ruark, Moses, Conner, & Tham, 2017). The purpose of drawing



attention to these manifestations is to offer a heuristic for designing a course. While the previous pedagogical perspectives have advantages and disadvantages, each can be altered and revised to include these principles in an effort to create symbolic-analytic workers who question knowledge production in an effort to disrupt dominant narratives whose foundations are position, privilege, and power over marginalized individuals (see Jones, Moore, & Walton, 2016).

**Acceptance of ambiguity.** Perhaps the fluidity of the field of technical communication might be linked to how technical communicators understand the way language functions. Certain phrases have different meanings (or no meanings at all) in different cultural contexts. Words have meanings beyond their denotations to an ever-growing list of connotations. Not only are words ambiguous, but so are the symbols that surround them: font choices, color, organization, and other stylistic factors. As the world moves towards globalization, these factors become ever more prevalent when communicating across cultures. Agboka (2012) cautions that a lack of awareness and acceptance of ambiguity risks marginalizing subcultures into larger cultural identities. A field that rests on the uncertainty and ambiguity of language and symbols might sound unstable, but this instability is what TPC does. Recalling Mehlenbacher (2013), technical communicators constantly create artifacts in ever-changing environments in different mediums for different stakeholders. When this ambiguity of the mundane couples with the uncertainty of wicked problems, a symbolic-analytic worker ought to be able to quickly assess the problem not to solve it necessarily, but to see the multifaceted nature

of the problem and to hear the voices of stakeholders—especially those who are traditionally marginalized.

**Productive failure.** Part of helping students understand the ambiguity of language and symbols involves trial and error. For many students, coming to university will be the first time they've had their own language and symbol systems questioned and challenged. One way to encourage students to challenge their own systems of knowledge—which will enable them to challenge other systems of knowledge—is by providing a safe space for failure. Even more than a safe space, we ought to provide an active pursuit of failure. This pursuit of failure is argued most clearly in Juul's (2013) book, *The Art of Failure*, where he turns to gaming to understand how individuals try new methods over and over until they succeed: "Failure brings about something positive, but it is always potentially painful or at least unpleasant" (p. 9). Pursuing failure not only seeks a paradigm-shift from a student's perspective but an instructor's perspective as well. Hyperpragmatism has sunk its claws in the classroom, and the corporate ideologies they reproduce demand success. Instructors need to disrupt this ideology by redefining what it means to fail. In other circles, this is called "ludic pedagogy": "Ludic creativity is playful, fearless, and awe-some. Teachers must not only tear down assumptions that may serve to oppress, but must also inject creativity back into the classroom" (Morris, 1999, p. 422). This alludes to a pedagogical principle I hold dear—classes should be fun. Classes should be a place where students can safely fail in order to learn—embracing failure, learning from it, and moving on. An active pursuit of failure recognizes values what a student learns what *not* to do rather than just what they should do. This, of course, means

an instructor must give time for students to fail, which means iterative projects in the classroom where students cannot be assessed based on final “success” but productive failure.

**Radical collaboration.** Given the ambiguity of language, the word “radical” has various meanings. Radical in one sense is something drastically different than what has come before it. As an example, Walton, Zraly, & Mugengana (2015) recount their interdisciplinary, international collaborative project in Rwanda, practicing what Agboka (2013) calls “participatory localization,” which is “user involvement, not as isolated user participation but as user-in-community involvement and participation in the design phase of products” (p. 42). This type of collaboration is not easy: “The complexity of community-based research is an unavoidable and even productive factor in conducting ethical research that upholds values of collaboration and shared power and for conducting rigorous research that is produced, shaped, analyzed, and presented with local partners” (Walton, Zraly, & Mugengana, 2015, p. 62). Not only did the two American researchers, Walton and Zraly, work with local subjects, but they published the article with their local contact, Mugengana, who guided them through the streets and cultures of Rwanda. This collaboration resonates with Agboka’s call for participatory localization where he encourages researchers to move beyond language translation in international projects, but translation of cultural factors as well (i.e. local customs, laws, values, medical conditions, economic issues, and sexual standards). This example, however, is one of conducting research, not necessarily teaching it in the classroom. Trips for entire classes across the world to perform participatory localization would be costly to say the least. Shifting the

perspective from hyperpragmatism to symbolic-analytic workers may not be radically new, but the outcomes are certainly radical. To disrupt systems of power calls for radical thinkers to take risks.

**Bias towards action.** Radical collaboration for change might be met with hesitancy from students and even instructors. Taking on powerful ideologies is not an easy task, and action is always implied. Just as a classroom ought to embrace ambiguity, it ought to embrace a bias towards action. As Katz (1992) discusses, all technical communication concerns itself with deliberative rhetoric, which always leads to action: “All deliberative rhetoric is concerned with decision and action. Technical writing, perhaps ever more than other kinds of rhetorical discourse, always leads to action, and thus always impacts on human life” (p. 259). Which actions should be taken is a different matter; however, by embracing ambiguity, failure, and radical collaboration, students can best prepare themselves to be symbolic-analytic workers rather than hyperpragmatists. Jones, Moore, and Walton (2016) call for researchers to “unabashedly embrace social justice and inclusivity” (p. 212), and if that begins in the classroom, those tendencies towards action can take root before students enter the workplace. In his critique of the political-ethical implications of TPC as a practice, Sullivan (1990) contends:

[T]hat teaching standardized formats and forms means teaching the technological mindset, and, thus, enculturating students into the military-industrial complex.

This conclusion further suggests that we implicitly accept present restrictions on public discourse about technology and fail to give students power to engage in social action. (p. 377)

Rather than defining technical communication as an art or skill—*techné*— Sullivan advises instructors to move towards a practice—*praxis*. Social action requires more than an understanding that technical communication is merely *praxis* over *techné*. To use practice as a bias towards action requires practical wisdom—*phronesis*: “a virtue defined as the ability to reason about ends rather than means. *Phronesis*,” Sullivan writes, “enables a person to deliberate about the good rather than the expedient and, as such, to act in the political sphere rather than the sphere of work” (1990, p. 378). Although Sullivan’s emphasis on the “ends rather than the means” may seem to counter the idea of productive failure and ludic pedagogy, it actually isn’t. What needs to be defined here are what ends, and in the classroom those ends—ultimately—are learning. By focusing on *phronesis* as practical wisdom, students can begin to see how the mundane artifacts they produce in their future profession contribute to wicked problems—for better or for worse.

### **The Course Design**

As an alternative pedagogical design, this course embeds these concepts of wicked problems, mundane artifacts, design thinking, and PESTEL. As I mentioned previously regarding productive failure, designing a class that is fun helps to engage with students—especially in a class for multi-majors. Many of the students who attend my Introduction to Technical Writing course come from several different majors seeking to take my class to fulfil their major’s writing requirement. At the beginning of each semester, I ask my students how many of them saw this class on their schedule and burst with ecstasy. Few, if any, raise their hands. In designing the class, I wanted my students to engage in a wicked problem that would impact each and every one of them that would

be fun and engaging—a zombie apocalypse (for more examples, see “Zombie Preparedness”). This section of the article outlines the structure of the course, class projects, and a method of assessment combining the principles of design-thinking with Cargile Cook’s six layered literacies.

### **Structure of the Course**

**Part one.** At the beginning of the course, I divide the students into six groups using PESTEL. As part of homework for the class, students watch the first seasons of *The Walking Dead* and *Fear the Walking Dead*. Together, the members of each group write down as many instances of their PESTEL area as possible, ready to come to class to discuss/defend their decisions. As the semester moves on and the class watches more episodes, each group becomes a subject-matter expert (SME) in their area as it pertains to a zombie apocalypse. Students are also encouraged to make explicit connections with their notes, their PESTEL area, and their home disciplines/majors. Below are some examples of what students might note from *The Walking Dead*, S01E03, “Tell It to the Frogs”:

<b>Political</b>	Patriarchy has become even more prevalent with the male characters taking leadership roles.	This patriarchy is challenged when Rick joins Shane's group.
<b>Economic</b>	Certain items (like guns, ammo, tools and food) have exponentially higher value. Rick risks his life and the lives of three others in order to recover a bag of guns and tools.	As the monetary system collapses, members of the group are forced to barter. For example, when Rick negotiates with Dale for his bolt cutters, he offers him a gun and scrap parts from the van.
<b>Socio-cultural</b>	Glenn, an Asian character, suggests that Rick, a white character, tell Daryl that his brother (both also white) has been handcuffed to a roof by T-Dog, a black character.	After seeing two male characters hunting playfully for frogs, the women in the show begin questioning the division of labor.
<b>Technological</b>	A vehicle is seen not as a single object but a source for parts.	Rick relies on the weak frequencies of walkie-talkies in order to try to maintain contact with Morgan.
<b>Environmental</b>	The group sought out high ground surrounded by mountains in order to maintain safety.	As the group realizes that the zombies are no longer human, they become part of the natural environment as something to be dealt with.
<b>Legal</b>	Glenn hotwires a sports car and uses it as a distraction for the others to get away. What was once a crime (grand theft auto) is now necessary for survival.	Rules and laws have been established in the camp, like when Shane tells Ed that adding more wood to the fire is against the rules.

**Part two.** As students collaborate with each other, they become SME in their areas, which is important for the next phase of the class. Students do not stay in their PESTEL groups for the duration of the class. Halfway through the semester I rearrange the groups, each including a separate SME. Each student as a SME is now an advocate on behalf of their PESTEL area. Together, these different SME's must navigate their respective interests while writing a \$1.5 million grant proposal to the CDC for the study and prevention of the zombie apocalypse.

### **Class Projects**

One particular challenge to teaching a class of multi-majors is selecting deliverables that fit across several disciplines. On one hand, I want my students to understand the basic conventions of genre, but I also want to teach it in a way that promotes action. This is especially important to consider given Sullivan's caution against teaching genre as form: "Unfortunately, genres in technical discourse seem to preclude the opportunity for citizens to speak simply as citizens" (p. 377). With Johnson-Eilola's (1996) call for symbolic-analytic workers, students can see the connections between genres, rhetorical situations, and a self-awareness as citizens. Although there are several options when it comes to selecting deliverables, I chose to focus on three genres that, when used intentionally, may be used to enact change.

**Poster.** With a focus on visual communication, students are asked to produce a poster that they might see on a campus bulletin board similar to what the health center might release. This project encourages students to take caution against a potential infection without alarming people into a panic. To give this project context in a non-



zombie reality, the class discusses the 2015 Zika virus outbreak. Students find posters from various regions, analyzing each from their various PESTEL categories. This particular genre works well for the ludic pedagogy discussed in the productive failure section. Many of these multi-major students have little experience in visual design; beginning with sketches and different software, I ask students to make different posters that match different tones: staying calm, frightening to action, humor, etc.

**White paper.** A traditional genre in TPC pedagogy. While students watch episodes of *The Walking Dead* and *Fear the Walking Dead*, they are also collectively producing an annotated bibliography relating to their respective PESTEL groups and existing global epidemics that might relate to how an organization might approach a zombie apocalypse. To avoid Spinuzzi's pseudotransactionality, the audience for their white papers are their groups. While focusing on a single facet of their PESTEL category, the purpose of the white paper is to explain how they plan to advocate on behalf of their group's interests.

**Grant proposal.** After the groups have been reorganized, members of the new groups must advocate on behalf of their stakeholders. With competing interests, students must learn to listen and compromise in order to write a successful grant. As far as genres, the grant proposal puts action at the forefront, but not just one type of action—actions that meet the needs of multiple stakeholders.

### **Assessment using the Six Layered Literacies**

Cargile Cook's (2002) framework provides a fantastic, nuanced approach for assessment, focusing on six literacies:

- 1) Basic: The ability to read and write
- 2) Rhetorical: Understanding the multifaceted nature of communication
- 3) Social: Working with others in a collaborative environment
- 4) Technological: Knowing how to use tools and which ones to use given context
- 5) Ethical: Considering all possible stakeholders, especially marginalized stakeholders
- 6) Critical: Recognizing ideological biases and power structures and a willingness to take action for those marginalized

Rather than rely solely on these literacies for assessment, I created a matrix that weaves them with the principles of design thinking in order to provide critical questions that instructors can use to ask students and for students to ask themselves through the semester projects. While these questions are not exhaustive, they offer a foundation to which teachers can turn to assess these projects.

	Acceptance of Ambiguity	Productive Failure	Radical Collaboration	Bias towards Action
Basic Literacy	Whose basic literacy is being measured? How do other groups see and use language aside from "Standardized English"?	Read your work out loud to each other. How often do you stumble over your words? Is it because they are too complicated?	Have you asked anyone to look at your work before you turned it in?	How might you create the same document to a marginalized community?

Rhetorical Literacy	Which audience is being addressed? Are any of my own ways of thinking contradictory? Does my audience face similar contradictions?	What would this look like if addressed to a CEO? A family member? A teacher?	What can the life experiences of your group members teach you about your own ways of thinking about the world?	What do you want your audience to do when they finish engaging with your deliverable?
Social Literacy	In group settings, how are you most productive and least productive? How can you and your group work to make this project applicable outside of this class?	How have you and your group resolved conflict? What have you learned from where things went wrong in the past?	Have you taken the time to understand the perspective of your group members? Should you ask people outside of your group for input?	How can you encourage your fellow students to make a compelling argument towards action?
Technological Literacy	What various kinds of software will enable you to complete this project? Which works best?	How many of those different programs did you try to use?	Did you go to someone for help? Did you search it on the internet? Ask a friend? Go to a workshop?	Which technology is most accessible to others in case they need to alter it in the future?
Ethical Literacy	When faced with ambiguity, how do you resolve to move forward? What systems of thinking are in place that made you resolve it that way?	When have you made ethical mistakes or misunderstandings? What have you learned from them?	What are your group members' approaches to ethical considerations? How can their views enhance your own?	When faced with an ethical dilemma, what should you consider in contemplating how to move forward?

Critical Literacy	What structures of power are in place? How much do you benefit from those power structures? How do you resolve those privileges?	When have you been subjected to a power dynamic? How did you overcome it?	What can others teach you about power structures that you may not know exist?	With an understanding of these power structures, what do you plan to do to take action towards them?
-------------------	----------------------------------------------------------------------------------------------------------------------------------	---------------------------------------------------------------------------	-------------------------------------------------------------------------------	------------------------------------------------------------------------------------------------------

### **Conclusion**

Several pedagogies exist in TPC, and all of them have one thing in common—helping students to learn. While many of those pedagogies engage with teaching TPC to students who want to become technical and professional communicators, teaching an Introduction to Technical Writing is a completely different course when teaching multi-majors. With a desire to teach the conventions and principles they need to be successful communicators, I wanted to design a course that merges design-thinking, PESTEL, and the two pillars inextricably connected in TPC—wicked problems and mundane artifacts. Due to the fluidity of our field, TPC instructors are in a unique position to make students aware of the multi-faceted nature of wicked problems and how the writing we do impacts those problems on several levels: political, economic, socio-cultural, technological, environmental, and legal issues. No matter what field these students enter, each of them will write. As symbolic-analytic workers who have developed their six layered literacies, students will be able to become more aware of those ambient, mundane artifacts that impact their lives. More than that, however, students will know that their writing has impact.

## References

- Agboka, G. (2012). Liberating intercultural technical communication from “large culture” ideologies: Constructing culture discursively. *Journal of Technical Writing and Communication*, 42(2), 159–181. <https://doi.org/10.2190/TW.42.2.e>
- Agboka, G. Y. (2013). Participatory localization: A social justice approach to navigating unenfranchised/disenfranchised cultural sites. *Technical Communication Quarterly*, 22(1), 28–49. <https://doi.org/10.1080/10572252.2013.730966>
- Barrington, R. (2016, June 30). What is a PESTEL analysis? Retrieved March 19, 2018, from <https://blog.oxfordcollegeofmarketing.com/2016/06/30/pestel-analysis/>
- Bryan, J. (1992). Down the slippery slope: Ethics and the technical writer as marketer. *Technical Communication Quarterly*, 1(1), 73–88. <https://doi.org/10.1080/10572259209359492>
- Bushneil, J. (1999). A contrary view of the technical writing classroom: Notes toward future discussion. *Technical Communication Quarterly*, 8(2), 175–188. <https://doi.org/10.1080/10572259909364658>
- Cargile Cook, K. (2002). Layered literacies: A theoretical frame for technical communication pedagogy. *Technical Communication Quarterly*, 11(1), 5–29. [https://doi.org/10.1207/s15427625tcq1101\\_1](https://doi.org/10.1207/s15427625tcq1101_1)
- Fear the walking dead. (2015). AMC.
- Freire, P. (1970). *Pedagogy of the oppressed*. New York: Seabury Press.
- Herndl, C. G. (1993). Teaching discourse and reproducing culture: A critique of research and pedagogy in professional and non-academic writing. *College Composition and Communication*, 44(3), 349–363. <https://doi.org/10.2307/358988>
- Johnson-Eilola, J. (1996). Relocating the value of work: Technical communication in a post-industrial age. *Technical Communication Quarterly*, 5(3), 245–270. [https://doi.org/10.1207/s15427625tcq0503\\_1](https://doi.org/10.1207/s15427625tcq0503_1)
- Johnson-Eilola, J., & Selber, S. A. (2013). Introduction. In *Solving problems in technical communication* (pp. 1–14). Chicago, IL: University of Chicago Press.
- Jones, N. N., Moore, K. R., & Walton, R. (2016). Disrupting the past to disrupt the future: An antenarrative of technical communication. *Technical Communication Quarterly*, 25(4), 211–229. <https://doi.org/10.1080/10572252.2016.1224655>

- Juul, J. (2013). *The art of failure: An essay on the pain of playing video games*. MIT Press.
- Katz, S. B. (1992). The ethic of expediency: Classical rhetoric, technology, and the Holocaust. *College English*, 54(3), 255–275. <https://doi.org/10.2307/378062>
- McNely, B. (2012). Big data, situated people: Humane approaches to communication design. *Communication Design Quarterly Review*, 1(1), 27–30. <https://doi.org/10.1145/2448917.2448923>
- Mehlenbacher, B. (2013). What is the future of technical communication? In J. Johnson-Eilola & S. A. Selber (Eds.), *Solving problems in technical communication* (pp. 187–208). Chicago, IL: University of Chicago Press.
- Morris, M. (1999). Toward a ludic pedagogy: An uncertain occasion. *Counterpoints*, 70, 412–424.
- Pope-Ruark, R., Moses, J., Conner, T., & Tham, J. (2017). Call for papers: Design-thinking approaches in technical and professional communication. *Journal of Business and Technical Communication*, 31(4), 520–522.
- Pringle, K., & Williams, S. D. (2005). The future is the past: Has technical communication arrived as a profession? *Technical Communication*, 52(3), 361–370.
- Rickert, T. J. (2013). *Ambient rhetoric: The attunements of rhetorical being*. Pittsburgh, Pa: University of Pittsburgh Press.
- Rittel, H. W. J., & Webber, M. M. (1973). Dilemmas in a general theory of planning. *Policy Sciences*, 4, 155–169.
- Rivers, N. A., & Weber, R. P. (2011). Ecological, pedagogical, public rhetoric. *College Composition and Communication*, 63(2), 187–218.
- Rozumalski, L. P., & Graves, M. F. (1995). Effects of case and traditional writing assignments on writing products and processes. *Journal of Business and Technical Communication*, 9(1), 77–102. <https://doi.org/10.1177/1050651995009001005>
- Schriver, K. (2013). What do technical communicators need to know about information design? In J. Johnson-Eilola & S. A. Selber (Eds.), *Solving problems in technical communication* (pp. 386–427). Chicago, IL: University of Chicago Press.

- Scott, J. B. (1995). Sophistic ethics in the technical writing classroom: Teaching nomos, deliberation, and action. *Technical Communication Quarterly*, 4(2), 187–199. <https://doi.org/10.1080/10572259509364596>
- Scott, J. B. (2004). Rearticulating civic engagement through cultural studies and service-learning. *Technical Communication Quarterly: TCQ; Oxford*, 13(3), 289–306.
- Selfe, R. J., & Selfe, C. L. (2013). What are the boundaries, artifacts, and identities of technical communication? In J. Johnson-Eilola & S. A. Selber (Eds.), *Solving problems in technical communication* (pp. 19–49). Chicago, IL: University of Chicago Press.
- Spinuzzi, C. (1996). Pseudotransactionality, activity theory, and professional writing instruction. *Technical Communication Quarterly*, 5(3), 295–308. [https://doi.org/10.1207/s15427625tcq0503\\_3](https://doi.org/10.1207/s15427625tcq0503_3)
- Stephens, E. J. (2016, April). *Advocating for the inexperienced: Teaching social justice and citizenship with limited time, resources, and experience*. Adobe Spark Page presented at the Association of Teachers of Technical Writing, Houston, TX. Retrieved from <https://spark.adobe.com/page/NW9HL/>
- Sullivan, D. L. (1990). Political-ethical implications of defining technical communication as a practice. *Journal of Advanced Composition*, 10(2), 375–386.
- The walking dead. (2010). AMC.
- Trimbur, J. (1989). Consensus and difference in collaborative learning. *College English*, 51(6), 602–616. <https://doi.org/10.2307/377955>
- Walton, R., Zraly, M., & Mugengana, J. P. (2015). Values and validity: Navigating messiness in a community-based research project in Rwanda. *Technical Communication Quarterly*, 24(1), 45–69. <https://doi.org/10.1080/10572252.2015.975962>
- Wickman, C. (2014). Wicked problems in technical communication. *Journal of Technical Writing and Communication*, 44(1), 23–42. <https://doi.org/10.2190/TW.44.1.c>
- Zombie preparedness. (2017, June 15). Retrieved March 21, 2018, from <https://www.cdc.gov/phpr/zombie/index.htm>

APPENDIX A  
PYTHON CODE

```
# -*- coding: utf-8 -*-
"""
Created on Fri Mar  9 20:19:51 2018

@author: bwebster
"""
from PyPDF2 import PdfFileReader, PdfFileWriter
from os import listdir
from os.path import isfile, join
from tika import parser
import os
import re

filenames = [f for f in listdir('C:\\Users\\bwebster\\Downloads\\PDF -
The Prison Journal\\PDF - The Prison Journal') if
isfile(join('C:\\Users\\bwebster\\Downloads\\PDF - The Prison
Journal\\PDF - The Prison Journal', f))]

filenames

i = 0

for doc in filenames:
    with open(str(doc), 'rb') as infile:
        try:
            reader = PdfFileReader(infile)
            if reader.isEncrypted:
                print(doc)
            else:
                x = reader.numPages
                os.mkdir('C:\\Users\\bwebster\\Downloads\\PDF - The
Prison Journal\\PDF - The Prison Journal\\SinglePageFolder\\' +
str(doc)+'_files')
                out_dir = 'C:\\Users\\bwebster\\Downloads\\PDF - The
Prison Journal\\PDF - The Prison Journal\\SinglePageFolder\\' +
str(doc)+'_files'
                print(i)
                i = i+1
                for page in range(x):
                    writer = PdfFileWriter()
                    writer.addPage(reader.getPage(page))
                    fn = 'output'+str(page)+'.pdf'
                    with open(os.path.join(out_dir,fn), 'wb') as
outfile:
                        writer.write(outfile)
        except:
            pass

doc_files = [f for f in os.walk('C:\\Users\\bwebster\\Downloads\\PDF -
The Prison Journal\\PDF - The Prison Journal\\SinglePageFolder')][0][1]

k = 0
```



```

for folder in doc_files:
    path = os.path.join('C:\\Users\\bwebster\\Downloads\\PDF - The
Prison Journal\\PDF - The Prison Journal\\SinglePageFolder\\', folder)
    singlepages = [g for g in os.walk(path)][0][2]
    for document in singlepages:
        if k%100 == 0:
            print(k)
            k = k+1
            text = parser.from_file('C:\\Users\\bwebster\\Downloads\\PDF -
The Prison Journal\\PDF - The Prison
Journal\\SinglePageFolder\\'+str(folder)+'\\'+str(document))
            if isinstance(text['content'], str):
                text_parsed = re.sub(' +', ' ',
re.sub(r'[\^A-Za-z0-9.]', ' ',
re.sub("(\\w[a-z])([A-Z])", r"\\1
\\2",
re.sub('-', ' ',
text['content'])))
                ).rstrip().lstrip().lower()
            else:
                text_parsed = 'hi eric'
                txtpath = str(document) + '.txt'
                out_dir = 'C:\\Users\\bwebster\\Downloads\\PDF - The Prison
Journal\\PDF - The Prison Journal\\SinglePageFolder\\' + str(folder)
                with open(os.path.join(out_dir,txtpath),"w") as text_file:
                    text_file.write(text_parsed)

```

```

..... # -*- coding: utf-8 -*-

```

Created on Sat Mar 10 13:16:42 2018

@author: bwebster

```

from tika import parser
import os
import re
from nltk.tokenize import sent_tokenize
from nltk.tokenize.punkt import PunktSentenceTokenizer, PunktParameters
from nltk.tokenize import word_tokenize
import pandas as pd
import random
import csv

```

```

doc_files = [f for f in os.walk('C:\\Users\\bwebster\\Downloads\\Manuals
Small\\SinglePageFolder')][0][1]

```

```

output = {}

```

```

for folder in doc_files:
    path = os.path.join('C:\\Users\\bwebster\\Downloads\\Manuals
Small\\SinglePageFolder\\', folder)
    singlepages = [g for g in os.walk(path)][0][2]

    output[folder] = {}
    for document in singlepages:
        if str(document)[-4:] == '.txt':
            with open(os.path.join(path,str(document))) as f:

```

```

        txt = f.readlines()
        punkt_param = PunktParameters()
        punkt_param.abbrev_types =
set(['mr', 'mr.', 'jr', 'jr.', 'sr', 'sr.', 'e.g', 'e.g.', 'i.e', 'i.e.', 'www.',
'www', 'ex.', 'ex', '1', '1.', '2', '2.', '3', '3.', '4', '4.'])
        sentence_splitter = PunktSentenceTokenizer(punkt_param)
        sents=sentence_splitter.tokenize(txt[0])
        output[folder][document]=sents

out_arr = []
k=0
for oput in output.keys():
    doc_val = output[oput]
    for flder in doc_val.keys():
        pg_val = doc_val[flder]
        for idx, sent in enumerate(pg_val):
            state = oput[:2]
            page_sub = flder[:-8]
            page = page_sub[6:]
            empty = ''
            outrow = [k,oput,state,page,idx,empty,empty,empty,sent]
            out_arr.append(outrow)
            k+=1

out_df =
pd.DataFrame(out_arr,columns=['uid','document','state','page_number','s
entence_number',
'is_punishment','is_reward','is_background','sentence'])
# sample = out_df.sample(150)

out_df.to_csv('C:\\Users\\bwebster\\Downloads\\Manuals
Small\\sample\\sample1.csv', index = False)

# -*- coding: utf-8 -*-
"""
Created on Sat Mar 31 15:50:26 2018

@author: bwebster
"""

# -*- coding: utf-8 -*-
"""
Created on Tue Jul  4 19:44:12 2017

@author: ben
"""

import random, time, sys, hashlib, re, os, operator, pickle
import pandas as pd
from gensim.models import word2vec
from collections import Counter

```

```

import definitions_v5_kt_recommendation as ktr
from nltk.corpus import stopwords

start_time = time.time()

##### System Arguments #####

# vec_size = int(sys.argv[1])
vec_size = 10

#####

input_file = 'C:\\Users\\bwebster\\Downloads\\Inmate Handbook\\Inmate
Handbooks\\PDFs\\sample1.csv'

df = pd.read_csv(input_file)

## train the bigrams tool and the model
w2v, ngrams = {}, {}
ngram_data = []

# create a list to fill with sentences for this job class to train on
# start the dictionary for the job class
##ngrams[jc] = {}

fixed_sentences = [sent.split(' ') for sent in df.sentence.tolist()]

# train the bigrams
bigrams = ktr.LEARN_NGRAMS(fixed_sentences)

# fit the bigrams
bigrams_corpus = []
for sentence in fixed_sentences:
    bigrams_corpus.append(bigrams[sentence])

# train the trigrams
trigrams = ktr.LEARN_NGRAMS(bigrams_corpus)

# fit the trigrams
trigrams_corpus = []
for sentence in bigrams_corpus:
    trigrams_corpus.append(trigrams[sentence])

# train the quadgrams
quadgrams = ktr.LEARN_NGRAMS(trigrams_corpus)

# fit the quadgrams
quadgrams_corpus = []
for sentence in trigrams_corpus:
    quadgrams_corpus.append(quadgrams[sentence])

print('Training the word2vec model', ("--- %s minutes ---" %
round((time.time() - start_time)/60,2)))

# create the word2vec model
model = word2Vec(workers=4, size=vec_size, min_count=3, window=7,
sample=.01, iter=10)

```

```

# build the vocabulary
model.build_vocab(quadgrams_corpus)

# identify the learning rates
alpha, min_alpha, passes = (0.025, 0.001, 20)
alpha_delta = (alpha - min_alpha) / passes

# keep a copy for shuffling
_docs = quadgrams_corpus[:]

# learning loop
for epoch in range(passes):

    # reshuffle before presentation
    random.shuffle(_docs)

    # train
    model.alpha, model.min_alpha = alpha, min_alpha
    model.train(_docs, total_examples=model.corpus_count,
epochs=model.iter)

    alpha -= alpha_delta

out_arr = []
for idx, row in df.iterrows():
    words = row.sentence.split(' ')
    bg_words = bigrams[words]
    tg_words = trigrams[bg_words]
    qg_words = quadgrams[tg_words]

    for k,v in Counter(qg_words).items():
        k = re.sub('[!@#$.]', '', k)
        if k in model.wv.index2word:
            if k not in stopwords.words('english'):
                vec = model[k]
                outrow = [row.document,k,v] + vec.tolist()
                out_arr.append(outrow)

out_cols = ['document', 'token', 'freq'] + ['v'+str(i+1) for i in
range(10)]
out_df = pd.DataFrame(out_arr, columns=out_cols)

out_df.to_csv('C:\\Users\\bwebster\\Downloads\\Inmate Handbook\\Inmate
Handbooks\\PDFs\\sample2.csv')

```