Disruption, Remix, and Mediation: Analyzing "Bot" Algorithm Case Studies to Better Understand Rhetorical Agency in the Digital Age

Aaron Geiger
butifandthat@gmail.com

Follow this and additional works at: https://huskiecommons.lib.niu.edu/allgraduate-thesesdissertations

Part of the Rhetoric Commons

Recommended Citation
https://huskiecommons.lib.niu.edu/allgraduate-thesesdissertations/7059

This Dissertation/Thesis is brought to you for free and open access by the Graduate Research & Artistry at Huskie Commons. It has been accepted for inclusion in Graduate Research Theses & Dissertations by an authorized administrator of Huskie Commons. For more information, please contact jschumacher@niu.edu.
ABSTRACT

DISRUPTION, REMIX, AND MEDIATION: ANALYZING “BOT” ALGORITHM CASE STUDIES TO BETTER UNDERSTAND RHETORICAL AGENCY IN THE DIGITAL AGE

Aaron Geiger, Ph.D.
Department of English
Northern Illinois University, 2020
Jessica Reyman, Ph.D., Director

Algorithms are ubiquitous components of digital age rhetoric and communication. Although they are integrated within our daily online functions, shaping our discourse and interactions with one another and with other algorithms, little is known about how they shape rhetorical agency. Most of our online communication occurs through social media on platforms such as Facebook and Twitter, and we rely on social spaces like Wikipedia for updated information that is shaped by social constructs. We presume that since we program algorithms, their communicating functions mimic human discourse and can be shaped similarly.

However, a rich text analysis of three different case studies looks at how algorithms vary in how they alter rhetorical agency, shaping discourse, power structures, and interactions in ways we have not seen before. Through this analysis and visualized with ontographs – geographic maps that show the distribution of actant interactions within an environment – the distribution of agency is traced between actants and compared to more traditional systems of human-to-human communication to tease out how algorithms behave differently.
This research examines the case study of the Facebook [loveMachine], an “art bot” that was created to provoke discussion about what the Facebook “like” button really means; a Twitter botnet dubbed the “Peñabots” that shaped a political election and presidency in Mexico; and the Wikipedia Addbot, which automated tens of thousands of “reverts” on the platform. Each case study offers a unique look at the rhetorical function of the social media bot algorithms within different platforms.
DISRUPTION, REMIX, AND MEDIATION: ANALYZING “BOT” ALGORITHM CASE STUDIES TO BETTER UNDERSTAND RHETORICAL AGENCY IN THE DIGITAL AGE

BY

AARON GEIGER
©2019 Aaron Geiger

A DISSERTATION SUBMITTED TO THE GRADUATE SCHOOL IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

DEPARTMENT OF ENGLISH

Doctoral Director:
Jessica Reyman
ACKNOWLEDGEMENTS

I would like to thank the following people for their time, feedback, and resources found within these pages:

Casey Boyle, Ph.D., University of Texas, Austin
James Brown, Jr., Ph.D., Rutgers University, Camden
Julien Deswaef
Charles M. Grimm, Georgia Highlands College
Aaron Halfaker, Ph.D., Wikimedia Foundation
Matthew Plummer-Fernandez, Ph.D.
Adam Shorland, Wikimedia Deutschland
DEDICATION

To my wife, Dr. Sarah Dee Geiger, and children Eva Anuhea, Jack Malae, Wynn Wailani, Atticus Aukai, and Willa Wai`oli. We live and love life wide and deep. Sarah, thank you for being in my corner all the way.

To my mother, Kathleen Pauahi, thank you for your patient love and for showing me the magic of books. Of all the gifts you’ve given me, those two will sustain me to the very end.

To my father, Stephen Michael, for your hard work and sacrifice, and for instilling in me an honor code and sense of moral goodness.

To Jessica Reyman, Michael Day, and David Gunkel, for your expertise and mentorship. My gratitude can never be fully expressed, but I promise I will pay it forward.

Finally, to the many professors, staff members, colleagues, and my own students who helped me along in this journey, this is a tribute to you and your contributions to my world of knowledge and experience.
Table of Contents

LIST OF FIGURES ............................................................................................................................................ vi

CHAPTER ONE: RHETORICAL AGENCY IN SOCIAL MEDIA BOT ALGORITHMS ................................. 1
  BACKGROUND ........................................................................................................................................... 1
  ALGORITHMS AND RHETORICAL ENTANGLEMENT .............................................................................. 8
  ALGORITHMS AS OBJECTS ...................................................................................................................... 13
  ALGORITHMS, EXPRESSION, AND ART .................................................................................................. 18
  REMIX ...................................................................................................................................................... 20
  RHETORICAL SIMULACRUM ..................................................................................................................... 22
  THE AUTOMATION OF SOCIALITY ........................................................................................................... 29
  RESEARCH QUESTIONS AND METHODS ............................................................................................... 30

CHAPTER TWO: THE FACEBOOK [LOVEMACHINE]: EXPRESSIVE AGENCY, INVENTION, AND EXPERIMENTALISM ........................................................ 38
  EARLY OBSERVATIONS ............................................................................................................................ 41
  PROGRAM DESCRIPTION AND HISTORY ............................................................................................... 49
  ACTANTS AND ACTANT CHARACTERISTICS ......................................................................................... 52
  ONTOLOGICAL MAPS ............................................................................................................................... 56
  ENVIRONMENT ....................................................................................................................................... 64
  ETHICAL CONSIDERATIONS ................................................................................................................... 66

CHAPTER 3: TWITTER PEÑABOTS: AUTOMATING MANY VOICES AS PART OF A STRATEGIC EFFORT ...................................................................................................................... 73
  EARLY OBSERVATIONS ............................................................................................................................ 73
  PROGRAM DESCRIPTION .......................................................................................................................... 77
  BOTNET LIFECYCLE .................................................................................................................................. 86
  AGENCY AND BOTS VS. BOTNETS ............................................................................................................ 87
  FROM BOT TO BOTNETS ............................................................................................................................ 94
  ACTANT CHARACTERISTICS ..................................................................................................................... 96
  ONTOLOGIES .......................................................................................................................................... 97
  ETHICAL CONSIDERATIONS .................................................................................................................... 101

CHAPTER 4: WIKIPEDIA ADDBOT: PLANNING AND GOVERNANCE VS. EXPRESSIVENESS ...... 107
  EARLY OBSERVATIONS ............................................................................................................................ 107
  PROGRAM DESCRIPTION .......................................................................................................................... 111
  CONFLICT AND AGENCY ........................................................................................................................ 112
<table>
<thead>
<tr>
<th>Chapter/Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>KINETIC VS. POTENTIAL AGENCY AND ENERGY</td>
<td>116</td>
</tr>
<tr>
<td>BOT GOVERNANCE AND AGENCY</td>
<td>118</td>
</tr>
<tr>
<td>ACTANTS AND ACTANT CHARACTERISTICS</td>
<td>118</td>
</tr>
<tr>
<td>ONTOLOGICAL MAPS</td>
<td>122</td>
</tr>
<tr>
<td>ENVIRONMENT</td>
<td>126</td>
</tr>
<tr>
<td>ETHICAL CONSIDERATIONS</td>
<td>127</td>
</tr>
<tr>
<td>CHAPTER 5: CONCLUSION: TOWARD A RESPONSIBLE FUTURE?</td>
<td>136</td>
</tr>
<tr>
<td>ALGORITHMS AND AGENCY</td>
<td>136</td>
</tr>
<tr>
<td>FINDINGS</td>
<td>139</td>
</tr>
<tr>
<td>Rhetorical Strategy</td>
<td>139</td>
</tr>
<tr>
<td>Expressive Behaviors and Meaning-Making</td>
<td>144</td>
</tr>
<tr>
<td>Examining Bots as Individuals or as Tools</td>
<td>146</td>
</tr>
<tr>
<td>Anonymity, Accountability, and Agency</td>
<td>149</td>
</tr>
<tr>
<td>A RETURN TO THE MORAL GOOD THROUGH DIGITAL LITERACY</td>
<td>153</td>
</tr>
<tr>
<td>WORKS CONSULTED</td>
<td>155</td>
</tr>
</tbody>
</table>
# List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Image from @commonsbot, a Twitter bot that collects and tweets images from the Flickr Commons Collection, tweeted to @IMG2ASCII.</td>
<td>3</td>
</tr>
<tr>
<td>2</td>
<td>@commonsbot image rendered into ASCII by @IMG2ASCII, a Twitter bot by Colin Mitchell (2014).</td>
<td>4</td>
</tr>
<tr>
<td>3</td>
<td>Tweets by van der Goot after police discovered the death threat issued by his bot.</td>
<td>10</td>
</tr>
<tr>
<td>4</td>
<td>Tweet from Twitter bot @poem_exe.</td>
<td>15</td>
</tr>
<tr>
<td>5</td>
<td>Example of nonsensical tweet from Twitter bot @poem_exe.</td>
<td>16</td>
</tr>
<tr>
<td>6</td>
<td>An algorithm's outcomes through experimentation vs. planning.</td>
<td>18</td>
</tr>
<tr>
<td>7</td>
<td>&quot;MagicRealismBot&quot;.</td>
<td>21</td>
</tr>
<tr>
<td>8</td>
<td>&quot;Token Homes,&quot; sculpture by Matthew Plummer-Fernandez, 2018. Photo used with permission of Plummer-Fernandez.</td>
<td>44</td>
</tr>
<tr>
<td>9</td>
<td>[LoveMachine]'s &quot;why?&quot; section.</td>
<td>47</td>
</tr>
<tr>
<td>10</td>
<td>The Web 2.0 Suicide Machine landing page.</td>
<td>50</td>
</tr>
<tr>
<td>11</td>
<td>Ontological map of rhetorical agency and immediate network surrounding the [Love] Machine.</td>
<td>58</td>
</tr>
<tr>
<td>12</td>
<td>Ontological map of rhetorical agency and the Facebook &quot;like&quot; button.</td>
<td>60</td>
</tr>
<tr>
<td>13</td>
<td>Sampling of &quot;Photos that I like,&quot; photo montage, during the [LoveMachine] experiment, courtesy of Julien Deswaef.</td>
<td>61</td>
</tr>
<tr>
<td>14</td>
<td>Screen grab from Iván Santiesteban’s webpage, where he linked to data he collected on bots acting to change public opinion weeks ahead of Mexico’s 2012 election day.</td>
<td>78</td>
</tr>
<tr>
<td>15</td>
<td>Screen grab from Erin Gallagher’s conference presentation video &quot;Mexican Botnet Dirty Wars.&quot; This image demonstrates hashtag attacks or &quot;poisoning,&quot; where bots attempt to disrupt a trending hashtag targeting Mexican president Enrique Peña Nieto’s nepotism.</td>
<td>81</td>
</tr>
<tr>
<td>16</td>
<td>Twitter botnet lifecycle.</td>
<td>87</td>
</tr>
<tr>
<td>17</td>
<td>Twitter user/bot @ImmoralReport (June 10, 2019).</td>
<td>89</td>
</tr>
<tr>
<td>18</td>
<td>Web tool analysis of @ImmoralReport’s tweets, tweet volume and schedule, and use of apps/services, created at <a href="http://makeadverbsgreatagain.org/allegedly">http://makeadverbsgreatagain.org/allegedly</a>.</td>
<td>91</td>
</tr>
<tr>
<td>19</td>
<td>Web tool analysis of @GrimmProspects’s tweets, tweet volume and schedule, and use of apps/services, created at <a href="http://makeadverbsgreatagain.com/allegedly">http://makeadverbsgreatagain.com/allegedly</a>.</td>
<td>91</td>
</tr>
<tr>
<td>20</td>
<td>Botometer score of Twitter user @ImmoralReport, June 11, 2019.</td>
<td>93</td>
</tr>
<tr>
<td>21</td>
<td>Trending hashtag ecosystem on Twitter.</td>
<td>99</td>
</tr>
<tr>
<td>22</td>
<td>How a botnet can effectively &quot;poison&quot; a trending hashtag.</td>
<td>100</td>
</tr>
<tr>
<td>23</td>
<td>Early stage bot ecosystem without governance guidelines.</td>
<td>121</td>
</tr>
<tr>
<td>24</td>
<td>Late stage bot ecosystem without governance guidelines.</td>
<td>121</td>
</tr>
<tr>
<td>25</td>
<td>Wikipedia bot governance ecosystem.</td>
<td>122</td>
</tr>
<tr>
<td>27</td>
<td>Bot access and authorization on Wikipedia [source: Wikipedia.org].</td>
<td>131</td>
</tr>
<tr>
<td>28</td>
<td>Archived feedback of Addbot status on Wikimedia.org [source: Wikimedia.org].</td>
<td>131</td>
</tr>
<tr>
<td>29</td>
<td>Google Images search results for &quot;bot&quot; [source: Google].</td>
<td>134</td>
</tr>
</tbody>
</table>
CHAPTER ONE: RHETORICAL AGENCY IN SOCIAL MEDIA BOT ALGORITHMS

BACKGROUND

The average person in an online space uses algorithms to perform functions that seem, by now, almost second nature and automated to the point of refinement. We use them to search for things, for instance, typing in a query to locate basic services and goods, or to explore and navigate the latest news in Twitter. It hardly strikes us that these algorithms are inherently value-laden and present a number of biases and issues that can be amplified, distorted, remixed, or silenced in a sea of textual noise. “Few computer users know anything about programming,” said Carolyn Miller, “they simply manipulate an interface” (2001, p. 255). We interact with algorithms in many different disciplines, and we carry on a relationship with them through design, use, feedback, and remixing, among other actions.

A “bot” is a reductive term for “web robot” or “robot”—an algorithm written for the purposes of automating various processes on the Internet. Bot algorithms serve as distributors of communication, carrying within them all of the trappings of visual and textual rhetoric, semiotics, artistic expressions, and linguistics. One example is the Twitter bot @ThreadReaderApp, which stitches together long Twitter threads—stories pieced together through short phrases or sentences posted as tweets. To use the bot, all you have to do is mention it within a Twitter thread and use the command “unroll”: “@ThreadReaderApp, please unroll.” The result is an easier and more convenient reading of what would otherwise be short, chopped-
up sentences or phrases. It automates and simplifies our lives, and it appears to be relatively simple. But it, too, is value-laden and can alter the course of communication from rhetor to audience. It alters our perception of timing, or kairos, of a deliberate rhetorical strategy from the Twitter author in exchange for convenience. This shift in timing – or kairos – is hardly perceptible, and likewise the shift in agency can be altered almost imperceptibly.

Another Twitter bot, @IMG2ASCII, programmed by Mario Klingemann, took tweet-submitted images (see Figure 1) and converted them to American Standard Code for Information Interchange (ASCII)—a series of teletype symbols that can be rendered visually from a digitized image (see Figure 2). Although the image is viewed in a simple manner as it transitions from one digitized image of an old photo (such as the example submitted by @commonsbot) to a similar likeness composed entirely out of ASCII symbols, the process is an interesting moment in procedural rhetoric and rhetorical agency that muddies the waters on what constitutes art, how agency is defined, and where meaning-making takes place. The first image was tweeted by a Twitter bot, the @commonsbot, as it sifted through the Flickr Commons Collection, a repository for digitized photos that are free for public use and publication. The @commonsbot tweeted the image to the bot @IMG2ASCII, which replied to @commonsbot with the altered ASCII image. No humans were directly involved in this exchange, although the bots were created by humans and shared on a platform that automated processes crafted by humans. Both of these bots, as well as the @ThreadReaderApp, carry tremendous communicative energy through the use of rhetorical agency, as exemplified by procedural rhetoric. The @commonsbot took photos that were converted to pixels and shared those pixelated images with the public. And it was, in 2014, caught in a perpetual loop of tweet-conversations with the @IMG2ASCII bot, which was converting image pixels to ASCII symbols. The communicative energy of an old photo – which
appears to be a U.S. government photo of what looks like the remains of a battlefield – resurfaced into the Flickr Commons Collection, most likely when a library digitized the photo. Then the @commonsbot selected the image to share with the @IMG2ASCII bot, which altered the image into ASCII. Finally, Twitter users shared, “liked,” or copied the image for other uses, or perhaps for remixing. In short, these simple exchanges allowed an otherwise obscure image to enjoy a new life on the Internet, and then remixed as part of an artistic process, even making its way into the pages of this dissertation.

![Image](image.png)

*Figure 1: Image from @commonsbot, a Twitter bot that collects and tweets images from the Flickr Commons Collection, tweeted to @IMG2ASCII.*
In order to better explore how these examples work, we need to examine the concept of rhetorical agency. Simply put, rhetorical agency transfers the energy of an action or a call to action, there is movement of persuasiveness from rhetor to audience that results in a measurable reaction, however slight or perceived. Markus Schlosser (2015) charts the philosophical arc of agency as beginning with Aristotle and, later, receiving renewed traction through the works of Enlightenment philosopher David Hume as he pushed back against rationalism. In the *Nichomachean Ethics* (Aristotle, 350BCE) Aristotle contends that actions are not limited to the
Aristotelian concepts of agency involved a more deliberative attempt at generating social change with a call to action. Belikian (2017), however, notes that many contemporary theorists view the Aristotelian model as “instrumentalist, as attributing to rhetorical utterance the causal efficacy of some philosophical slingshot” (n.p.), that these scholars “have outstripped both the notion that rhetoric might express the resistless will of the rhetor, and that notion that rhetoric might stage-manage the thoughts and behaviors of subjects imagined as stable, fully formed, self-sufficient” (n.p.). Belikian, citing Christine Gardner (2011), says that “although rhetorical agency can’t nowadays be treated as the province of a sovereign subject, it can at least be studied as ideology, or as a power, or as responsibility, or – perhaps a bit disingenuously – as human potential, or as a resource, or as performance, or as illusion” (n.p.). However accurate as these terms may be, they are poetic and abstract and offer little in the way of concrete reality, quantifiability, or a meaningful way to study the transfer of agency.

More helpful are definitions put forth by Carolyn Miller and Laurie Gries. Agency has also been defined as the “kinetic energy of performance that is generated through a process of mutual attribution between rhetor and audience” (Miller, 2007, 137). Laurie Gries (2015), channeling the visual rhetorical agency of the “Obama Hope” poster and the many similar images generated after it, called agency “a dynamic, distributed dance that cannot be turned off and on like a water fountain” (70), a rejection of previous, linear concepts and definitions of agency as models of action and reaction. When considering agency and digital technologies, Jeremy David Johnson (2017) cautioned us about a “crisis of agency,” since algorithms further trouble notions of agency by casting doubt on what makes any behavior uniquely human” (p.
This question of human-ness poses a critical question for rhetoricians and scholars regarding ownership and responsibility over algorithms and the ethics of creating such inventions.

In 2004, Cheryl Geisler wrote in a landmark report from the Alliance of Rhetoric Societies (“How Ought We to Understand the Concept of Rhetorical Agency?”) that “recent concern with the question of rhetorical agency arises from the postmodern critique of the autonomous agent” (p. 10). She noted that, among the 40 scholars convening to address rhetorical agency, the following consensus was reached: “As rhetoricians, we generally take as a starting point that rhetoric involves action” (p. 12). Agency, therefore, is a central component to rhetorical inquiry, no matter the debate on the skill of the rhetor, the conditions under which the agent is acting, the situation under which an audience receives a communication, or whether or not there is an agreement upon the conditions and definitions of rhetorical action. As a concept, agency is elusive, but it is, indeed, tangible and actionable, and, almost ironically, it is the automated feature, as one example, of the algorithm that is now showing us where and how agency changes. Through repetitive actions and lack of nuance, automated processes bludgeon us into recognizing these changes. This is, of course, but one variant of how simple algorithms express themselves through rhetorical agency. Additionally, when a social media algorithm does something different from what it is intended, or when a different actant behaves in a manner that is unpredictable or altered as a result of the algorithm, the redistribution of agency becomes tangible and observable.

My study of algorithms adds to, and challenges, these concepts and definitions of agency by following both the predictability and unpredictability of social media algorithms as they exist in the digital ecologies of the omnipresent social media platforms of Facebook, Wikipedia, and Twitter. Through a careful plotting of major and minor actants within the sphere of influence of
the social media algorithm, I demonstrate not only the dynamic dance and fluidity of agency as it shapes and reshapes discourse and other communication media and methods, but also that agency can be deliberately shut off, redirected, corrupted, or amplified through both strategic planning and unintended consequences. The water faucet metaphor mentioned by Gries is, in the case of algorithms, only partly true if you consider the kinetic metaphor of the water pressure behind the faucet. It seems, in certain cases, platforms, humans, and other algorithms can shut the agency water valve off for a time or reduce it to a trickle. Social media bots help to demonstrate different characteristics of rhetorical agency, which includes the ability to stifle the kinetic energy of rhetorical action, persuasion, or meaning-making.

It is important to note that although other environmental factors carry a responsibility for algorithms and the kinetic energy behind them – the regulating of algorithms on a platform or the social responsibility of programmers, for example – there are deliberate strategies involved and currently evolving around algorithms as persuasive devices that alter our discourse. Bots aren’t necessarily hierarchal or privileged over other agents, but they are emerging as strategic tools through both deliberative strategies and accidental/experimental discovery. However, their usefulness as tools exceeds that of traditional notions of static artifacts of hardware. In *The Media Equation* (1996), Reeves and Nass come to two important conclusions: that “media equal real life” and that “people respond socially and naturally to media even though they believe it is not reasonable to do so, and even though they don’t think that these responses characterize themselves” (p. 7). These notions transcended the preceding thought of media as tools, since “people don’t have social relationships with tools” (p. 6). The nature of social media bots is such that they are reliant upon social interaction in order to be effective as a tool.
Social media bots display a number of agency-changing properties that alter our perceptions of human-ness and the human quality of rhetorical agency, rendering obsolete the Aristotellean concepts of rhetor and audience, action and reaction. This dissertation demonstrates how bots obfuscate and disrupt, amplify, replicate or mimic artistic expression, augment and remix, entertain, and serve as mediators in dissoi logoi relationships. balancing out two competing arguments with measured responses. A social media bot – as an algorithmic device – marries both the conventions of communication and rhetoric with logic-driven commands that are similar to those of machines, but machines that alter many aspects of communication. They are, in short, fascinating devices that alter everyday conversation, finances, politics, medicine, research, sociology, history, critical race theory, women’s studies, and a never-ending variety of disciplines, areas of research, and communications.

ALGORITHMS AND RHETORICAL ENTANGLEMENT

Because they are written and executed by humans for the purposes of communication and automation, algorithms and humans are inextricably linked together. It’s important to consider that algorithms are often created to be indistinguishable from actual human communication and therefore obfuscate that the origins of such communication originated with a person. However, even though programmers can imbue their algorithms with human-like qualities, those creations are ultimately tethered to the programmer inasmuch as Boeing creates airplanes or Smith & Wesson creates handguns. But algorithms confound these types of industrial conventions when they take on anthropomorphic qualities and real people interact with them under the premise that users treat the algorithms as human or human-like. Algorithms that are treated more as tools are regarded differently and shape rhetorical agency in different ways. The more an algorithm is
treated as a tool rather than a social media bot, the more it appears that humans remove their emotional connections from it. By calling a tool a tool, or a bot a bot, the more we are able to compartmentalize our social interactions with it.

The programmer, therefore, must take into account that anthropomorphizing algorithms carry an extra burden of responsibility. In February of 2015, Twitter user Jeff van der Goot’s twitterbot composed tweets based on a Markov chain algorithm using textual input to string together phrases and segments of sentences to create new utterances. Some of the tweets were nonsensical, but then the bot randomly put together a new phrase – a death threat – that it tweeted out. Since the tweet no longer exists, I cannot relay the exact death threat, but it was enough to get the attention of the Dutch police. In real time, van der Goot tweeted out his 2015 encounter with the police (Figure 3). He also wrote, “But apparently *I’m* responsible for what the bot says, since it’s under my name and based on my words.” He seemed surprised that he would be held accountable for the bot’s language. The police had him delete the tweets and twitterbot, and they seemed confused that an algorithm had been the source of the problem.

Furthermore, the twitterbot actually tweeted in response to another account, which also happened to be a bot. The police reacted to an algorithm’s randomized choice of words (the death threat) as it threatened another algorithm.
Programmers code algorithms within a language familiar to themselves with the express notion of communication—whether it be financial, artistic, geospatial, textual, etc. However, algorithm utterances or output can vastly differ from human intent. Predictive models of classic rhetoric—and early digital rhetoric—are now obsolete. Gries (2015) and Marilyn Cooper (2012) dismiss the notion of causation and predictive rhetoric, where we mistakenly believe “causation
is a linear process in which an agent’s action prescribes a certain result” (p. 437). Gries explains that biology is but one actant or component in an assemblage. “Agency,” therefore, becomes “an act of change that arises from an entanglement of human and nonhuman entities and other environmental factors, each of which is but a phenomenon of ongoing historicity” (p. 68). These entanglements can seem so complex because humans are unable to understand the infinite number of outputs and variables executed (a simple text bot drawing from a corpus of 100 words could theoretically generate more differing strings of words than there are stars in our galaxy), but that doesn’t mean strategic or unplanned changes in agency cannot be affected or changed.

Social media bots – like a text bot or chat bot¹, or a simple process-automating bot – are communication-heavy algorithms that people rely on to automate and simplify our lives. And people have quickly learned to use these bots toward particular ends over machines, ourselves, and each other, altering hegemonic structures. Brunton and Hessenbaum (2015) research and write about Twitter bots as obfuscators—automated algorithms filling channels of text with disruptive “noise” to the point where legitimate communication is indistinguishable from false or disruptive communication. Brunton and Hessenbaum also say that obfuscation “is a tool particularly suited to the ‘weak’—the situationally disadvantaged, those at the wrong end of asymmetrical power relationships,” which can be summarized by saying that Twitter bots and social media bots can serve as agents of empowerment to the disadvantaged (p. 9). Among several examples, you will see in the case study of the Peñabots that the Mexican government, at

¹ A text bot is a simple algorithm that uses different types of text-generative models based on a bank of textual input. I mentioned van der Goot’s twitterbot above—it uses a Markov chain algorithm to assemble phrases and pieces of sentences to make new utterances. A chat bot, however, is a little more complicated, and it bases textual responses according to queries by another human or algorithm. If you “chat” with a Verizon employee online, for example, many of the responses are fed to you by a chat bot algorithm according to your question.
the time headed by President Enrique Peña Nieto, used the tactic of the disadvantaged against the disadvantaged. By obfuscating activist trends by poisoning hashtags and drowning out activist messages with textual noise, the Mexican government employed the same strategies as the activists in order to have citizens and Twitter users question the legitimacy of activist messages, hashtag trends, and even geographic meeting spots.

Big Data algorithms, such as Google’s hierarchal search utility, favor hegemony and power structures that reflect the biases of the algorithms’ creators. Google searches for “Latina women” or “Asian girls” retrieve disproportionately hypersexualized content, for example (Noble, 2018). The Google search engine also reflects capitalist economies, allowing paid advertisements to appear first, followed by a ranking that favors digital capitalism hegemony—websites that drive traffic, “clicks,” and views. Social media bots, which I have chosen to research, generally do not possess complicated algorithms or groups of algorithms and lend themselves to easier study than proprietary, complex algorithms belonging to companies. They also tend to be more experimental, which allows for more dynamic study when considering properties and concepts of rhetorical agency. The process of experimentation by programmers demonstrates a more transparent push-and-pull of user and object interaction with the algorithm, and the programmer makes adjustments or observations about her or his program. Because many social media bots are simple programs, programmers replicate them easily, and semi-computer-literate laypersons comprehend how to code these algorithms with little training. Platforms like Github – a repository of open-source codes, feedback, and interaction – greatly simplify the coding process.
ALGORITHMS AS OBJECTS

Each social media algorithm case study bot I researched is different in many ways, but at root the algorithm must be treated as an object—the sum of its code and inner parts. In an object-oriented litany, a textbook full of instructions could be deconstructed to the type of paper within its pages, the typesetting, font, binding, cover artwork, and dust jacket. Each of these things is value-laden. The font, for example, conveys meaning. The general whimsy of the comic sans font is well documented. Serif fonts portray meanings from a different time period, such as the textbooks of the 1920s or digital text of the 1990s. Sans serif fonts allude to modernity and ease of screen reading. The assemblage of these things creates a book. But if an algorithm, laden with instructions like a textbook is dissembled, the output becomes meaningless, so dependent is each line of code upon another. Social media bots, therefore, need to be treated as individual objects. There are instances within the code that can change a font type, color, size, or placement with a simple command or tag, but often these commands or tags become so enmeshed within the entirety of the code that through the creation of an ontograph – the visual geography and layout of interconnecting actants within an ecosystem – even simple algorithms become counterproductive from a human’s perspective. This is important to note for the following case studies and for the purposes of predefining algorithms with instructions and intent. Therefore, from an ontological framework, algorithms as value-laden objects can be visualized as exerting agency and functioning as actants in a given system.

Jessica Reyman (2018) defines algorithms as “software programs or computer codes that process data, perform automated reasoning, and then output information in a transformed manner toward a desired end” (p. 113). They operate by a set of logic-driven rules. We rely on the algorithms to perform complex operations on a level that far exceeds human capability, even
though we are crafting the rules. Because we are the creators of this technology, we imbue algorithms with Miller’s “kinetic energy,” and we use them in what Reyman calls participatory “meaning-making” that affects “human communication, understanding, and behavior on both small and large scales” (p. 115). In addition to this meaning-making, algorithms carry ontological weight worthy of being parsed out in research as we search for the subtle and not-so-subtle ways in which algorithms alter our interactions. John Muckelbauer (2016) uses Bruno Latour’s “famous speed bump example” to illustrate this point:

> Of course, one might justifiably say that the only reason a speed bump is effective is because drivers will extrapolate an argument from it: “If I do not slow down, I will damage my car.” Nevertheless, a speed bump is not in itself an argument—it is merely a mass of concrete. But that does not make it any less persuasive—if anything it makes it *more* persuasive. The point here is that everyday physical structures may not exactly be arguments (and they are certainly not primarily linguistic—though they are surely integrated with language), but they are undoubtedly persuasive. As a result, it is important to consider these types of objects as crucial components of rhetoric. (p. 36)

From the framework of rhetoric as discourse, algorithms act as both a piece of technology and an implementation of discourse (Mittelstadt et al., 2016). One example that highlights this relationship is @poem_exe, a Twitter bot that writes poetry. It is a piece of technology, coded by Twitter user @inky, who used Tweepy and Python languages to marry the algorithm with the Twitter Application Program Interface (API). From a classical standpoint, @poem_exe is a piece of technology. Plato’s *technê* can be loosely defined as “a bringing forth of the skills and trade of a craftsman.” The algorithm also brings forth something *poietic*, reconciling matter and time and person with the world. Therefore, epistemologically, technology – and the algorithm – is an object of the mind and arts and is created by skillful tradecraft.²

---

² Personal notes taken from lecture by Dr. John D. Schaeffer, Northern Illinois University, 2016.
The poem_exe bot utilizes Markov algorithms within the “parent” algorithm that string together lines of text in a semi-Thue system, which is a system that utilizes grammar rules to recreate new lines of text that sometimes contain meaning. This is to say, the Markov algorithm built into the Twitter algorithm retrieves grammar patterns on Twitter, assigns each of them a value, and rearranges them according to those values in phrases and sentences that more often than not can make sense to a human audience. In this case, @inky’s Twitter bot writes “poems” based on grammar patterns (see Figure 4).

![Tweet from Twitter bot @poem_exe.](image)

The poem seems like it is telling us there is a little snail in my soup bowl that I’m about to eat, and they look like pearls of bright dew. But the three lines of text were taken from completely different phrases and strung together by the algorithm. The pattern isn’t completely random – the Markov algorithm assigns number values to grammar patterns – but the poem/output is random enough that the chances this poem will appear again, even when pulling from a small sample of text, are remote. The algorithm inventor/programmer will never be able
to predict the textual outcome of this poetry bot, and neither will the audience. We must also consider a poem that doesn’t make any sense to put this in context (see Figure 5).

Does the value of the poetry bot change when we are unable to assign any sense to the poem? Wherein does “meaning-making” lie? When the algorithm makes sense, it acts as a piece of technology, and it serves as an implementation of discourse, but arguably only when it makes sense to the audience. And how do we assign rhetorical responsibility when a programmer creates an algorithm that is designed to be completely unpredictable compared to an algorithm that has a carefully planned outcome that largely functions within its parameters? These questions are important when considering the distribution of agency as programmers create bots and send them forth for different purposes.

Ontologically, @poem_exe must be considered differently than from a linear, rhetor-to-audience communication model that also views the algorithm epistemologically. An ontological...
framework suggests that we consider how the algorithm interacts with other objects (Barnett and Boyle, 2016), how we perceive the algorithm’s output through different moods and environments (Cooper, 2016; Mazis, 2008), and the historical and cultural weight of the algorithm itself, for example, considering such things as literacy rates or literary access when examining the value of textual outputs (Muckelbauer, 2016).

In addition to an ontological approach, we must consider the procedures of intent—what does a programmer mean to do when she originally writes the code? By determining intent, the rules of an algorithm’s code transcend error and glitches and give responsibility to the programmer. There are very different potential outcomes when a programmer experiments with a bot algorithm compared to creating one that is carefully planned for a specific purpose within a given system. These different outcomes also have implications when studying rhetorical agency, ownership, and ethics. Figure 6 shows the distribution of human interaction with social media bots after a programmer executes the program with either an experimental outcome or planned outcome in mind. The purpose of this flow chart is to demonstrate the unplanned outcomes of social media bots, which disrupt intent and fosters experimentalism, where rhetorical agency more closely resembles art rather than discourse. Carefully planned social media bots could very well be called tools of automation and discourse, rather than viewed as (and called) bots. I developed Figure 6 as an early result of initial data collection surrounding several case studies, three of which are carefully described in the following chapters. Through an assessment of these case studies, it became apparent that there are two macro-level types of social media bots, those created with the purpose of experimenting and testing boundaries and limits and those carefully planned and created with a specific function or functions in mind. In a sense, one type of bot
lives in a sphere of anarchy where everyone participates in its lifecycle without responsibility (Lanham, 2006), and the other within a system of rules and laws.

**Figure 6:** An algorithm's outcomes through experimentation vs. planning.

**ALGORITHMS, EXPRESSION, AND ART**

Both intended and unintended results are subject to experimentation—through remix, refinement, memes, commentary, and many other outputs that redefine the intent of the original program. Once an algorithm is executed, each actant the program comes into contact with becomes part of the meaning-making process: adapting, remixing, editing, reflecting, amplifying, rejecting, and/or corrupting the bot for their own agendas and purposes. This is the dynamic dance of agents and agency that is simultaneously both difficult to control, and yet those who are
able to control this fluidity by redirecting it stand the most to gain from shifting hegemony from one area to another. Control over these power structures means to hold significant power over discourse and action. We also see this type of struggle within the world of art, where artists use different media to critique and effect change and action.

Once I began to catalogue the lifecycle of the Facebook [loveMachine], I realized that its programmer – Julien Deswaef – self-identified as a digital artist who desired to make cultural and philosophical statements with his artwork. In this case, he created the [loveMachine] to achieve these ends. Understanding algorithms as artistic devices, I decided, is important. However, artistic philosophy within this space proved to be a daunting task. This is exemplified by setbacks Deswaef experienced when he attempted to find someone to curate his artistic experiment. he was told that he wasn’t truly an artist and his algorithm existed in a medium not easily displayed or curated within a traditional or conventional space.

However, it’s important to convey the points about expression and invention present in works of art as they pertain to algorithms, particularly to algorithms that are purposefully designed to be either expressive or inventive artistic devices. I elaborate further within the next chapter about expression and meaning-making and how users either willfully or unwittingly treat expressive bots as human. Creating a bot with the purpose of disguising it as human, or to confound the conventions of human and algorithm interaction – as in the case study of the [loveMachine] – brings with it different questions and philosophical digressions, as you shall see.
“Remix’ generally refers to the practice of recombining preexisting media content – popular songs, films, television programs, texts, web data – to fabricate a new work” (Gunkel, 2016, p. 1). Social media algorithms figure prominently within the sphere of remix. Online users, for example, quickly absorb viral news content and then respond with photoshopped memes, bastardized quotations, jokes, commentary, videos, and links to music, which are distributed and redistributed, often providing several layers of feedback and commentary to a particular news item in a single day. Some bots are specifically crafted to create, facilitate, archive, or manage these types of meme cycles online. Others, like the aforementioned @Image2ASCII bot on Twitter (Figures 1 and 2), experiment with content, remixing to create something new and different. Working with Natural Language Generation, a brother and sister team – Chris Rodley and Ali Rodley – created the @MagicRealismBot on Twitter. It crafts “a magical story every four hours” and tweets it to its followers. Based upon the concept of José Luis Borges’s magical realist writings, the bot reassembles textual elements into a sentence or two that often seems like the opening or thesis to a magical realist novel (Figure 7). “A schoolteacher in Mongolia is famous for exacting revenge upon people named Richard.” “A 12-year-old watchmaker falls into a swimming pool filled with childhood dreams. Nobody misses him.” “A schoolmistress is writing a list of people she plans to kill: A supermodel, a philosopher and a violinist.” Within these processes of digital remix are experimentations of appropriation, social commentary, redirection, activism, and an infinite list of reactionary responses. Even the simplest of recombined textual elements can cause us to question ownership, authorship, credibility, and intent.
On Twitter, the limitation of characters is currently 280, which forces Twitter users to adapt creative textual and visual strategies in order to offer memorable or interesting commentary. A poorly written or controversial headline tweeted out by a news agency might be retweeted with a rewritten/remixed headline meant to draw attention to another issue or the
“real” issue with the tagline, “There, I fixed it for you.” Bots can be programmed to manipulate these types of strategies, remixing and recombining textual and visual elements for a variety of desired effects. This is, to say, that rhetorical agency is greatly affected by remix, and further complicated by the automated dissemination of remixed works.

RHETORICAL SIMULACRUM

An algorithm should be regarded as an actant of agency during the planning, coding, and execution processes and then again as other network actants and agents respond to it once a command has been executed. Algorithms contain silent lines of code – or a “null” input – that remain useless until executed with an action or command. The term *actant* has differing connotations, depending upon the researcher and discipline. I prefer to lean toward the sociological definition in the actor-network theory (ANT), by Bruno Latour and Michel Callon, wherein an actant is a rhetorical component born from actantial, semiotic analysis, it is a component that mediates or translates (Latour, 2005). As in literary and rhetorical theory, actants are objects that exert agency in a network. Algorithms are assembled lines of code, but the assemblage forms a package that carries within it an embodiment of great agential potency, depending upon other actants and actors within a given system. Until algorithms are executed, however, they exist as objects-in-reserve.

Heidegger (1977) said that objects-in-reserve are revealed by “unlocking, transforming, storing, distributing, and switching” (p. 7). Because Heidegger passed away in 1976, before the microcomputer was popularized, I argue that we can include “executing” as part of this canon of
Platonic technē. Heidegger also referred to technology as responsible for something else—in this case other actants within a rhetorical system (1977), and Justin Lewis (2016), through this framework, referred to technology as exerting agency by the simple fact that we cannot experience the world without it. However, when we limit our worldview of agency as dependent upon human interaction for execution and as humans dependent upon technology to exert agency and change, then we overlook some of the inherent values imbued within technology—in this case algorithms as technology. To wit: algorithms are ontological objects or “things” worthy of exploring as actants of agency and change. Furthermore, algorithms complicate the notions of human-ness and objects. Humans, Latour (2005) noted, are capable of weaving through complicated and shifting positions of competing object-actants, we can reassess shifting situations and give weight to certain objects and importance. When simple algorithms can mimic the flexibility of humans, we need to take note.

Algorithms exert agency, for example, by assigning a hierarchy of results or outcomes dependent upon code. Just as traditional rhetors assign weight and importance to objects, programmers can code algorithms to mimic similar processes. As humans, we create lists; we organize, assign, prioritize, and reassess how the objects in our lists affect us and other people and objects around us. Drivers see the rhetorical weight of Latour’s metaphorical speed bump ahead of their cars and adjust their speed. At that moment, they recognize the agency of the speed bump and note that if they don’t slow down they could injure themselves, experience discomfort, and potentially damage their vehicle. This is a momentary human assignation of hierarchy. Algorithms can be programmed to react similarly. Programmers can assign simple social media algorithms of basic if, then commands in a ranked order to build a list of priorities according to hierarchy. Simple algorithms cannot adjust to rapidly changing conditions, but they
can be programmed to act like they are. This is, in effect, a *rhetorical simulacrum* or *simulacra*—a simulation, exerting agency through procedural action.\(^3\)

In his text, *Of Remixology* (2016), David Gunkel cites Sterne in noting that both a copy and original are “products of the process of reproducibility. The original requires as much artifice as the copy” (p. 74). Gunkel writes that “original” and “copy” are “already … by-product[s] of production” (p. 74). A simulacra offers a copy or a replacement of an original production, yet the reproduction, in the form of a simulacra, offers subtle differences, both tangible and intangible. Gunkel calls upon other scholars\(^4\) to demonstrate the anxiety and discomfort surrounding the moment when a simulation replaces that which created it. I recall a video version of *This American Life* that featured a Texas couple’s love for their gentle Brahman bull that acted more like a pet puppy than a dangerous bull. Ralph and Sandra Fisher named the bull Chance, and it appeared at parties and in movies, and children loved to play with it. After Chance died they were devastated, but scientists and researchers at Texas A&M University offered the Fishers a second “Chance” through the science of experimental cloning. The copied Chance looked like and sounded like the original Chance, but the Fishers quickly realized that it didn’t behave the same and was prone to violent outbursts like normal, temperamental bulls. The simulacrum Chance lulled Ralph into a false sense of security when performing what used to be routine tasks like feeding the animal in its pen. Ralph felt a conflict when looking at this detached, dangerous version of his beloved bull that looked exactly the same. Similarly, when a

\(^3\) A *simulacrum*, according to Jean Baudrillard, is “no longer that of a territory, a referential being, or a substance. It is the generation by models of a real without origin or reality: a hyperreal … It is no longer a question of imitation, no duplication, nor even parody. It is a question of substituting the signs of the real for the real” (1995, pp. 1-2).

\(^4\) Baudrillard (1983, 2, 5); Mark Taylor & Esa Saarinen (1994).
human programmer builds a social media bot to mimic a human action and that bot performs that function – but is unable to completely accurately respond like a human – there is an element of remix, of agency disruption, and a questioning of ownership. All of this is accomplished through simulating human actions through automation, which is a function of procedural action.

By using video games and research programs as examples, Ian Bogost’s work on procedural rhetoric explains the computational procedures and processes that exert persuasive forces upon us. Using MIT’s ELIZA program as an example of a natural language processing (NLP) algorithm, Bogost (2007) wrote that the mid-1960s program was an example of procedural authorship; a computer responded to written queries with realistic language that mimicked a psychologist responding to a patient. This was done partly by assigning hierarchal values to input queries developed with user feedback.

*Eliza* seeks out keywords such as “I am” or “you” and manufactures transformations based on them. The computer program has no real understanding of the meaning of the user’s input; rather, it is taking that input and spinning it into a possible conversation. *Eliza* is a machine for generating conversations according to procedures. (p. 11)

Bogost greatly simplified how ELIZA functioned; the program possessed more depth and complexity, but it still assigned values to keywords as it responded to typed input:

It’s true. I am unhappy.
DO YOU THINK COMING HERE WILL HELP YOU NOT TO BE UNHAPPY?
I need some help, that much seems certain.
WHAT WOULD IT MEAN TO YOU IF YOU GOT SOME HELP?
Perhaps I could learn to get along with my mother.
TELL ME MORE ABOUT YOUR FAMILY.
My mother takes care of me.

This is part of the example used by Bogost in his text *Persuasive Games: The Expressive Power of Video Games* (2007). The text above is a fragment of an original exchange between
ELIZA (in caps) and a “patient.” “The creation of computer scientist Joseph Weizenbaum, ELIZA was named to reference the Galatea and Pygmalion myths, an acknowledgement that he was ‘teaching’ the inanimate machine to speak. His computer program, a natural language processor (NLP), was comprised of two components: a language analyzer to interpret what was said to the program, and a script, which applied transformational rules to keywords in order to produce a response” (Weil, 2017, p. 22). Bogost noted that although the ELIZA program was clearly not a substitute for an actual Rogerian psychologist, it was an “expressive agent” used for the purpose of creating a believable agent, attempting to achieve something far grander than a simple keyword-searching program. Bogost’s point was that “procedurality is not limited to tool-making” (p. 11). Expressive social media algorithms tend to be anthropomorphized, but non-expressive social media algorithms tend to be viewed as utilitarian tools, processes, or procedures. The differences between expressive and non-expressive algorithms have great ramifications upon agency. A rhetorical simulacrum is problematic when we consciously or unconsciously assign it the same values as we do with humans, when in fact it cannot process those assignations in a conventional, rhetorical manner.

“Algorithms can be seen as rhetorical insofar as they exert a persuasive influence upon what is held to be important or true in our social, cultural, political, and economic interactions” (Ingraham, 2014, p. 63). Social media bots fulfill a number of rhetorical functions that bridge these interactions, reaffirming Ingraham’s assertion that they persuade us of what is important in our social, cultural, political, and economic interactions. And, it must be said partly in jest and partly backed by serious study, what is most important to us is the self—our place within the world and how we figure into it. And it is this attentiveness to the self that obscures other rational thinking. Reeves and Ness (1996) drew attention to a televised presidential debate set in
a format where candidates received questions from audience members. At one point a candidate left the podium and walked toward the question asker, but on television, the candidate appeared to be walking toward the viewer. The result, Reeves and Ness concluded, was that this candidate’s comments were viewed favorably not necessarily because of his answers, but because he seemed to approach the physical space of the viewers. This is, overall, a summary of social niceties that have to do with the self. Reeves and Ness use their own book, *The Media Equation* as an example within a space of social behaviors and politeness:

> Consider this research example, a prototype of our work on the media equation. If either of us called you on the phone to ask how well you liked this book, you would likely be polite and say that it was fine. But if someone else were to make the same phone call, the evaluation might be less positive. This is an example of a simple social rule: People are polite to those who ask questions about themselves. (p. 5)

My ten-year-old son, when interacting with an Amazon Echo, through the AI personality “Alexa,” commanded it to “shut up” when it was rambling on about a Wikipedia definition, and my eight-year-old daughter and five-year-old son reprimanded him for his rudeness. First, we aren’t allowed to say “shut up” in our house, especially out of anger. The ten-year-old’s choice of words was particularly shocking. Second, my younger two children felt genuine empathy for Alexa. “Don’t say that to her!” cried out my daughter. “That’s not nice!” said my younger son. They reprimanded the elder son as forcefully as they reprimand me when I scold the family dog. “It’s not real,” was his reply.

This leads me to another important observation from Reeves and Ness (1996) regarding social cues and politeness to computers, both of which inform current rhetorical agency in the digital world. They studied politeness toward computers through the use of textual expression and voice. “If machines don’t deserve our positive regard, there should be no differences in
evaluations of the computers based on which one asked the questions. If computers are social actors, however, then participants who responded to the same computer that taught them should be polite, and uniformly so, just as if the machine were a real person with real feelings” (p. 23). The results of the studies concluded that participants who provided answers on the same computer that issued them gave more favorable responses than when they provided feedback on a different computer. Even more interestingly, the results didn’t change much when the computer used a human voice to solicit evaluations instead of asking via text on a screen. “The conclusion: Users are polite to computers whether they use text or voices” (p. 25). Our automatic inclinations of social propriety and niceness toward computers figures into how agency can be shaped online, that is if we make the theoretical leap from computer to social media bot, especially one that engages with us.

I’d like to add that rhetorical simulacra – these social media bots – also affect our interactions with art—its creation, interpretation, execution, and remixing properties. I posit that algorithms could be viewed as artistic actants themselves, challenging us to question the aesthetic of our epistemic interactions. Ingraham (2014) said that because we are thinking of algorithms as rhetorical, it means we are “recognizing that they are neither infallible nor disinterested, but rather motivated by quite specific epistemic standards that can radically delimit what counts as valid or meaningful in various spheres of our intellectual, social, and material world” (p. 64). As I shall demonstrate, social media algorithms do fit these definitions put forth by Ingraham, and the statement, “exert a persuasive influence upon what is held to be important or true” (emphasis mine). is a critical observation in our current world of disinformation, obfuscation, and increasing need for credibility. In short, these types of algorithms carry with them values and potential rhetorical agency and energy.
Incorporating the research of several mathematicians and computer science scholars, Mittelstadt et al. (2016) agree that algorithms are “inescapably value-laden,”\textsuperscript{5} and their “operational parameters are specified by developers and configured by users with desired outcomes in mind that privilege some values and interests over others”\textsuperscript{6} (p. 1). Mittelstadt et al. cite a 2013 study by Sweeney that shows how algorithms “inadvertently discriminate against marginalized populations … as seen in delivery of online advertisements according to perceived ethnicity”\textsuperscript{7} (pp. 1-2). As I will demonstrate through case study analysis later in this dissertation, it appears that while algorithms are very efficient at quantifiable calculations, programmers are 1) offering the wrong input – by accident or on purpose – and expecting algorithms to quantify queries that are best served with a qualitative response (Facebook [loveMachine]; 2) programmers are asking algorithms to quantify and solve problems without understanding – or because they understand very well – that they are inserting their own biases in programming, user-experience design, and feedback stages (Wikipedia Addbot, Twitter Peñabots); and 3) these algorithms carry another set of ontological values that programmers generally do not consider (Facebook [loveMachine], Twitter Peñabots, Wikipedia Addbot).

THE AUTOMATION OF SOCIALITY

Robert Gehl and Maria Bakardjieva’s anthology of social media bots, \textit{Socialbots and their Friends: Digital Media and the Automation of Sociality} (2017) brings together computer

\textsuperscript{5} Brey and Soraker, 2009; Wiener, 1988.

\textsuperscript{6} Friedman and Nissenbaum, 1996; Johnson, 2006; Kraemer et al., 2011; Nakamura, 2013.

\textsuperscript{7} Barocas and Selbst, 2015; Birrer, 2005; Sweeney, 2013).
scientists, digital artists, researchers in communications and rhetoric, techno-philosophers and others to discuss the impacts and parameters (and future implications thereof) of “socialbots.” Gehl and Bakardjieva begin their collection with a definition of a socialbot from Boshnaf et al., 2011:

A socialbot is an automation software that controls an account on a particular OSN [Online Social Network] and has the ability to perform basic activities such as posting a message and sending a connection request. What makes a socialbot different from self-declared bots (e.g., Twitter bots that post up-to-date weather forecasts) and spambots is that it is designed to be stealthy, that is, it is able to pass itself off as a human being (p. 93).

Gehl and Bakardjieva draw particular attention to the fact that a socialbot is “designed to be stealthy” and that it can engage others by passing itself off as a human.

RESEARCH QUESTIONS AND METHODS

RQ1: How do social media bots affect rhetorical agency?

RQ2: Are there areas within the lifecycle of a social media bot where ethical applications can or should be implemented?

RQ3: By treating algorithms as rhetorical objects, what ontological values can be discerned through rich descriptions (and subsequent ontography) of these case studies?

The assertion that algorithms are exerting rhetorical forces among humans, objects, and other algorithms can be further explored and strengthened through research. Case studies of social media bot algorithms offer a look at several stages of a lifecycle of an algorithm. Through
selected case studies, I applied “thick description” … “an in-depth description of the entity being evaluated, the circumstances under which it [was] used, the characteristics of the people involved in it, and the nature of the community in which it [was] located” (Becker et al., 2012).

I selected three different case studies to showcase two different elements of social media bot algorithms: 1. the platform or ecosystem in which the bot resides and 2. the purpose for which it was created. The case studies look at three different social media platforms: Facebook, Twitter, and Wikipedia. The purposes for which the bots were created include art experimentalism, disinformation/obfuscation, and automating menial tasks. It was my intent to diversify the bot case studies as such in order to demonstrate not only their various functions but to be able to compare and contrast how they affect rhetorical agency.

During the process of collecting considerable amounts of data while researching case studies of social media bots, I began to create lists of things – actants, programs, platforms, systems, and other interacting parts – that contributed to the flux and change of rhetorical agency. Lists form the basic component of object-oriented philosophy. Bogost (2012) noted that “the naming [of] objects is only one ontographical method, the most basic one” (p. 39). The categorization and naming of things – the creation of lists – help to create divisions between actants—a basic premise of object-oriented ontology (OOO). The idea of lists, or litanies, “saves us from privileging humans, imposing a narrative structure fraught with anthropocentric and patriarchal biases” (Zabrowski, 2016, p.58). This partition of objects allowed me to create ontographs of rhetorical objects within a given system as I tracked the movement of agency—action, influence, persuasion, revision, execution, etc. This is different from classification in that ontologies allows for fluid hierarchies and offer no object or actant an advantage over one
another, given that rhetorical situations change with time and perspective. An ontology allow us to see shifting variables of agency.

“From the perspective of metaphysics, ontography involves the revelation of object relationships without necessarily offering clarification or description of any kind … a record of things juxtaposed to demonstrate their overlap and imply interaction through collocation” (Bogost, 2012, p. 38). I use rhetoric and ontological practices together to demonstrate how each actant exists as a separate entity, filled with Gries’s “kinetic energy,” yet without a corresponding actant. Each actant-object within an ontological framework is viewed individually, free from the constraints of other actants, which allows them to simply be “things.” Scot Barnett and Casey Boyle (2016) warn us of not recognizing the importance and value of objects, especially the seemingly mundane: “If we continue to think of things exclusively in terms of language, appearance, and representation – as epistemological objects – we will likely go on believing that human beings alone determine the scope and possibilities of rhetoric and that humans, as a consequence, are the only true legislators of nature. This is to the peril of all things” (p. 4). It seems that programmers are exerting willpower and influence over their lines of code, but these case studies demonstrate that their influences depend upon certain conditions, and there are limitations imposed upon these conditions by other agents. Are programmers the great arbiters of their code? Or should they treat their algorithms as individual agents among other agents within a given framework?

David Gunkel (2017) sums these questions up into a tighter space as he looks at what is to come in the increasingly complicated and interconnected digital spaces: “At what point, for instance, might a socialbot, an algorithm or other autonomous system be held responsible for the decisions it makes or the actions it deploys? When, in other words, would it make sense to say,
‘It’s the computer’s fault?’” With a mind toward the future, Gunkel posits these questions as an opening toward a bigger question of ethics: “… at what point might we have to seriously consider extending something like rights – civil, moral or legal standing – to these devices” (pp. 230-231). Agency, and the actionable consequences of the transfer of agency, figures heavily into this argument as we examine how bots alter our rhetoric and communication.

With these thoughts in mind, I conducted three case studies with the purpose of creating litanies of actants within a given system, the social media algorithm, the programmer, and other forces that connect and communicate with the algorithm as they push and pull, reframe, edit, and react during the lifecycle of the algorithm. This allowed me to more clearly see the distribution of rhetorical agency. By examining the value and rhetorical agency of even simple algorithms I suggest that as rhetoric and composition teacher-scholars, we will understand and realize how little our determinist nature holds sway over digital communication. The three case studies I researched are: 1. the [loveMachine] on Facebook, 2. the collective Twitter botnet known as the Peñabots, and 3. the Wikipedia Addbot.

Measuring the distribution of agency within a given social media algorithm’s platform (Facebook, Twitter, or Wikipedia in these cases) isn’t easily quantifiable. However, a ric-text analysis of a large corpus of data collection allows for simplified ontographs and litanies and their relationships with one another. Creating a simple, visual ontograph out of the lists of actants and agents allowed me to quickly reference relationships and agency dynamics. These visual interactions help rhetoricians to better understand the flow of agency and how it changes with the introduction of different actants. I developed these lists when I discovered them coming into contact with one another within a network or digital ecology. For example, Julien Deswaef, the creator of the Facebook [loveMachine], received feedback from others, either directly or
indirectly. His code also interfaced with Facebook using an assemblage of programming languages and interfacing applications. The algorithm affected different groups of people in different locations and in different ways, which I explain in the following chapter. By treating these aforementioned things as objects in the spirit of OOO, I must first create a list of them. Secondarily, these lists isolate the objects in and unto themselves. A visual ontograph, then, demonstrates how and where these objects interact with one another and with each other. Ian Bogost, writing in *Alien Phenomenology* (2012), offers several examples of visual ontographies, such as the “exploded view” of a parts diagram for the components of a Shimano three-speed internal gear hub for a bicycle as they exist in their individual pieces and where they are in the arrangement of connectivity for the gear hub. “An ontograph is a landfill, not a Japanese garden. It shows how much rather than how little exists simultaneously, suspended in the dense meanwhile of being” (p. 59).

With both the increasing complexity of algorithms and the complexity of assemblages in which they reach and connect, the question becomes not how *should* they be examined from a rhetorical standpoint, but how *can* they? The use of case studies of bot algorithms that have already been through the process of concept and invention, execution, output, and redesign or remix can answer that question. In addition, we can also analyze the creator-rhetor/inventor, her intent, audience response, and outcomes. The purpose of a critical analysis will both elucidate the strengths and weaknesses of examining rhetorical agency with humans and non-humans and allow us to rethink how agency is understood from the standpoint of rhetoric and ethics.

I conducted interviews where I was able in order to better understand the intent of the algorithm as well as gain a much deeper knowledge into the process of conceiving, programming, deploying, and updating the bot algorithms as they went through a lifecycle. In the
case of the Facebook [loVMachine], I coordinated email interviews, followed by extensive video interviews with Julien Deswaef, a Belgian residing in New York. Although Deswaef kept detailed blogging records of his experimental bot, and several other artists and programmers offered their own interpretations of Deswaef’s experiment, video interviews provided key contextual moments and demonstrated Deswaef’s thinking in response to different elements of change within the Facebook platform. Facebook conducted what was essentially digital warfare against the bot, and Deswaef had to adapt to changes in Facebook’s bot and automation detection algorithms.

Regarding the Peñabots, which were logged as tens of thousands of coordinated or semi-coordinated Twitter bots, there was no person to interview since the bots were criminal in nature, ethically unsound, and purposefully anonymous. However, information that catalogued the bots, bot behaviors, interactions with others, and probable origins was abundant and provided by a number of different sources that were academic, journalistic, and even academically sound by the standards of social science. The tweets were curated by researchers and publicly available for reading.

For the Wikipedia Addbot, I conducted email and video interviews with Aaron Halfaker – a Wikimedia researcher – and email interviews with bot programmer and employee of Wikimedia Deutschland Adam Shorland. I was able to corroborate some of their details and compare those to published articles on Addbot and Wikimedia. I also counterbalanced these narratives with published research that viewed the Addbot as a “violent” bot that was controversial in nature. With these primary sources of data collection, I performed a detailed and systematic analysis of all the records I could of each case study bot. By synthesizing these
secondary sources with my primary sources, I offered a rich textual analysis of these bots and the systems in which they resided.

This type of description gave me a holistic approach so that I could make my inferences ontologically and across several disciplines. The observations I recorded weren’t necessarily limited to my perspective, since I collected emails, comments, interviews, and papers from other people, and they are now available to the community for individual interpretations and conclusions as print, web, and archival data. The case studies that I selected consisted only of social media bots, but these bots were diverse in planning, intent, and outcome. Through multimodal data collection, I used at least four of the six types of data collection outlines by Becker et al. (2012) in each case study. These included documents, archival records, interviews, direct observation, participant observation, and artifacts (2012). Through these different types of data collection, I made every attempt to corroborate different information about each case study, and from different points of view.

In each case study, I recorded the data collection goals of Becker et al. under the areas of:

- **Early Observations**: I opened each case study with an early observation that led me to each type of case study, incorporating anecdotal information, scholarly research, or personal reflection. I asked the following questions: Why and how did I arrive at the decision to pursue each case study? What does each case study have to contribute to rhetorical agency and ethical behavior?

- **Program Description**: I wrote out a detailed description of the social media bot, its title, its purpose or inferred purpose, and a summation of its history. I also included programming languages, revisions, and other details relevant to its history.
• **Actant Characteristics**: This included the programmer and the people or objects involved that exerted influence on the program—either by design or interaction (inadvertent or purposeful). I created a list of actants for each case study so I could better understand the dynamics of the social media bot and how it interacted with other actants.

• **Environment**: In lieu of “the nature of the community in which it is located,” I shortened this subheading to “environment.” This means the various digital ecologies in which the program was coded, executed, and interacted with. Ontologically, I collected as much data as possible in order to recognize various groups and subgroups, actants, and objects that were affected by or affected the program.

• **Actants and Agency**: After data collection and description, I analyzed the effects of rhetorical agency that were exerted on various actants. Once the actants were identified, I compared the push and pull of agency among each actant or series of actants. This was done so I could demonstrate the network of influence between actants, particularly as we search for “meaning-making.” I generated a rudimentary map of influence, power dynamics, and transfer of agency dynamics.

• **Ethical Considerations**: Finally, I performed an analysis of ethical considerations as different actants worked to augment, disrupt, obfuscate, remix, amplify, suppress, or exert a number of altering influences on a social media bot. I identified where these changes took place, attempted to discern the intent of these changes, and the ethical considerations thereof. It is my hope that other rhetoricians and researchers will be able to use this extrapolation to more fully research and teach about rhetorical agency in an algorithmic age.
CHAPTER TWO: THE FACEBOOK [LOVEMACHINE]: EXPRESSIVE AGENCY, INVENTION, AND EXPERIMENTALISM

Facebook is omnipresent. It affects nearly all aspects of the Internet as it threads social lives into the fabric of commerce, entertainment, the media, healthcare, politics, and nearly everything else, turning social interactions, agents, and actants into commodifiable objects. Each “like” is a value that is entered into an algorithm or series of algorithms as part of procedural rhetoric. Facebook uses those “likes” as quantifiable data, which are, in turn, converted to digital currency as Facebook pairs commercial businesses with users. Let’s use a fictional Facebook account, which I call “Annie,” as an example: She “liked” the pages of the *Washington Post* and *Mother Jones*, comedian Sarah Silverman, Southern Illinois University-Edwardsville, a local elementary school PTA, and the St. Louis Cardinals Major League Baseball team. She publicly announced her anniversary date to her husband of nine years, her two children’s birthdays, and tagged several Facebook users as family members. Therefore, a Facebook algorithm has, in theory, assembled a marketing profile of her that states to third parties that she is most likely a white, heterosexual, progressive woman and mother, age 27-40, college-educated, from Missouri or Illinois, who uses Facebook somewhat frequently to share news, pictures of family, and other information related to her job and hobbies. Over the period of several years, Annie continued to provide much of her personal information without realizing her data was being aggregated for sale. Facebook is able to create a constantly updated profile of her that far exceeds marketing research data from a decade prior. All of Annie’s digital labor is transferred to Facebook, which
then sells access to her data to marketing firms, political parties, and any other business that desires access to specific and aggregated personal information.

Within the framework of Facebook, users express their ideas, opinions, and thoughts by sharing text, images, and videos with a network of friends. Over time, a certain type of culture emerges that is specific to Facebook. You are able to choose what you “like,” affirming to all within your network that you approve of some things. Additionally, the absence of “liking” something might be a measure of what you don’t like. If you “like” Barack Obama, but you haven’t “liked” Donald Trump, you’ve expressed that you’re most likely an American Democrat not only to the Facebook algorithms but to your friends and family. If you “like” Pope Francis as well as your local Catholic church, but you haven’t “liked” any other religion or tenet of Christianity or other churches, then you’ve indicated that you’re most likely a Catholic, along with many of the presuppositions that come with Catholicism. The duality of the “like” button, and absence thereof, has shaped digital culture since 2006.

What happens, then, when a social media bot is created for the sake of artistic experimentalism by gaming the “like” button on Facebook? What is the result when a bot ends up “liking” everything? How does this affect rhetorical agency and the transfer of agency between actants and agents? One of my first case study interests turned out to be one that continuously challenged my thoughts and expectations as I investigated how some social media bots served as functions of digital art as they performed actions of experimentalism. In fact, the creator of the Facebook [loveMachine] social media bot – Julien Deswaef – identified himself as an artist, yet he found it to be a challenge to convince art galleries and institutions to accept or identify his algorithm as a work of art. The concept of algorithms as art, I discovered, had implications for agency through the use of experimentalism. Furthermore, the [loveMachine] bot
sought to deliberately complicate more traditional modes of rhetorical strategy. By following the trajectory and life cycle of the bot, I was able to see how humans adapted their rhetoric when faced with this disruption. By tracing agents and actants as they exchanged actions with one another, I was able to assess the transfer of rhetorical agency around the [loveMachine]. By using the terms agent and actant, I pay homage to Bruno Latour (2005) and his nod toward people (agents) who are acting, but not necessarily on behalf of someone else; they are actants, but not all actants are agents. Actants are artifacts that are objects or creations that can influence agency within a system.

This case study (especially when contrasted to the Wikipedia Addbot) became a great example of how expressive algorithms that mimic human behavior challenge conventions of rhetorical agency. It also led to insights of the unpredictable outcomes of experimentalism as it challenged rhetorical traditions. It did these things in the following ways:

1. When human users knew the [loveMachine] was a bot that took over Deswaef’s Facebook account, some of them still treated the account as being run by a human, or an amalgamation of human and algorithm, expressing disgust, dislike, frustration, or amusement, for example. Some wrote to the bot’s wall, expressing their emotions and feedback.

2. When human users didn’t know the [loveMachine] was in control of Deswaef’s Facebook account, they assumed there was a motive behind the random “liking” of pages, photos, and comments, and sometimes the human users attempted to apply reasoning behind the eccentricities behind “liking” everything.

3. The algorithm’s experimental intent created results that weren’t predictable in any sense, exemplified by the curating and building of gay communities in Tel Aviv.
and Islamabad who interacted with the algorithm as they sought out other gay men to connect with.

**EARLY OBSERVATIONS**

When programming algorithms, there appear to be two opposing forces on an X-Y grid: planning vs. experimentalism. The less planning involved, the more experimental the algorithm. The more planning involved, the more the algorithm is viewed as a tool rather than as what Bogost (2007), called an expressive procedure, or in this case, as a procedural program mimicking human-ness in an artistic way. In the art world, a similar argument pitting planning vs. experimentalism could be made prior to the advent of the microcomputer. Consider Jackson Pollock—he originally studied under the guise of being a regionalist—a style devoted to conservative, appealing depictions of American small-town sensibility. But after viewing the abstract expressionist movement as it unfolded in Europe, he decided to experiment with artistic procedure, which resulted in his now-famous drip-splash paintings. The art world was fascinated. Pollock was explicit in his experimentalist philosophy in the following quote from 1947, featured in a 1998 exhibit at the Museum of Modern Art in New York:

> When I am in my painting, I’m not aware of what I’m doing. It is only after a sort of “get acquainted” period that I see what I have been about. I have no fears about making changes, destroying the image, etc., because the painting has a life of its own. I try to let it come through. It is only when I lose contact with the painting that the result is a mess. Otherwise there is pure harmony, an easy give and take, and the painting comes out well. (Pollock, 1998)

Pollock’s paintings were and are value-laden, artistic objects, or artifacts. They are also rhetorically dependent upon the situation of the person viewing the works—how the viewer is feeling, and what they are experiencing in life when observing the art; their history and previous
experiences with art; their perceptions of the artist as a person; their knowledge of the painting and understanding of the time period; their ability to see and see a full range of colors; and a number of other environmental and physical factors. When Pollock’s paintings were photographed, turned into prints, and made available in print media, the dissemination of his works entered into a dynamic, fluid relationship with the public. His works were open to public and scholarly criticism and feedback, and further experimentation by other artists. This dynamic distribution of energy was rhetorical agency in action.

Pollock’s celebrity status was partly because of the push-and-pull of his art’s agency as the public sphere grappled with how to approach Pollock’s work: was it art, was it bad art, or was it another work in a growing line of confusing abstract art? Was it pleasing to view? Did it add value to the conversation of art and commentary?

In 1949, when *Life* magazine asked if Jackson Pollock was “the greatest living painter in the United States,” the resulting outcry voiced nearly half a century of popular frustration with abstract art. Some said their splatter boards were better than Pollock’s work. Others said that a trained chimpanzee could do just as well. A Pollock painting, one critic complained, is like “a mop of tangled hair I have an irresistible urge to comb out” (Ouelette, 2001).

Ouelette’s article went on to explain that there was a scientific complexity to Pollock’s paintings, according to physicist and art historian Richard Taylor. Taylor suggested through data analysis that Pollock’s drip-splash paintings exhibited fractal patterns that were aesthetically pleasing and complicated, not unlike nature. Coast-lines and patterns left behind by chaotic systems exhibited similar fractal patterns seen in the drip-splash outlines of Pollock’s works, which varied in fractal complexity, Taylor asserted. There was a natural beauty to the paintings not readily apparent to observers, but aesthetically pleasing, nonetheless.
In short, Pollock’s experimentalism and the agency of his art still reverberates 75 years later. Agents continue to apply new actant-methods of observation to understand, appreciate, remix, or discredit his paintings. His works have fostered mimicry: current painter Robert Miskines has made a career out of copying Pollock’s methods. But we don’t have to look toward artists when industry has incorporated the drip-splash aesthetic into many folios of interior decoration and design, with prints and panels readily available through online stores like Etsy. The “mop of tangled hair” has since become ubiquitous in everyday American culture.

What does this type of artistic experimentalism have to do with social media bots and agency? There is a category of artists who are loosely defined as digital artists, glitch artists, or new aesthetic artists who use digital technologies to test the limitations of information systems in various formats of experimentalism as they make cultural or aesthetic statements. The philosophical leap between rhetor, author, programmer, and artist isn’t necessarily a lengthy chasm, especially when we consider agency in action. Artist Matthew Plummer-Fernandez, for example, created the three-meter sculpture “Token Homes” (figure 4) by “glitching” a stereolithography file (a computer-aided drafting file format) and printing the design on thermoplastics via a 3D printer. Plummer-Fernandez created the sculpture so that people could purchase houses they could not live in, “creating a housing estate where the housing function is removed” (Youngs 2018). Living in London, Plummer-Fernandez found himself to be part of the “rent culture,” unable to purchase a home or property due to high real estate prices. Because of high prices Londoners are no longer purchasing properties to live in, instead choosing to view real estate as an investment property, or a habitat to temporarily reside in.
Social media bots fulfill many types of functions, and one of those functions is art—a near limitless appeal to the emotions and senses. Art more often than not challenges the conventions of rhetorical agency as an artistic artifact or agent, redirecting our attention as it disrupts convention. Whereas Pollock challenged the current aesthetic through visual drippings, Plummer-Fernandez harmonized technology and sculpture with social commentary. This leap
into using technology as a social action isn’t necessarily new, but the rapid movement forward into the digital age is opening new frontiers in art experimentalism online.

I thought it important to look for potential case studies where the algorithm programmer created a bot to serve an artistic function, and discovered early on the Facebook [loveMachine], which serves as the case study for this chapter. The [loveMachine] was an experimental bot designed to game the vast and complex Facebook social media platform as the designer, Julien Deswaef, sought to make an artistic statement on digital values. Carefully curated over the course of a year, the bot interacted with thousands of users. Much like static works of art, the bot engaged with a diverse audience, provoking discourses both positive and negative. It also cultivated many reactions and uses the designer did not – or possibly could not – conceive. And, like many of the more traditional media of art, the bot was created in response to something—in this case it was developed as a response to the Web2.0 Suicide Machine (which I will explain more in detail), as well as the triteness and mechanical action of users “liking” things on Facebook.

The bot serving as an art intermediary begs us to return to Cheryl Geisler’s views on rhetorical agency: “Even if one concedes that rhetorical action is efficacious, a question can still be asked concerning the rhetor herself—to what extent does her skill matter (p. 13)?” Geisler ponders the role of a rhetor taking stock of resources and marshaling them into a conscious effort of deliberative strategy and responding to changing conditions (2004). Orators and writers levy agency by assessing their audiences and summoning forth the appropriate resources to call them to action. We need to consider the rhetorical velocity of when speech or essay becomes an artifact manipulated by print and digital media in an act of remix and circulation. How is a rhetor able to respond to shifting circumstances? Even further, how is rhetorical agency altered when a
rhetor’s artifacts are manipulated by someone other than the speaker or writer? In the 1950s we couldn’t take a Pollock painting home with us to cut up, paint on, obscure with paint thinner, or use in experimental settings. But today we can take a high-definition image of one of his paintings and perform an unlimited amount of experiments with Adobe Creative Suite by printing it and manipulating it and transferring it over the digital cloud endlessly. I argue that digital manipulation has pushed the rhetor-as-artist farther away from engaging with an audience, however, the audience has become more engaged as rhetors as they mix and remix digital copies through the use of digital applications. Granted, Pollock is dead, but his voice today would be competing against digital and industrial reproduction and manipulation of a scale impossible to reconcile. This type of emergent reproduction, commercialism, and industrial ennui was captured soon after by Andy Warhol’s pop art diptychs and prints.

What happens when an artist is working with digital media to challenge the environment in which the artwork resides? The Facebook [loveMachine], conceived of and programmed by Julien Deswaef, is a Facebook bot—an autonomous program designed for one purpose in mind: to click the “like” button on everything in a Facebook feed. In his “Why?” section of his production notes (see figure 7), Deswaef offers what seems to be a tongue-in-cheek answer to the purpose of his bot:

So, because I have so many fb_friends I want to please and so little time, I’ve decided to automated [sic] my “likes.” So Facebook will have to deal with it (whatever that means) and my fb_friends will be comforted that what they do is right. And I couldn’t wish them better (n.p.).
To wit: Deswaef created the bot to reassure his friends that everything they did was “right,” and they should derive comfort by the process of posting. Although it seems that this was Deswaef’s original intent, I believe that he quickly discovered that his social art experiment was more complex than he anticipated. He theorized Facebook’s commodification process would be disrupted. He also wanted to bring attention to the fact that the process of “liking” something on Facebook had become dumbed down to a simple, semiotic image. “We are using the ‘like’ button all of the time, and for different purposes—’hey, I exist’; ‘hey, I like this’; ‘have you seen me?’ If I ‘like’ everything [with the bot], it solves my problems” (personal correspondence, 2018). Additionally, Deswaef created the bot to disrupt Facebook’s oversimplifying “like” button and procedures, while calling attention to the oversimplification and laziness of this sort of discourse.

Deswaef adds that his bot was inspired by a workshop he participated in with “moddr_” the creative group behind the creation of the Web2.0 Suicide Machine—a bot designed to help...
humans create “digital suicide” by eradicating their online personalities and media. Deswaef is a digital artist, experimenting with digital parameters and cultural phenomena. His bot serves as a good case study to explicate—it plays upon the conventions of rhetorical agency by disrupting the process of communication and expression. In Facebook, particularly in 2014 when the [loveMachine] was officially launched, the “like” button was still novel and used as a form of digital and emotional currency. Other users measured the interactions of their posts by likes and comments. Businesses, groups, and celebrities also measured their popularity by broadcasting how many likes they received on their web pages. The blue “thumbs up” became synonymous with digital approval and popularity.

When Facebook users “like” something, they are expressing a digital variation of approval or acknowledgment by clicking the thumbs-up symbol. Imagine, then, of a human-to-human interaction, where Person 1 says, “It’s my birthday today,” and Person 2 gives a thumbs up. This personal interaction is relatively similar to a digital interaction that would appear on Facebook. Conversely, consider a real-life interaction that mimicked the [loveMachine] bot: What if Person 1 said, “My partner just left me for someone else,” and Person 2 gave a thumbs up? This interaction just defied conventional rhetoric. Person 1 disclosed a rather personal statement, where s/he appeared vulnerable and open to receiving support, but instead of offering support, Person 2 dismissed the personal statement with a confusing gesture. Person 1 is now enmeshed in an alien part of discourse—figuring out why Person 2 gave such an unexpected response. Did Person 2 not understand? Did Person 2 not like Person 1’s partner or relationship? Was Person 2 being mean-spirited? Instead of receiving support for her statement, Person 1 is now forcing herself to find strategies to figure out the meaning behind the thumbs-up.
The disruption of discourse only takes one click from a bot, but the labor of strategizing how to respond to this disruption far exceeds that of a click that was automated by a bot. In any normalized discourse, the distribution of agency would be shared among peers or friends. “My partner just left me for someone else,” said Person 1. “I’m so sorry, how terrible,” said Person 2. “How can I help?” Together they are working toward solving a problem together in a shared experience. Behind the digital veil, however, when a bot is pretending to be Person 2, yet it is purposefully disrupting discourse, Person 1 is not able to participate in an equal distribution of agency. Instead, Person 1 is the recipient of a false call to action—attempting to figure out the meaning behind the thumbs up. Meanwhile, the bot continues to automate these disruptions, offering hundreds of confusing thumbs-up an hour to dozens upon dozens of real Facebook users.

PROGRAM DESCRIPTION AND HISTORY

Julien Deswaef’s original concept was to “build a small bot that will interact with [his] fb_friends.” (2011). The first instructions for the bot were simple: once Deswaef logged into his Facebook account, the bot found the first “like” button available to it (through another person’s status, or a friend-of-a-friend’s status) and the bot would click it. The [loveMachine] was conceived as a type of answer to the Web 2.0 Suicide Machine—a bot that assisted users in purging their Web presence. The Web 2.0 Suicide Machine landing page does not mince words (figure 8).
Whereas the Suicide Machine would remove friends, delete information, change passwords, and update your status to let everyone know you’ve “committed Facebook suicide,” the [loveMachine] was designed to “like” everything—representing the opposite extreme of a digital suicide. Deswaef said that the Suicide Machine was popular in artists’ circles in Belgium, where he was residing at the time. He met one of the hacktivist-artists of the group moddr_, who introduced him to the website-automation software Selenium. Deswaef, like the moddr_ artists, used Selenium to test user interactions on Facebook, a complex platform. Deswaef said, “The goal is to avoid the Facebook API [Application Programming Interface]; it is the master API and tries to control everything. I wanted to avoid being ‘flagged out’ (detected as performing automated functions) by Facebook.”

Interested in the commodification and rhetoric of the “like” function in Facebook, Deswaef sought to disrupt the system by having the bot automate the process of liking statuses, images, and other media posted by other users.

---

1 Video correspondence with Julien Deswaef, January 2019.
The like button was more than just liking something on Facebook. [It] was all over the news, and all over the media. Nobody was really talking about the user experience. The like button was misunderstood, and it was more than just liking a post. It is a metaphor for Facebook. I found myself asking, “Once you automate this thing and you start liking everything your friends do, what are the consequences?”

At the time this bot was programmed and executed, Facebook had only two interaction options: by comment or by liking a post. The [loveMachine] clicked the like button for all posts, no matter the content, time, or circumstances. Deswaef knew that he would disrupt normal Facebook interactions by sometimes performing counterintuitive actions, which is why he kept a blog over the period of a year, but he wasn’t quite sure what to expect. He thought the program might be fascinating to some friends, and he figured he might alienate other friends. He was unprepared, however, for a number of interactions.

The [loveMachine] was “officially” launched February 4, 2014, on Facebook’s 10-year anniversary, after trial runs and testing by Deswaef. He used the Facebook nom-de-plume of “Juego Requiem,” and he seeded the algorithm with 447 of his own friends, 14 followers, 908 “liked” pages, and 11 groups. The bot ran in three-hour intervals each day, liking as much content as possible during those intervals. Deswaef also accepted any friend requests that he received during the bot’s year-long run, while not interfering if anyone “unfriended” him (asking them why, or attempting to be Facebook friends with them again). After Deswaef reached 1000 friends (day 212), he changed his privacy settings to “public.”

---

2 Video correspondence with Julien Deswaef, January 2019.
ACTANTS AND ACTANT CHARACTERISTICS

The following actants and their characteristics are noted here because I used them to create ontological maps of rhetorical agency and lines of communication and persuasion (figures 9, 10). The actants are defined in the maps as people (blue), code/algorithms (yellow), and algorithm functions (red). The use of the ontological map follows the spirit of Object-Oriented Ontology (OOO) and reductive flat ontology. A flat ontology gives objects a place in the “democracy of objects,” reduced to a basic form that is independent of human interaction.³ As I mentioned in the introductory chapter, creating a list of all networked things – objects – allows them to exist temporarily by themselves and free from the influence of other actors. They exist within a space where we can regard them for what they are and not their potential.

What is a flat ontology, and how does it appear in the world of computing and digital spaces? Bogost used an object familiar to him – an early 80s Atari video game cartridge of “E.T.” – as he revealed its flat ontology, which, summed up significantly, incorporates user-input radio frequency modulations; its mask ROM hardwired into an etched wafer; a molded plastic cartridge; a consumer good, packaged and marketed; a system of rules or mechanics designed for a user experience; an interactive experience; a collectible; and an artifact of the videogame crash of 1983 (2012, p. 18-19). In the introductory chapter, I mentioned the ontology of a book: the binding, paper stock and quality, typesetting, font, etc. Even though the book can be regarded as an object, it can be further reduced to these separate objects as part of the book’s assemblage. Bogost states that the video game is not a single object, but an assemblage of Latourian “irreduction” (p. 19)—comprised of the aforementioned objects and actants that carry weight

individually and as part of a network, although these weighted merits are not necessarily equal. I
applied the process of irredution to the digital ecology of the [loveMachine] as part of
Facebook. By doing so, I was able to map actants as they existed in relation to one another in a
dissembled state of being.

As humans exhibit interaction with one or a number of these components, they form
networks of agency. The list below denotes primary actants in a networked assemblage of the
Facebook [loveMachine]:

- Julien Deswaef – Programmer and digital artist.
- Matthew Plummer-Fernandez – Fellow digital manipulation artist residing in
  Belgium. Plummer-Fernandez provided feedback throughout Deswaef’s
  inception and execution processes. Plummer-Fernandez also participated in
  other methods of artistic experimentation on Facebook, including glitching text.
- Facebook Platform – The Facebook API, unnamed moderators, and overall
  brand. Facebook constantly updated its algorithms to detect, root out, and shut
  down automated functions within Facebook. Deswaef had to revise his code
  and functions of the [loveMachine] in order to safely operate without attracting
  the attention of Facebook’s own automated functions. For example, he would
  run the [loveMachine] for three hours at a time, or the large number of “likes”
  would be “red-flagged.”
- Facebook friends – Human users and other bots interacting with “Juego
  Requiem” (Julien Deswaef). Deswaef changed his name to Juego Requiem, but
he kept his friends, liked pages, and other media as “seeded” information to experiment with.

- “Photos that I like” Facebook function – A utility of Facebook in 2014, which would great photo montages/collages of “liked” images from friends and friends of friends, and publish them on Juego Requiem’s wall. The collages became indicative of patterns of Facebook friends and groups that interacted with the bot.

- Tel Aviv gay community, a network of Facebook friends – The “photos that I like” function and “like” function managed to assemble members of a community into social spaces as they used images and likes to extend their community. In this case, mostly younger, adult men from the Tel Aviv area in Israel used the “photos that I like” function to find other men and plan nightclub events and interactions.

- PhantomJS – “A headless web browser scriptable with JavaScript” (phantomjs.org). It is mostly used for page automation, screen capturing, headless website testing, and network monitoring (2019).

- CasperJS – A JavaScript scripting and testing utility that is installed within PhantomJS.

- Other Facebook bots – At times the [loveMachine] connected with and interacted with other Facebook bots, which were not easily or readily defined, but are mentioned for the purposes of research.
• Islamabad gay community, a network of Facebook friends – Drawn together due to circumstances similar to the Tel Aviv gay community, the Islamabad gay community was more secretive due to cultural oppressions. They faced severe repercussions for expressing homosexuality, and they apparently relied upon social media to discreetly reach out and connect with other gay men in Islamabad in the world, and to express themselves.

• Anna Li – Li was Deswaef’s first GitHub “pull request,” which meant that she was contributing code and feedback in order to make the [loveMachine] bot run better. It marked a collaborative turning point for Deswaef. Li wrote on Github in 2014: “The newly added install.bat calls install.ps1, which auto-downloads and installs phantomjs and casperjs to the [loveMachine] directory. Created for people who don't have phantomjs and casperjs or know how to add environmental variables, which are like, most people...”

• Andrew Nakas – He noted Deswaef’s “liking problem”; “The love machine has kind of turned into a hate machine. It likes things then unlikes them now for me giving people notifications with no visible likes.” This feedback allowed Deswaef to rethink his bot’s interaction with Facebook, which was constantly revising its code in order to detect and shut down bots and automated functions.

• Three-step “dislike” – Deswaef blogged about the three-step version for disliking posts, which was the precursor to the “angry” emoji that appeared several years later. Disliking a post was not a simple click or process, which distorted real reactions to posts as being either liked or ignored.
• Robin Dunbar – Researcher Dunbar wrote that “148 is precisely the maximum number of friends with whom a person can have a stable relationship at a given point in his life. Quantity is defined by Robin Dunbar, a British anthropologist, who determined, after studying behavior in primates, that this value was relative to the size of the neocortex” (March 21, 2014).

• Friends of friends – Because the bot would like anything in its feed, it would introduce itself to Facebook users through the sharing of photos, media, and comments by a common friend or page. This is how the bot gained new friends at such a fast pace. It’s important to note this sort of networking is critical for understanding rhetorical agency, networking, and persuasion.

**ONTOLOGICAL MAPS**

The use of ontologies is helpful to trace actants and rhetorical agency. Because Deswaef kept a detailed blog, we are able to easily trace more of the network surrounding the [loveMachine]. For example, there were key individuals that shaped Deswaef’s thinking, programming, and revision efforts on his code. Anna Li was the first GitHub “pull request.” Andrew Nakas noted a “liking problem,” where the [loveMachine] would like a post, but Facebook would covertly delete the “like.” Until Nakas pointed this issue out, Deswaef assumed his bot was functioning normally—by all measures it was, but Facebook’s algorithm recognized the automation and sought to undo Deswaef’s bot’s work. These individuals worked to shift rhetorical agency and the role of other actants—namely Deswaef as rhetor, and the [loveMachine] as an object-rhetor.
In Figure 9, I have denoted the actants and agents outlined by Deswaef through interviews and his blog, and followed up on information others may have published in reference to either Deswaef or the [loveMachine]. Hegemonic influences are triangulated on Deswaef, his bot, and Facebook—three main actants that exhibited a push and pull of rhetorical agency. The most repetitive and consequential actions were between Deswaef and Facebook—Deswaef sought to disrupt the intent of Facebook through the use of his bot, and Facebook sought to suppress disruption. As Facebook’s algorithm was refined to detect automated content, such as the [loveMachine], its code was adjusted, which resulted in the removal or suspension of the bot. Deswaef would adjust his bot’s code to different parameters, such as the number of “likes” it would give within a certain time frame, and attempt to stay within Facebook’s automation detection limitations.

Figure 10 is a closer look at the transfer of rhetorical agency as the [loveMachine] and Facebook competed for control of the “like” button and function. I placed the objects in this format to most efficiently display how certain networks of unfamiliar Facebook users befriended “Juego Requiem,” just as Facebook acted to eliminate the [loveMachine]. Because the [loveMachine] would “like” everything it came across, it confounded the Facebook feature “Photos that I like.” This feature created a collage of images based upon what users liked, through the use of the “like” button. The collage became a barometer to which Deswaef could see what randomness the bot engaged. However, Deswaef noticed more and more images of what appeared to be gay, Middle Eastern men (figure 11). This discovery led Deswaef to the realization that instead of achieving randomness, the bot fostered the creation of outside social networks, which included two differing gay communities: one in Tel Aviv, Israel, the other in Islamabad, Pakistan. On September 4, 2014, Deswaef posted that he had surpassed 1000 friends,
Figure 11: Ontological map of rhetorical agency and immediate network surrounding the [love] Machine
and it was at this point he noticed that many of his earlier friends had come from the Tel Aviv gay community, and he speculated that this group somehow connected a more covert gay community in Islamabad to the [loveMachine].

At first I was getting a lot of friend requests from a bright and colorful gay community in Tel Aviv. I would get messages to go to private parties. Then I noticed I began to get a lot of requests from [gay community members] in Pakistan. Islamabad. As they became friends the bot would like their posts and somehow connect the gay community together. Or maybe they were just looking for one another. I don’t know.4

Deswaef wrote in his blog then that he inferred that the religious suppression of homosexuality in countries like Pakistan might have been the reason for such a pronounced Facebook presence. Facebook appeared to be a safe space for gay men to connect with one another, and the [loveMachine] helped to provide that catalyst of connectivity.5

Theoretically, the formation of these gay networks might exist not just because of the “like” button and the [loveMachine], but through a combination of the “like” button and the Facebook feature “Photos that I like.” As the [loveMachine] began to automate the “liking” process, it might have caught the attention of men posting selfies online as a means to cautiously put themselves out there for someone else to acknowledge. In a suppressed, conservative society hostile toward gay men, the digital spaces of social media platforms can help amplify the need to reach out to other gay men also cautiously attempting to connect with others as they search for friendship or relationships. As the bot “liked” these images, the men befriended the bot, who they only knew as Juego Requiem, another male that “liked” their images. Facebook then curated the “likes” with the “Photos that I like” feature, further amplifying the notion that Juego

4 Video correspondence with Julien Deswaef, January 2019.

5 Deswaef’s blog post, September 4, 2014.
Requiem was comfortable “liking” several images of young, Middle Eastern men. The more the network grew, the safer Juego Requiem seemed to be.

Figure 12: Ontological map of rhetorical agency and the Facebook “like” button
Figure 13: Sampling of "Photos that I like," photo montage, during the [loveMachine] experiment, courtesy of Julien Deswaef
From the standpoint of rhetorical agency, the bot’s automating feature had an unintended consequence—it both connected to at least two networks of gay men, and it created a space where these men seemed to find safety or comfort. This last part is purely speculation, since I was not able to connect with any of these men in the networks to ask them directly, but the [loveMachine] didn’t appear to connect any other group or subsect of people, as far as any evidence points to, according to Deswaef and my observations and data collection. When the bot “liked” a celebrity, or religious icon, or piece of pop culture, it didn’t connect to a network of fans or acolytes. It didn’t connect Juego Requiem to a network of women, even though Deswaef’s likeness, through the bot, was “liking” photos of women, too. The bot was indifferent, but it still offered a space for these men to congregate and connect online. What does this say about rhetorical agency? Automated, digital processes can open up cultural spaces for those who exist in oppressive environments. The power of automation can have the effect of projecting strength, acceptance, and assertion, even if these processes are not human at the core. As I will demonstrate in the following chapter on the Peñabots, a chorus of voices can project unification and agreement in areas that have a minority of voices. In the case of the [loveMachine], the minority happened to be two distinct communities of gay men. Deswaef’s algorithm performed a function that was not considered safe for gay men—it openly and explicitly “liked” photos of men, perhaps projecting bravery. As more men “friend” Juego Requiem, the more the [loveMachine] “liked” the photos of the men within its feed. This created a feedback loop of safety, comfort, and community—empowering gay men to connect with one another through the nexus point of Juego Requiem. The Peñabots, however, used an automating function to project disinformation under the guise of community safety, as I shall explain in Chapter Three.
We see automation in other social media spaces, where algorithms bludgeon us with information, such as disinformation or fake news. When we see disinformation repeated over and over again, and connected to other disinformation links and media, we are unable to cross-reference them to what we know is true, and we end up doubting ourselves. This only exists because of an absence of authority and power—a lack of education or experience, for example. One or two people, for example, might share disinformation with us at work, but we are able to change environments and networks and ask someone else about the validity of these claims, or compare these claims to a reputable source. But in the era of automation and algorithms, a social media bot, or group of bots, can overwhelm our ability to discern what is real and not, projecting authority and agency simply by overwhelming us. A bot or botnet can also project authority and agency through automation for good or benign causes. Matthew Plummer-Fernandez, one of the digital artists with whom Julien Deswaef connected, created the Twitter bot @happyb0t, which was self-learning, expanding its vocabulary based upon finding tweets with “:)” in the contents (the pre-emoji way of expressing a smile, or of happiness). The @happyb0t projected happiness and positivity—along with some confusing statements that also seemed upbeat and positive—with smiley faces and exclamation points, attracting over two hundred followers. It automated happiness and created a space for others to enjoy discourse about the whimsy of the bot and the happy statements it offered.

The gay men of Tel Aviv and Islamabad were denied authority and agency in different degrees, which affected how they connected with one another. According to Deswaef, the Facebook Tel Aviv gay community used the social media platform to connect and socialize not only in the digital space, but they used the digital spaces to plan to meet in physical spaces, too. The Facebook Islamabad community appeared to be more interested in simply connecting, which
was mostly limited to a digital space. The laws forbidding homosexual activity (such as sodomy) in Pakistan stem from a combination of Victorian-era British colonial law and sharia law, and can result in prison sentences or the death penalty, though such sentences are rare (Carroll and Mendos 2017). Israel, conversely, has recognized same-sex relationships for 30 years, decriminalizing them in 1988. The absence of agency in the Tel Aviv gay community appears to be a natural by-product of being a minority group with some social stigmas and judgments attached. The Islamabad gay community experiences a much larger void of agency and authority because of strict laws and a religious authority that upholds such laws and stigmas.

ENVIRONMENT

Deswaef attempted to adhere to the following rules, unless changes in the Facebook algorithm forced him to adapt or deviate:

1. Accept any friend requests (except from click farms and underage [people]).
2. Accept all events invitations.
3. Accept any page suggestion but unfollow it immediately afterwards.
5. Never post pictures of friends or family and never tag anyone in a photo.
6. Never accept or send game invitations.
7. Never publish anything that I don’t consider public.
8. Only use the loveMachine for liking content.\(^6\)

\[^6\] Deswaef’s blog post, February 19, 2014.
It is important to note that when the [loveMachine] was launched, the digital environment of Facebook only allowed for two types of immediate feedback of posts: either by commenting or by liking a post. At present, users can express their opinion through emoticons—a thumb’s up “like”; a heart “love” icon; a laughing face; a face that expresses “wow”; a crying “sad” face; and a red, surly “angry” face. Additionally, there are a number of ways users can now comment on a post or status update, such as with a Graphics Interchange Format (GIF) image, a static image, video, emoji, or stickers, for example. During the [loveMachine]’s tenure, only text comments were available. Therefore, the “like” button carried much more ontological and semiotic weight in 2014, when the [loveMachine] was officially launched.

One of the first issues that Deswaef encountered was that many friends did not initially understand why Deswaef—now under the pseudonym Juego Requiem—would behave in such a manner that he would like posts that theoretically had no type of relationship to the friend. But the liking of posts had a darker side to them. When a friend posted that a family member had died, the [loveMachine] automatically liked the post, which, in 2014, was a rather distasteful thing to do. The “like” conveyed that Deswaef was possibly exhibiting a number of odd reactions: he could be

1. Expressing that he liked that a friend’s family member died
2. Attempting a distasteful joke

---

7 Deswaef noted that during the course of his year-long experiment, Facebook introduced what Deswaef called a three-step “dislike button”—by clicking the “x” in the upper right corner of another person’s comment on your content, Facebook would remove the comment and prompt the user to provide feedback to the original comment, with the pre-filled text of “Hey, I didn’t like this comment” ready to send.
3. Not fully grasping the seriousness of the post, and by liking it he was expressing callousness.

4. Suggesting that he was acknowledging the passing of a friend’s family member with a “like,” which often denotes a user has read or understood something.

Regardless of the outcomes, the reaction to the automatic liking of the post was negative. Deswaef also noted that a Facebook user – Mohammed Afzul of Pakistan – did not appreciate that his comments were being liked, writing, “Don’t like my comments:-@ / If you have to like, then like my pics only.” But when Facebook friends would question why a post was being liked, for whatever reason, Deswaef saw this as an artistic success. “It was a joy to receive these questions, because they were proving my point.”

ETHICAL CONSIDERATIONS

When considering ethics and rhetorical agency, we must return to the intent of the algorithm. What is the algorithm meant to do? In this case, Julien Deswaef intended for the bot to disrupt normal discourse and socialization on Facebook by liking any post it came across. For the most part, aside from disruption and intervention on behalf of Facebook, the bot performed its duties. However, Deswaef had a secondary motive: to provoke people to think about how they used the “like” button in Facebook, but with no end result in mind. Deswaef said, repeatedly, in interviews and in his blog posts, that he did not know what to expect, and he wanted to see what was going to happen. One could argue that Deswaef, like Jackson Pollock, worked within a

---

8 Video correspondence with Julien Deswaef, January 2019.
realm of experimentalism, or art for art’s sake. Pollock set out to disrupt a fluid, yet somewhat fixed system of painting styles, types, and tropes. Arguably, he may have meant to confound these systems within the painting he was working on, but once the painting was part of a networked system with the public, he disrupted discourse. Deswaef, however, set out to directly disrupt discourse with the public, and the bot was the tool to perform this disruption. But is it ethical to levy disruption without planning for different outcomes? Never mind those who create malware or computer viruses with the intent to harm. Consider the artist who wants to make a statement.

The ethics of disruption and invention is problematic if we adopt a modern moral philosophy. Is an invention a means to an end, or is it an open-ended argument? G. E. Moore asked in the *Principia Ethica* what type of actions should we be executing, and what types of things should exist for their own sake? (Moore 1903; MacIntyre 1998). The issue at hand is that in our [loveMachine] ontology, the non-human actants are possessive of latent agency, but only enacted when a human actant becomes part of that network. Heidegger’s plane sat idle on the runway, awaiting someone to fly it. Like the plane, actant-objects exhibit a push and pull of agency within a given digital work, and those objects include algorithms.

Consider the computer virus, trojan, or worm: the very mention of the 2017 ransomware virus *WannaCry*, especially in the days following its release, caused panic in businesses. *WannaCry* disabled networks and computer systems, offering to return precious data in return for a ransom payment. The algorithm didn’t have to be present in a computer system in order to cause panic. It caused entire information technology departments to change policies and strategies without ever physically engaging with their systems. Like the Latourian speed bump, *WannaCry* produced a behavior in response to what it *could* do. But until that virus was executed
within a network, it was actually useless. A speed bump serves no purpose unless it must be
crossed by an automobile. We exert agency upon the objects and actants around us each day.
What changes when we create algorithms to perform these functions for us? We project
responsibility to the creator(s) of WannaCry with responsibility for the creation of the algorithm-
object, primarily because they purposefully created the software to extort money from businesses
by holding data hostage. The ransomware’s execution set off a chain of events, but the process
began with its creation. Does our perspective on responsibility change if the creator(s) of
WannaCry developed the ransomware to test the security of a corporation, but someone took that
software and used it to extort money? What if WannaCry was the result of experimentation,
where the programmer developed the algorithm to see how different information systems
reacted, and unexpectedly infected other computers?

It seems that the process of invention and experimentation absolves the programmer-
agent, to a certain degree, of responsibility. Felony crimes in the United States are approached
with degrees of severity. Manslaughter is much less serious than second-degree murder, which is
less serious than capital murder. The root of the felony is causing the untimely death of another
individual. Should we approach programming responsibility and potential agency-changing
properties with the same degrees of responsibility? After all, there are varying degrees of
cybercrime. Van der Goot’s twitterbot that mistakenly issued a death threat is less serious than
hacking credit card numbers, which is less serious than having a bot commit a “swatting”
offense—tricking emergency services into responding to a false threat. The process of invention,
after all, needs to incorporate others (LeFevre 1987; Brooke 2009), so how do we reconcile
agency with one another if not through causation and degrees of responsibility?
Through the ontology map of the [loveMachine], Facebook, and Julien Deswaef, I demonstrated that these three forces levied agency upon one another. Facebook, through the use of the “like” button, provoked a desire within Deswaef to respond to this system that he regarded as a false signifier within the “giant Turing test” of Facebook. Hit interaction was situational and reactionary, and, in the spirt of the *Principia Ethica*, he wanted to do something *intrinsically good*. In our interviews Deswaef expressed a strong dislike for Facebook, both personally, and from a sociological perspective. Facebook was responsible for a dumbing down of communication, for commodifying relationships and interests, and for abusing privacy and personal information. The bot was to be a sort of self-fulfilling prophecy of self-destruction. Even though he created the [loveMachine] in response to the data-killing Web2.0 Suicide Bot, the [loveMachine] was used in a matter befitting a suicide, as exemplified by how he wanted to end the bot’s lifecycle:

I let the bot die by itself. I wanted to collect all that experience and do something. My Facebook account is now gone. It was part of my attempt to get rid of [my] Facebook [page]. It was a tactic to introduce noise into my profile. I wanted to make it into a mess that was no longer of interest to me or other, but that ended up taking a while, [because] I got 5,000 friends in a year [as a result of the bot] (personal correspondence 2018).

Deswaef believed that disrupting Facebook, even at a very small and experimental level, was intrinsically good for other Facebook users. But this good-ness is indefinable, and subject to outside forces, negative and positive, such as when the bot liked obituaries, or when it connected gay communities in otherwise hostile environments. While the overall goal of the bot was to provoke discourse, awareness, or a number of other holistically good outcomes, the immediate response to obituary “likes” was that of confusion and hurt. Conversely, the positive nature of

---

9 Deswaef’s blog post, July 23, 2014.
the bot brought an otherwise ostracized community of gay men together. Deswaef made a conscious effort to continue with the experiment, weighing outcomes against what he perceived as smaller, or somewhat inconsequential, events. The friend-actant attempted to convey to Deswaef the displeasure of the bot’s actions, and Deswaef attempted to reconcile by explaining the art project. In fact, the annoyances of the bot led several friends to “unfriend” Deswaef. He explained that he did attempt to let his original Facebook friends know at the onset that they would be subjected to this art experiment. Whether they knew this or not is unquantifiable, but the loss of friends was quantifiable. However, where Deswaef lost friends, his bot managed to accrue Facebook friends simply by liking content of friends and friends-of-friends, surpassing 1000 within a few months, and over 5,000 over the course of a year.

The bot, through automation, created a new set of expectations. Instead of Deswaef’s predictable behavior, which he cultivated on Facebook, the bot provoked a wide variety of responses—amusement, annoyance, anger, curiosity. As time passed, Deswaef’s original friends either adapted to the new interaction with the bot, or they quit their Facebook friendship. “They would say to me, ‘I don’t like it, because I don’t know if it’s you or if it’s your project that I’m [interacting] with.’” But when new friends arrived, drawn in by the bot’s liking feature, they arrived knowing what the expectations of “Juego Requiem” were: this persona would like everything they posted, he didn’t comment, he didn’t say or do anything negative, he was online almost every day, and he may or may not have shared similar networks. Once that ecosystem began to take over, the friendship circle and networks grew.

---

10 Video correspondence with Julien Deswaef, January 2019.
Ethically, Deswaef did make an attempt to inform his friends that there would be a change in digital activity due to an art experiment, but in the vast space of the Internet many friends either didn’t see this warning, they didn’t understand it, or they forgot. The bot challenged some of their value systems and appreciation for predictable behavior. The bot’s purpose was to provoke discourse and disrupt Facebook, which was perceived by Deswaef to be an abhorrent system that marketed off of emotional narratives and expression, so intrinsically, Deswaef was inventing something good or beneficent. But in social networks, he caused minor harm or provoked negative reactions, in some instances; in others the positive nature of the [loveMachine] led new friends to believe Deswaef was a good person worth befriending. The art experiment worked, in a sense. It also cultivated new social circles, which didn’t have the effect of disrupting Facebook’s commodity of “likes.” Through interviews, Deswaef expressed interest and appreciation for how the gay communities adopted the [loveMachine] as a means to connect with one another.

Also of note are the gay communities of Tel Aviv and Islamabad—outlier-actants that were completely unexpected by-products of the [loveMachine]. As Deswaef noted, it appeared that the Tel Aviv gay community – a thriving, established community that even held its own Gay Pride Day and parade – seemed to be the precursor to the Islamabad gay community joining the [loveMachine]’s social circle. Deswaef also noted how the religious power structure of Pakistan acted against homosexuality, driving gay men to adapt and find new ways of connecting with one another to express their own humanity. The [loveMachine] served as a catalyst for these men to connect with one another. What would happen if a government actor working for Pakistan was able to use the [loveMachine] to locate, discover, and imprison or execute gay men in Islamabad? The intrinsic good that Deswaef intended could have ended up being disastrous, even
if completely unpredictable. Would he still be responsible, ultimately? This is a critical moment in rhetorical agency scholarship—the ownership of such programs and the responsibility of the programmer and inventor. As we will see in other case studies, the flat ontology of even the simplest algorithms demonstrates that humans have little control over the unlimited possibilities of interactions each line of code can have with one another and other objects and actants.
CHAPTER 3: TWITTER PEÑABOTS: AUTOMATING MANY VOICES AS PART OF A STRATEGIC EFFORT

EARLY OBSERVATIONS

Information warfare is part of the social media landscape. This was made apparent by the 2016 presidential election in the United States. Twitter users have been sounding off for years about the presence of trollbots, spambots, and fake news distribution algorithms. The problem is so endemic that some Twitter users, academic researchers, and entrepreneurs have attempted to devise ways of detecting, doxxing,¹ and reporting malicious bots. Twitter, always reliant upon bots as quantitative units that count as advertising numbers in marketing policies and prices, struggles between allowing automation while purging tens of thousands of fake accounts in order to satisfy the Twittersphere—the users who constitute the majority of actual discourse. The underlying problem, however, is that the automated accounts are acting either individually to disrupt normal discourse or in large groups to shape topics and narratives. Rather than use traditional advertising, for example, to steer policy, for pennies on the dollar one person can shape policy through social media through the use of bots:

Can one person controlling an identity, or a group of identities, really shape social architecture? Actually, yes. The Web Ecology Project’s analysis of 2009’s post-election protests in Iran revealed that only a handful of people accounted for most of the Twitter activity there. The attempt to steer large social groups toward a particular behavior or cause has long been the province of lobbyists, whose “astroturfing” seeks to camouflage their campaigns as genuine grassroots efforts, and company employees who pose on Internet message boards as unbiased consumers to tout their products. But social bots

¹ Maliciously publishing an anonymous Internet user’s identifying information, such as name, address, phone number, other accounts.
introduce new scale: they run off a server at practically no cost, and can reach thousands of people. (Isaacson, 2011, n.p.)

Twitter bots are, for better or worse, an inherent part of the Twitter ecosystem that veteran users have learned to negotiate around. Some Twitter bots are designed to have a positive or neutral impact when they perform tasks such as sharing news, liking content (Chu et al., 2012), analyzing Twitter data (Satuluri 2013), or warning the public about natural disasters (Akimoto 2011; Haustein et al, 2015). However, many bots are used for tasks that humans label as deceitful, duplicitous, or nefarious. They are used to propagate viruses (Mirani, 2014; Haustein et al., 2015), disrupt dissent through obfuscation (Brunton, & Nissenbaum 2015), silence activists (Gallagher, 2015), spread fake news, and falsely increase user statistics and “influencer” status (Confessore et al. 2018).

Bots working together in a strategic fashion are collectively known as a botnet. Botnets are fascinating rhetorical devices. They can overwhelm a system with automated information, pushing a narrative that may not otherwise receive recognition. Bots carry great weight when combined with one another for a common purpose. We see these examples with marketing advertisements, creating social trends and topics, and through the use of propaganda and disinformation. It is with the more nefarious deeds that I am concerned when examining agency and botnets, as exemplified by this chapter’s case study.

But first, please allow me to explicate the value of voluminous automation and disinformation. The following example is entirely fictional. If I were to fill a stadium with actors whom I paid to say something similar to one another, such as, “There was never a ship called the Titanic,” or “The Titanic never sank because it never existed,” you would inherently know these types of statements to be not true. You learned about the Titanic in school. You’ve seen pictures
of the wreckage in the pages of National Geographic. You enjoyed the James Cameron movie, Titanic, and the British predecessor, A Night to Remember. But when person after person tells you what you thought you knew wasn’t true, you might begin to question your own logic. “Have I actually seen the wreckage?” “How do I know it’s true?” “Plenty of movies have been based on fictional stories because we love drama.” We are used to comparing truths to one another as we search for credibility, but when many different voices convey the same mistruth, our ability to compare through different points of view diminishes. The more time we spend in a given ecosystem filled with these mistruths or disinformation, the more likely we are to cast aside what we know to be true. Some botnets are employed to fulfill the strategy of coordinating this rhetorical strategy of usurping truth. And there are millions of bots on Twitter at any given year, fulfilling a variety of roles. Repetition can be weaponized.

No one is able to count how many bots there are on Twitter, but in 2015 it was reported that 16% of Twitter users exhibited a “high degree of automation” (Zhang & Paxson, 2011, p. 102) in an environment where bots aren’t easy to identify, especially when they mimic human social media behavior (Boshmaf et al., 2011; Haustein et al., 2015). Twitter self-reported that it had 305 million users in 2015, which means, according to Zhang and Paxson (2011), that almost 49 million of those accounts exhibited high levels of automation. In May and June of 2017, Twitter had suspended over 70 million accounts and was “locking” more than 10 million “suspicious” accounts per week as part of a new initiative to combat social media fraud (Confessore & Dance, 2018)—a purposeful intent to deceive others by posing as someone else (outside of fan, parody, or tribute accounts) by conduction information or financial scams or deliberately misleading the public, as delineated by the Twitter Terms of Service (2019). At the end of the second quarter of earnings in 2017, and before the implementation of its new policy,
Twitter self-reported a membership 328 million users. Clearly, there was an upward trend of bots and suspicious accounts that Twitter felt was hurting its overall image.

Automation isn’t necessarily an indicator or metric with which we can count bots on Twitter. Some accounts sit silent or generate little content over months or years, waiting for a moment to be put to use. In 2017, researchers Juan Echeverria and Shi Zhou published a paper on a relatively quiet botnet they discovered. Dubbed the “Star Wars botnet,” there were over 350,000 Twitter accounts periodically posting quotes from *Star Wars* literature; it was a botnet that had yet to be activated for a coordinated purpose. More interestingly, the fake accounts had been active since 2013 – four years at the publication of the paper – most likely waiting for a bidder to purchase them to be put to use. “It is known that pre-aged bots are sold at a premium on the black market. This means the Star Wars bots are perfectly suited to be sold as fake followers because they are already four years old and therefore more ‘valuable’” (n.p.).

Echeverria and Zhou (2017) discovered this botnet by first sampling data that showed geolocation coordinates—Twitter allows users to choose to broadcast their geolocation through tweets or by tagging pictures, events, or video with locations. By plotting these locations, the researchers looked for tweets that were located in areas that were uninhabited or uninhabitable—“seas, deserts, and frozen lands” (n.p.). The researchers assumed that the “uniform, low-density blankets of tweets over uninhabited areas were likely created by Twitter bots” (n.p.). Secondarily, Echeverria and Zhou manually reviewed the Tweets coming from these anomalies, and they noticed a pattern of *Star Wars* quotes. In their experience, some bots generated tweet content by pulling quotes from literature or other sources of textual language in order to confuse bot-detection algorithms. This particular botnet was designed to maintain a low profile in order to avoid Twitter’s detection algorithms. The bots were trained to tweet in a manner that
somewhat resembled human activity. “It seems the Star Wars bots were deliberately designed to circumvent many of the heuristics underlying bot detection methods. For example, contrary to popular assumptions on bots, the Star Wars bots do not have any URLs in their tweets, they never mention or reply to other users, and they only follow a small number of friends in comparison to random users” (Echeverría & Zhou, 2017, n.p.). But a big question remains: for what purpose was the botnet to serve, ultimately? We don’t know how much a botnet as large as the Star Wars bots would cost on the black market, but we are learning more each day about the importance of coordinated social media campaigns attempting to change public opinion. Someone wanting to change public opinion, shape discourse, change election outcomes, or push policies could find great value in a large botnet. The example of the Star Wars botnet brought me to the case study of not one social media algorithm or bot, but a network of similar or identical bots that are used to coordinate behavior.

PROGRAM DESCRIPTION

In 2012, bots spread political propaganda for Mexican presidential candidate Enrique Peña Nieto during his campaign, which led to the nomenclature “Peñabots,” but the infrastructure and willingness to use botnets was already in place long before Nieto began his presidential campaign. In 2006, President Felipe Calderón barely defeated Andrés Manuel López Obrador, partly because large corporations paid large amounts of money to craft and broadcast attack advertisements against Obrador (McKinley, 2006). Kirk Semple and Marina Franco (2018) reported in the New York Times that some analysts believe that the resulting Mexican election reform laws from 2007 caused a change in strategy to control social media: hardball politics moved from traditional marketing arenas to the Internet, and social media was the vehicle of the
new information warfare battleground. Web developer Iván Santiesteban collected and published data on 20,000 bots (see Figure 14) in the weeks leading up to Mexico’s 2012 presidential election with Nieto (Santiesteban, 2012, Semple & Franco, 2018). By 2014, there were, by researcher Alberto Escorcia’s account, at least 75,000 Peñabots acting on behalf of then-President Nieto (Gallagher, 2015).

<table>
<thead>
<tr>
<th>Listas de cuentas falsas</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Hemos encontrado más de 20,000 bots. Son estos:</strong></td>
</tr>
<tr>
<td>* 20 mayo 2012, 4,900 bots: <a href="#">descarga lista</a> o <a href="#">vela en Pastebin</a>.</td>
</tr>
<tr>
<td>* 22 mayo 2012, 1,478 bots: <a href="#">descarga la lista</a> o <a href="#">vela en Pastebin</a>.</td>
</tr>
<tr>
<td>* 24 mayo 2012, 952 bots: <a href="#">descarga la lista</a> o <a href="#">vela en Pastebin</a>.</td>
</tr>
<tr>
<td>* 29 mayo 2012, 8,758 bots: <a href="#">descarga la lista</a> o <a href="#">vela en Pastebin</a>.</td>
</tr>
<tr>
<td>* 10 junio 2012, 1,117 bots: <a href="#">descarga la lista</a> o <a href="#">vela en Pastebin</a>.</td>
</tr>
<tr>
<td>* 30 junio 2012, 1,478 bots: <a href="#">descarga la lista</a> o <a href="#">vela en Pastebin</a>.</td>
</tr>
</tbody>
</table>

Figure 14: Screen grab from Iván Santiesteban’s webpage, where he linked to data he collected on bots acting to change public opinion weeks ahead of Mexico’s 2012 election day.

The Peñabots switched from political propaganda during Nieto’s campaign to hashtag poisoning – flooding a feed with similar Twitter hashtag topics in order to render them ineffective – in an effort to keep Nieto in power and within the public’s favor. In November 2014, for example, the bots actively worked to silence Twitter opposition to President Nieto as he sought to install family members to prominent positions within his administration. The public was criticizing Nieto’s nepotism. As Gallagher noted in her 2015 conference presentation video *Mexican Botnet Dirty Wars*, the Peñabots attacked the hashtag campaign in two ways: 1. By
“filling the trend with spam, which is a censor in itself; it drowns out the real conversation,” and 2. “… gaming the Twitter algorithm by altering the novelty index.”

More specifically, by flooding Twitter with bots that are continuously repeating the same hashtag, spam-detection algorithms are triggered within Twitter, thereby removing the trending hashtag from Twitter’s “Trends for You” window. Gallagher used an example of the trending hashtag “#SobrinaEPN” to illustrate how the botnet poisoned the hashtag campaign. In Figure 15, several bots are listed in a Twitter hashtag search using #SobrinaEPN as the search query. “Sobrina” is Spanish for “niece,” and “EPN” is the Spanish acronym for the president, and the hashtag refers to Nieto installing his niece at the state-owned energy company Pemex. Each bot is coordinating with one another, overusing #SobrinaEPN. Human Twitter users rarely use the same hashtag twice in one tweet. Here, the bots were clearly attempting to saturate the Twitter novelty index with the same hashtag, triggering Twitter’s spam-detection algorithms. Instead of trending, #SobrinaEPN was pulled from any “Trends for You” window, since it was removed for acting suspiciously. What originally began as an activist hashtag created to raise awareness, #SobrinaEPN was demoted to obscurity, lost in a sea of noise.

The hashtag, as exemplified by the Peñabots, is a rich commodity in what Richard Lanham (2006) would call an “information economy” built upon “attention engineering.” Lanham labeled the World Wide Web a place where, counterintuitively, the more people that collectively participate in its processes without responsibility, the “greener its grass grows,” which is the opposite of what occurs in a physical space, where participation and a lack of responsibility will leave the fields barren and overgrazed. In the World Wide Web, “the more people graze on it for their own purposes, the bigger it becomes and the greener its grass grows. It thus combines the power of the free market, where individual gain leads to collective benefit,
with the cooperative ownership of the cultural conversation” (p. 13). We see phenomena that
echoes these thoughts, such as “cancel culture,” where the collective masses band together to
oust someone contradictory, controversial, or unsavory—shutting down advertising or businesses
or causing a celebrity to retreat out of the public eye. Hashtags in this world of collective non-
responsibility are effective processes that offer a lot of power and agency to those who wield it
the most effectively and collectively. Successful, organic hashtags are built on activism,
entertainment, reactions, and other pathos-driven responses. Hashtags allow for organizing and
building communities, and one way to disrupt organization is to use hashtags in a similar way to
disrupt community building, which we call hashtag poisoning. And the Peñabots were very good
at this sort of tactic.

The Peñabots, Gallagher (2015) noted, performed a variety of functions:

1. Hashtag spamming
2. Creating fake trends
3. Spreading smear campaigns
4. Sending death threats
5. Spreading political propaganda

In addition to hashtag spamming or hashtag poisoning, the same botnet is able to create
its own trends by not oversaturating the Twitter platform with too many hashtags. By staying
beneath the algorithm’s threshold of automation and hashtag saturation, the botnet can create a
fake trend. For example, Buzzfeed News reported in 2018 how Mexican citizen Carlos Merlo
Figure 15: Screen grab from Erin Gallagher’s conference presentation video “Mexican Botnet Dirty Wars.” This image demonstrates hashtag attacks or “poisoning,” where bots attempt to disrupt a trending hashtag targeting Mexican president Enrique Peña Nieto’s nepotism.
founded his disinformation company, Victory Lab, in 2011 at the request of a politician who needed help with his Internet campaign. By 2018, Merlo claimed to control four million accounts on Twitter, “many of which were purchased from Russian bot agencies and given Mexican-sounding names” (Broderick & Arredondo, n.p.). Merlo demonstrated to Buzzfeed News reporters how he fabricated a hashtag (#GanaConVictoryLab, or “Win With Victory Lab”) and made the hashtag trend on Twitter—entirely through the use of fake accounts, or bots. He charged clients around $10,000 for this service, which, at the time, he claimed Twitter would charge around $100,000 (Broderick & Arredondo, 2018). I mention Merlo as a cautionary tale when we attempt to ascertain where the Peñabots originated. Merlo claimed that Victory Lab worked for different political parties as they sought to gain an advantage in information warfare over their opponents. Broderick and Arredondo also reported that Victory Lab “is one of the estimated hundreds of homegrown Mexican Cambridge Analytica-like marketing firms that are constantly filling up the country’s social media platforms with junk” (n.p.). Additionally, foreign actors also have a vested interest in disrupting campaigns hostile to their country’s ambitions, bolstering campaigns who share similar interests to their country or pitting political parties against one another in order to gain diplomatic or economic advantages. This not only includes well-known actors such as Russia or Ukraine, but presumably also the United States. Other foreign countries, like Mexico, have used pro-government botnets as they manipulate public perception. Such examples have appeared in Turkey, Egypt, and Syria (Finley, 2015).

As the United States saw in 2016-2018, smear campaigns and the dissemination of political propaganda by foreign botnets were quite effective in changing the political landscape. New Knowledge, a small cyber security firm, took note of Russian successes in automated smear campaigns and used the tight 2018 Alabama senate race between Roy S. Moore, a Republican,
and Doug Jones, a Democrat, as a sort of litmus test to try some of the same tactics. The New York Times obtained and published a story in 2018 about an internal report by New Knowledge that claimed they set up a Facebook page to attempt to shut out Republican voters and “enrage and energize Democrats” (Shane & Blinder, 2018).

The project’s operators created a Facebook page on which they posed as conservative Alabamians, using it to try to divide Republicans and even to endorse a write-in candidate to draw votes from Mr. Moore. It involved a scheme to link the Moore campaign to thousands of Russian accounts that suddenly began following the Republican candidate on Twitter, a development that drew national media attention.

Although the size of the disinformation, or “false flag,” operation conducted by New Knowledge was reported by The New York Times to be minimal from a financial and investment scope, their attempt at connecting Russian bots to Moore’s campaign received significant attention. Ironically, Moore’s Twitter account did have a significant amount of automated accounts following him during his campaign, but the identities of the bot owners were never verified or confirmed. New Knowledge’s experimental bots also tied Moore to already public news of nine women who publicly claimed Moore engaged in either sexual or inappropriate misconduct with them, even when they were underage girls. It is difficult to determine how effective the Facebook “bot” accounts were at dividing Republicans through the experimental Facebook page (even as they violated Facebook’s Terms of Service on automated accounts), but New Knowledge’s experiment makes it clear that disinformation strategies through automated accounts are here to stay and have become part of the political landscape.

Botnets are also used to send threats of violence, sexual violence, and murder. One example that was backed up by visual data from Alberto Escorcia, as reported by Erin Gallagher (2015), included “Mexican academic and blogger Rossana Reguillo,” who was the subject of
repeated and massive threats of violence through the use of memes, images, and text by the Peñabots. Gallagher said this was in response to Reguillo’s activist stance as she protested the disappearance of 43 male students from Ayotzinapa Rural Teacher’s College—a tragedy that marked the disappearance and presumed murders of leftist, male educators at the hands of police. The incident was a dark mark on Nieto’s presidency and one that his administration was determined to eliminate or, at least, obfuscate. “International investigators arrived in 2015 at the invitation of the Mexican government. But after they contradicted the official version of events, they said, the government began a campaign of harassment and stonewalling that made it impossible for them to do their work” (Villegas, 2018). The Peñabots added to the campaign of harassment, specifically with the targeted death threats to Reguillo. It was clear the incident would not be solved in a manner that favored the administration, so the botnet went to work to destabilize resistance and activism. By targeting Reguillo with such a massive force of voices, it was sending a clear signal not just to Reguillo, but to anyone else who might protest with her: you will be harassed and threatened at a level never seen before. The sheer volume of harassment was enough to drown out any other voices of support, or of any other topic of discussion.

Twitter bots are simple programs that are usually coded in languages such as Ruby, Python, and Node.js. They can also be automated through the Twitter platform itself, without any personal coding, to perform tasks such as sending an automatic direct message (DM) when someone follows your account or redirecting users to a new Twitter handle or craft “Out of Office” replies. Twitter’s Terms of Service (ToS) offers “Do’s” and “Don’ts” of automation, appearing to steer users toward innovation, helpfulness, and good user experiences; and it warns against violating policies or user privacy, going outside the Twitter API, abusing other users, or spamming other users, for example (Twitter, 2017). However, Twitter is notorious for being
abused, with millions of accounts being purged in the past two years for violating the ToS and various policies, including violations of automation, abuse, and coordinated spamming.

It is unclear how Twitter enforces its various policies, especially when considering automation, ownership, abuse, and authority. Twitter has been quick to suspend some accounts that are deemed abusive, as exemplified by other Twitter users complaining that one of their favorite accounts (a celebrity or activist, for example) has been suspended or removed. In 2016, self-proclaimed “alt-right” activist and Breitbart tech writer Milo Yiannopoulos was permanently banned from Twitter in response to persistent, racist remarks to and about Saturday Night Live actor Leslie Jones:

The permanent suspension [was] the social network’s strongest response against those who break its abuse and harassment policies, but we know very little about how Twitter decides when to use it. Twitter generally doesn’t comment on questions about individual accounts, citing privacy and security reasons. Users who are perma-banned from Twitter are, as a practice, not told which tweets of theirs prompted the ban—only that they are banned, that their accounts will not be restored, and which part of Twitter’s rules the company said were violated. Twitter also does not have publicly available guidelines for the threshold that must be met for a permanent ban. (Ohlheiser, 2016, n.p.)

Similarly, bots pushing spam, retweets, disinformation, high volumes of tweets, or other policy violations are often deleted without warning or provocation. It appears that abusive behavior will prompt an explanation, but automation violations might not receive an explanation of suspension or permanent deletion.

Twitter states that if a user wishes to act as a developer, using algorithms and bots for education, entertainment, or assistance through the Twitter Application Program Interface (API), then they must be approved as a developer, subjected to Twitter’s developer guidelines and policies. These policies enforce certain constraints such as “rate limits,” the use of abusive or
hateful language, a warning against reverse engineering the Twitter API, or use or manipulation geographic data.

**BOTNET LIFECYCLE**

How do botnets work? The vast majority of malicious botnets echo tactics like that of the PeñaBots — they attempt to shape policy and discourse, spread fake news, target users for harassment, and poison hashtags — and follow one of two paths (see Figure 16). First, a “farm” builds fake accounts, which, for creating Twitter accounts, means following several steps: creating a username and profile, uploading an image or avatar, verifying information through an email service, and providing a phone number (which can be generated via a third-party application). Then the farm sells each account to a disinformation campaign. Carlos Merlo, the previously mentioned kingmaker of the Mexican Internet, said he purchases accounts for fifty cents apiece (Broderick & Arredondo, 2018). At this point the disinformation campaign uses a number of different third-party applications to disseminate and coordinate botnet attacks. Broderick and Arredondo reported that Merlo uses Microsoft Excel to push narratives through his botnet—usually a message or hashtag, or series of hashtags, but worded differently in separate accounts and tweets to avoid complete repetition. The second method is when a botfarm builds the accounts already linked together on Twitter and “seeds” them over time to build trust, users, and connections. These botnets are more valuable and are available to be sold as a unit, in bulk. The Star Wars botnet was one possible bundle of bots that could be sold for a high price to a bidder, most likely through bitcoin (Echeverria & Zhou, 2017).
AGENCY AND BOTS VS. BOTNETS

It seems that Twitter botnets are built of large groups of disposable accounts used for a short time before being detected and removed, suspended, or deleted by Twitter. As Erin Gallagher noted in her talk (2015), the hashtag poisoning bots pushed specific hashtags in a method that openly broadcast their automation, employing Twitter’s algorithm to remove the trending algorithm. These bots were designed to be discarded. They are low-level peddlers of information, meant to overwhelm a system with information in a coordinated fashion. Some bots, such as @ImmoralReport (Figure 17), however, are individually as powerful from the standpoint
of rhetorical agency as a large group of coordinated bots, and these accounts differ in apparent use of automation, which, in and of itself, can be difficult to track.

Twitter’s bot-detection algorithms are proprietary, and little information has been released to the public, other than headlines that Twitter deleted massive amounts of accounts for exhibiting a particular type of behavior: too many tweets, tweets that are abusive or threatening, accounts that haven’t been used for a certain time, etc. Users, researchers, and activists have attempted to build programs and web tools to help others find and report bots. But these programs often conflict with one another—one program might be able to detect a high probability that other programs might not due to different search parameters.

One example that demonstrates the difficulty in tracking bots and automation is the user @ImmoralReport, which has been repeatedly noted by Twitter users as an automated account that pushes controversial topics and hashtags while continuously violating Twitter’s Terms of Service. The account operates at all hours of the day, pushing large amounts of information (over one million tweets) since July 2015, and is currently operating as of June 10, 2019 (Figure 17). It’s possible that @ImmoralReport tweets just enough to stay below an invisible tweet threshold that the Twitter bot-detection algorithm uses as a baseline. The account appears to be designed to push as much content, as often as possible, to as many users as possible (it follows over 30,600 accounts, and it is followed by over 31,300 accounts).
Based upon different detection methods, it is difficult to prove that @ImmoralReport is either a bot or an automated account. Self-proclaimed data scientist @conspirator0 and colleague @ZellaQuixote – anonymous activists who spend time detecting and reporting trolls and bots on Twitter – offer a web tool for bot detection at http://makeadverbsgreatagain.org/allegedly. The tool offers a simple automation analysis of Twitter users and was accurate when I tested it against my own account and other known accounts of fellow doctoral students. When tested
against longstanding, suspected bot accounts, the tool helps illuminate behaviors that suggest automation. @ImmoralReport, for example, is revealed to tweet at all hours of the day and often at least 35-55 times an hour (see Figure 18). The web tool notes that a third-party application tweets on behalf of “StopMadness2,” and is active every hour of every day. Although it is unclear what “StopMadness2” actually is, we can infer that it is a form of publishing tool meant to tweet and retweet content specific to @ImmoralReport’s agenda. But what about other third-party apps that are used commercially?

George Tech University researcher and instructor Charles Grimm noted that he uses dlvr.it – a publishing and marketing automation service – to schedule and automate some of his tweets (Personal correspondence, 2019). The same web tool used for generating a data report for @ImmoralReport detected that some of Grimm’s posts (10.8%) were published via dlvr.it, and his posts were color-coded by the web tool to the time of day in which they were published. He regularly scheduled posts for 8am, 10am, or 3pm (Pacific Standard Time), coinciding with high web traffic times (see Figure 19). Also of note on Grimm’s data report is that there is a clear indication of when he is inactive or asleep, his tweet volume is barely more than five tweets per hour at most, and his primary method of tweeting is directly through the Twitter Web App. By most metrics, Grimm’s account would appear to exhibit very normal human behavior.
Figure 18: Web tool analysis of @ImmoralReport’s tweets, tweet volume and schedule, and use of apps/services, created at http://makeadverbsgreatagain.org/allegedly.

Figure 19: Web tool analysis of @GrimmProspects’s tweets, tweet volume and schedule, and use of apps/services, created at http://makeadverbsgreatagain.com/allegedly.
Other bot detection services, however, do not easily identify @ImmoralReport as a bot. The Botometer, for example, doesn’t confidently characterize @ImmoralReport as a bot.

Botometer, a program and research collaboration between the Center for Complex Networks and Systems Research (CNetS) and Indiana University Network Science Institute (IUNI), uses “a machine learning algorithm trained to classify an account as bot or human based on tens of thousands of labeled examples” (botometer.iuni.iu.edu, 2019, n.p.). According to the Botometer FAQ, the researchers and programmers made a complete automation probability (CAP) score, which gives the probability of an account being a bot, based on automation.

This probability calculation uses Bayes’ theorem to take into account an estimate of the overall prevalence of bots, so as to balance false positives with false negatives. For example an account might have a relatively high bot score of 4, but only a 30% probability of being completely automated. CAP may look like a conservative estimate, but one must consider that there are a lot more humans than bots, so you’d better be sure before accusing someone of being a bot!

The Botometer delivers a moderate probability score of bot behavior (see Figure 20), ranking @ImmoralReport as 2.3 out of 5. More interestingly, the Botometer’s CAP score is listed at only 8%, but the language-independent features are flagged as problematic. Its English-specific features are more moderate to normal, with content scores at 2 out of 5, and sentiment scores at 2.5 out of 5. Using the Botometer’s algorithm, based off of peer-reviewed research and publications, it’s easier to see why Twitter has not suspended the account based on automation thresholds and conflicting data on violations of their Terms of Service. Individual categories may suggest a reasoning for suspension, but holistically, @ImmoralReport looks to be just skirting the rules.
One additional aspect of the Botometer is the ability to examine the account’s friends and followers to determine a probability of how many bot accounts are networked with the account. @ImmoralReport also has a lot of connections to apparent bot accounts. Why does this matter?

Bots need to connect to human users in different ways in order to shape opinion, alter discourse, and influence ideas and policies.

Bot Sentinel, which touts itself as a nonpartisan bot-tracking website and Chrome extension, claims to use “machine learning and artificial intelligence to classify Twitter accounts who exhibit trollbot behavior … then add accounts that exceed a certain threshold to a publicly available database than anyone can browse” (botsentinel.com, 2019, n.p.). Bot Sentinel, according to its online FAQ, defines a trollbot as “a classification [they] created to describe human controlled accounts who exhibit toxic troll-like behavior. Some of these accounts frequently retweet known propaganda and fake news accounts, and they engage in repetitive bot-like activity” (n.p., 2019). Bot Sentinel is careful to cite that researchers disagree on the pragmatics of language use and situations and say that their machine-learning algorithm searches for accounts that break the rules of Twitter. Bot Sentinel administrators do not publish their code, nor do they offer concrete insight to their parameters—only global guidelines:
Instead of creating a model based on our interpretation of a troll or bot, we used Twitter rules as a guide when selecting Twitter accounts to train our model. We searched for accounts that were repeatedly violating Twitter rules and we trained our model to identify accounts similar to the accounts we identified as “trollbots.” Note: Ideology, political affiliation, religious beliefs, geographic location, or frequency of tweets are not factors when determining the classification of a Twitter account. (n.p.)

They also define their own differences between “propaganda bots,” “propaganda trolls,” and “trollbots.” Although their language is somewhat confusing, they infer that “bots” are algorithm-driven and they either work alone or in groups. Propaganda bots, for example, push political agendas, whereas trollbots work to divide opinions through toxic behavior. Linguistically, however, these two classifications shouldn’t be exclusive of one another.

FROM BOT TO BOTNETS

Botnets, however, have the added strategic and rhetorical value of volume, mostly consistent with coordinated tweets that are similar in style, content, or purpose. Automation, particularly voluminous automation, can disconnect the audience from performative aspects of agency. When one account tweets something controversial, perhaps by posting something false, the “collective good” might offer a series of rebuttals or reply with factual information. When an algorithm-driven botnet of 1000 accounts tweets the same thing in different ways, the collective good is overwhelmed and can’t possibly offer rebuttals. The false posts sit unchallenged. Any neutral observers might notice this information is being omnipresent and unchallenged and are thusly persuaded by the misinformation.

Bots, conversely, offer little in the way of nuanced rhetorical strategies, which is their weakness. A botnet, focused on a particular outcome, cannot individually respond to individual challenges well. Its performative agency is flat and often recognizable to someone who has
participated in Twitter for some time. But what bots lack in performative energy, they make up for with a continuous assault of repetition and volume. Put together as part of a cohesive message, they can amplify or obfuscate messaging as long as they work within constraints and parameters set forth within a network. Twitter has shown that it levies algorithmic constraints, but it does not allow the public to see or know the actual metrics and limitations of those constraints. We have to consider these constraints and strategies in this new age of digital agency in the realm of bots and botnets if we are ever to combat the weaponization of automation.

Please allow me to illustrate an example of how performative, rhetorical strategies are adjusted to consider constraints. A speaker might be told they will be speaking for 60 minutes to an audience of ten in a small room on one day, then for 10 minutes to a theater of 5,000 people. The strategy of the speech will shift performatively. Likewise, a writer will shift performative strategies based upon their medium and audience. And just as a speaker or writer shifts their strategies around constraints, a programmer must code their bots and algorithms according to the platform’s constraints.

The Peñabots, in this framework, are programmed to operate in volume, but within the algorithmic constraints of Twitter’s platform. A violation of the ToS offers the chance that Twitter will discover a bot, or even the connected botnet, and remove it. Twitter bot programmers must be careful about how many times a bot can tweet or reply; what type of content it is amplifying, obscuring, or countering; what time of days it operates (is it functioning 24 hours a day, unlike a human being?); and keep the content from being overtly offensive or controversial.

Continuing to use the performative analogy, botnets are not that dissimilar to a speaker talking in front of a camera or microphone or a writer updating a blog, complete with hyperlinks.
The messaging is replicated and disseminated to the masses, using technologies that augment the automation of the messaging. Control over this automation is a subject unto itself, but important nonetheless to think of when considering hegemony and control over who gives the messaging and to whom the messages will be given. I previously mentioned Mexican “Internet kingmaker” Carlos Merlo, who willingly demonstrated his control over trending topics and discourse on Twitter. He was part of a power structure of politicians and influencers that wished to alter, amplify, and obfuscate messaging online.

**ACTANT CHARACTERISTICS**

- **Erin Gallagher** – Researcher, multimedia artist, and translator. She produced a lecture that identified Peñabot tactics, experiences, and interactions.

- **Alberto Escorcia** – Journalist and researcher for Buzzfeed News and Lo Que Sigue. Escorcia uses data visualization platforms to track bots and their networks as they attempt to influence Mexican life, politics, and information.

- **Bot Sentinel** – Self-described as a nonpartisan bot-tracking website. Bot Sentinel also exists as a Twitter bot that publishes information about automated accounts collected through the website.

- **Botometer** – Developed by researchers through CNetS and IUNI, it provides a CAP score that helps to inform human social media users of the probability a certain Twitter account is, indeed, a bot. This information is based upon such factors as language, followers, and automation.
• Allegedly Web Tool – Used for tracking automation and third-party apps of Twitter accounts, this web tool is unverified but appears to be accurate when tracking tweets, tweet times, volume, and applications used for tweeting. Activists and self-described data scientists @conpirador0 and @ZellaQuixote claim to use this tool to provide real-time feedback about automated accounts to the public.

• Twitter APIs – API is short for application program interfaces, ways for programs to talk to one another and exchange information and requests. Twitter uses five types of endpoints within this framework: tweets and replies, direct messages, accounts and users, ads, and publisher tools and software development kits (SDKs).

• Human Twitter Users – Users on Twitter number in the hundreds of millions, according to Twitter, which identified “active monthly users” in February of 2019. These users form networks with one another, interact with hashtags, and exchange information in a dynamic ecosystem.

ONTOLOGIES

The main feature of the Peñabots I’d like to note is how the bots are created as disposable, rhetorical tools to poison hashtags, game the Twitter “trends” algorithm/novelty index, and drown out the discussion between people and groups of people by creating a sea of white noise. Figure 21 demonstrates how Twitter users employ a hashtag in order to break into a sphere of reciprocating media and information. As Twitter users begin to build momentum by using a particular hashtag, such as the aforementioned example of #sobrinaEPN (bringing to light the nepotism of Enrique Peña Nieto as he installed his niece into a position of great power in Mexico), activists were attempting to recreate an uproar against what they perceived to be
nepotism and corruption. By pushing the hashtag into the trending topics and hashtags feature of Twitter, activists stood to reap powerful rewards, i.e., media recognition from a variety of sources. Potentially, their collective voices could be circulated among both local and national media outlets. They also could raise awareness to the regional/national Twitter trending topics bar, meaning that everyone within Mexico’s Twittersphere would at least see the trending topic. Additionally, their hashtag would then appear in powerful and popular search engines such as Google and Bing, further raising awareness.

But as the users began to trend the hashtag on Twitter (see Figure 21), the Peñabot network began to echo the #sobrinaEPN hashtag by overwhelming the index with automated responses, as I delineated earlier (Figure 22). The simple act of amplifying the hashtag to an extreme level of retweets and tweets trips the unknown variables of Twitter’s automation-detection algorithms, effectively suppressing the hashtag and keeping it from trending. This term is called hashtag poisoning—a term I credit to Alberto Escorcia and Erin Gallagher to describe a deliberate rhetorical strategy of seeing a topic trend through the hashtag feature on Twitter and disrupting the trend by flooding Twitter with a massive repetition of the hashtag. The “poisoning” trips Twitter’s automation-detection algorithms, which then shut down the trending hashtag. It’s a counter-strategic move designed to end momentum before it grows exponentially. The Twitter automation-detection algorithm, in theory, helps to keep automated content from trending, and to keep organic, human content and discourse relevant. Otherwise, bots could generate all content and eliminate collective, sociological movements on Twitter.
Figure 21: Trending hashtag ecosystem on Twitter.
Figure 22: How a botnet can effectively "poison" a trending hashtag.
ETHICAL CONSIDERATIONS

It appears that in the digital age we have superseded moral obligation, since immoral actions are now rewarded through emotional and monetary changes in currency, not through punishment or rebuke. Twitter users, for example, learned early on that outrage generates online traffic, or clicks. These clicks amount to “retweets,” “likes,” and “follows” on Twitter. The more digital traffic that is generated, the more influence a Twitter user has on an audience. This translates to a number of capitalizing enterprises: advertising revenue, votes, commerce, and agency, for example. As a result, emergent, extremist American voices on social media learned they no longer needed to adhere to the traditional rhetorical notions of ethos, pathos, or logos, but they could rely on rhetorical strategies that sow discord and disunity. They play the part of provocateur, using devices such as feigned outrage and hyperbole, for example. The more discussion a tweet generates, the more a user’s position rises in indexing algorithms. A trending topic involving a controversial tweet may cause a large group of Twitter users to be offended and angered, but the return currency is that the offending Twitter user will gain followers—not just supporters, but also those interested to see what might be said next. In short, offending tweets can be just as lucrative as those that are uplifting or inspiring. When advertising dollars or other forms of capital are tied to influence and hegemony, those who generate controversial content might earn more than someone who thrives on positivity and creativity. In short, it pays to kick the hornet’s nest.

Because of this shift in rhetorical strategy, where Twitter users commonly get attention by swearing at politicians and celebrities and chide reputable figures with offensive words and images, botnets gain a new advantage in agency. Not only does automation have advantages over
human output and interaction, but negative rhetorical strategies that are often employed by the “outrage machine” have been adopted by trolls, trollbots, and propaganda bots. These bots generate and circulate content that mimics negative discourse that is often indistinguishable from that of real Twitter users who are attempting to capitalize on negative interactions. Some Twitter bots, simply through the act of using negative strategies to divide and taunt, can attract thousands, if not hundreds of thousands, of Twitter users and other bots. The @immoralreport is, possibly, one such example. And, as the @immoralreport gains followers, the ability to reach and influence others grows at an exponential rate.

Smear campaigns that the Peñabots and other Mexican Twitter bots employed during the presentational campaigns of 2012 and 2016 receive a twofold advantage: first through volume and automation and second through gaining users through partisan tactics. Users and bots that employ partisan tactics and explicit, political propaganda receive the added benefit of being readily recognized by members of a particular political base, who might feel compelled to “follow” the account. More than likely, other followers will share similar interests. The partisan accounts become an attractive meeting place for others to connect and commiserate, not all that unlike we saw with the [loveMachine] and how it connected gay men to one another through Facebook.

Bot automation also employs disinformation as a tool to combat rationalism. In the arena of physical warfare and combat, guerilla-style warfare is highly successful pitting small, mobile groups of combatants against large, traditional armies. Guerillas employ unconventional warfare, using every advantage possible—terrain, weather, spying, assassination, sabotage, propaganda, destruction of supplies, etc. In the digital age, the smaller, outnumbered forces – such as the self-described “alt right” movement – are engaging in their own type of unconventional warfare.
Their types of information warfare and combat include upending the validity of truth, facts, and data by overrunning other users with false data until it becomes difficult to validate actual truths. Other forms of digital combat include bots and botnets, employed to sway voters, divide intellectuals, and create an illusion that some issues are much larger than they are in reality.

Morality is not a driving force in guerilla warfare—victory is. We arrive at certain truths, in part, by comparing them to other truths. But when bots and botnets disseminate untruths, false news, propaganda, skewed data, and other means to combat logic and logical processes, we become mired in a battle for the very foundation of truth.

The Peña-bots, and other botnets, also disrupt our ability to process information by overloading our cognitive perception. And we aren’t the only ones who suffer from this overload. Algorithms that look for patterns through quantity are susceptible to manipulation.

Collin Gifford Brooke (2009) writes that in the digital age we might not respond to the issue of an overabundance of data through cognitive storage, but with persistence. As issues develop, we can respond by adapting with new solutions. Brooke uses the example of weblogs, which, at the time of his 2009 text, had grown exponentially over the previous five years. Brooke said that he followed over 100 sites daily and found it difficult to keep track of updates and changes in weblogs. The Internet community responded to this immense amount of data by developing Really Simple Syndication (RSS) feeds for readers and followers of weblogs (and news headlines). Rather than rely upon his own memory—a limited form of storage—Brooke turned over the automation to RSS.

The feeds from many blogs can be collected by feed readers or aggregators, and there are both desktop and web-based versions of these programs … Effectively, aggregators check weblogs and keep track of whether a particular user has accessed the most recent content. They check our blogs so that we don’t have to, in the same way that most mail programs can be set to inform a user when there is new mail in the inbox. (p. 158)
The simple syndicating algorithm used automation to parse the overwhelming quantities of data so that Brooke could make informed choices without wasting time or relying upon his memory or other cues.

But Twitter botnets function under the protective umbrella of the very company that allows them to broadcast content. The soft and somewhat ambiguous rules of Twitter leave room for interpretation. As we saw with Milo Yiannopoulos, it took several significant violations of the Terms of Service to ban him permanently from Twitter. Several other prominent right-wing celebrity trolls constantly escape suspension, even though unknown Twitter users are suspended for violating Twitter’s ToS for much less severe infractions. And, as exemplified by Twitter activists such as @conspirator0 and the researchers behind the Botometer and Bot Sentinel, Twitter doesn’t necessarily suspend automated accounts committing policy and term violations, even when the activists lay the accounts at Twitter’s doorstep in an act of good faith—they seem to be the community watchdogs, if you will.

Whereas the general Internet population both created hundreds of thousands of weblogs and learned how to syndicate them, the Twitter public is subjected to the ecosystem of Twitter’s platform. Through the model of Brooke’s (2009) persistence, I noted the methods in which public citizens attempted to circumvent the problem of botnets, trollbots, and spambots. Bot Sentinel, Botometer, and user @conspirator0, for example, interact in different ways with Twitter’s API and feed as actants in the Twittersphere, but they are not directly connected with policy. Their influences are indirect – suggestions, really – as they attempt to sway the public, as well as Twitter employees, with accounts that habitually break or shirk the rules set forth by Twitter itself.
Twitter is not willing to explicitly collaborate with data collectors, even when faced with overwhelming evidence of violations of Twitter’s Terms of Service. Popular figures regularly flaunt these rules, which can be best exemplified by extremist tweets from politicians, celebrities, and law enforcement members. Twitter has found itself reluctant to ban these figures, presumably because the company relies upon its users to generate revenue through advertising. If Twitter were to ban Donald Trump for inciting violence against the press or individuals, for example, the backlash would be severe. His account is followed by over 60 million Twitter accounts. Although some data analytics companies or research centers, such as the Pew Research Center and SparkToro have noted that @realDonaldTrump’s account is comprised of over 60% fake accounts, that still leaves millions of users who might leave the platform or view it negatively. This affects Twitter’s bottom line.

Rhetorical agency on Twitter, it can be surmised, is affected by capitalism. In exchange for marketing revenue through advertising, Twitter plays a balancing game of placating users while softening its own rules that allow botnets to form and interconnect with extremist voices. Bots, after all, do count as users, and Twitter is able to falsely claim a larger viewership and community to advertising companies, the public, and investors. This comes at a cost to users. A Twitter user might use the service expecting to interact with advertisements as a form of social currency exchange, but there is a hidden cost of surrendering agency to bots and botnets. Bot accounts, therefore, are forms of theft—they inflate false values, steal productivity, affect time spent, and minimize rhetorical agency. Twitter employs a unit of measurement called the mDAU (monetizable Daily Active Users) to account for revenue; they claim this has increased by 11%.

---

10 Quartz, October 12, 2018, by Ana Campoy; SparkToro, October 9, 2018, by Rand Fishkin; Pew Research Center, April 19, 2018, by John Gramlich.
each year, even though their monthly average users (MAU) has decreased in 2019 (Twitter, 2019). Since its numbers have declined, Twitter has been more selective about what forms of data and analytics it releases to the public.

After Russia’s election interference began to unfold, lawmakers and the public turned a more scrutinizing eye to social media companies like Facebook and Twitter. In response, Twitter launched an information and suspension campaign, promising to tighten restrictions and controls. In May of 2018, Twitter claimed to delete 9.9 million accounts per week. It stated that they were responding to user-driven reports of abuse, instead of adapting a more proactive approach to identifying problematic accounts. “We focus on developing machine learning tools that identify and take action on networks of spammy or automated accounts automatically. This lets us tackle attempts to manipulate conversations on Twitter at scale, across languages and time zones, without relying on reactive reports” (Roth & Harvey, 2018, n.p.). Roth and Harvey also wrote that Twitter’s effectiveness in removing problematic accounts was a result of technology and process advancements. However, regular citizens were adapting to problematic bots and behaviors long before Twitter claimed to have made improvements in technology. For example, during the same month Roth and Harvey claimed Twitter had increased the deactivation of violating accounts by 214% – May 2018 – the researchers at Botometer merely updated their data collection points to make their program more effective. Although Twitter’s bot and spam identification algorithms theoretically operate at a higher level of sophistication and complexity, they have either been consistently outmatched and outpaced by public citizens determined to assist other Twitter users, or Twitter has been purposefully ineffective in their malingering, opting not to purge, suspend, or remove bots, spam apps, and suspicious accounts.
CHAPTER 4: WIKIPEDIA ADDBOT: PLANNING AND GOVERNANCE VS. EXPRESSIVENESS

EARLY OBSERVATIONS

Marilyn Cooper (1986) is generally credited with coming up with the notion of "ecologies" within the field of composition and writing. As writers connect with one another, they make and remake ecologies in response to situations. Later, in his seminal work on digital rhetoric, Doug Eyman (2015) drew on several researchers\(^1\) as he used the scientific term "ecology" as "a metaphor for complex, interconnected relationships" within the field of writing studies (p.84). Eyman continued to say that "the key elements of ecological study – relationships, interaction, complexity, and community – easily map onto qualitative studies of writing and rhetoric in both epistemological and ontological terms" (p. 85). In the era of digital networks and overlapping systems, I posit that bots exist within ecosystems, and those ecosystems exist within digital ecologies. Citing Kling, Eyman says that there are two important properties of ecosystems: "they have energy flows and they cycle materials" (p. 86; emphasis Eyman’s).

Continuing the metaphor, Eyman describes the energy flow of a natural ecosystem as it is "enacted through the uptake, use, respiration, reformation, and reuse of the basic ecological components (e.g., plants, animals, water, carbon, nitrogen); the energy flows provide the engine for these material cycles through input and consumption (of solar/heat energy)” (p. 86). If we

view digital spaces as ecologies, we should also take into account energy flows and the cycling of materials. Bots, existing within ecosystems – digital ecosystems – act as specialized components that take rhetorical agency – the kinetic energy of a call to action – and distribute it among agents and actants within that ecosystem.

By using the term *ecosystem*, I am referencing a self-contained system defined by an information ecology framework: "a system of people practices, values, and technologies in a particular local environment. In information ecologies, the spotlight is not on technology, but on human activities that are served by technology” (Nardi & O’Day, 1999, p. 49). This type of limited system “shifts focus from technology as to tool to technology-in-use (that is, activity can be seen as a synergistic relationship between digital media/technologies and human actors) and it focuses the lens of inquiry on a finite context (which is useful for the development of research methods)” (Eyman, 2015, p. 88). Eyman, by way of Spinuzzi and Zachary, argues that a genre ecology better serves dynamism, rather than static forms, and the “interconnectedness of genres” (p. 173). Eyman also argues that circulation ecologies better serve web-based text for the purposes of accruing “cultural and social capital” (p. 92). However, since I am attempting to look at a static place in time (i.e., Wikipedia in 2013), Nardi and O’Day’s information ecology framework serves my purposes, as well as helps me to define a limited space and time for governance-based algorithms and ontologies.

Digital ecosystems within digital ecologies can reflect planning and governance practices to manage the transfer of resources and the cycling of materials, or a lack thereof. In a natural ecosystem, specialized agents assist with the transfer and cycling of materials, such as a forestry manager or conservation warden. These specialists exist because of training, tax dollars, and a community willing to employ them to assist with natural resources. Within digital ecosystems,
we often employ stewards of information, automation managers, policy makers, etc., in order to increase efficiency and safety among digital communities.

Bots, capable of dramatically altering the distribution of energy and the cycling of materials, are sometimes carefully managed. These types of bots are usually functional and tool-like in scope, fulfilling “a range of routine, utilitarian functions aimed at helping users of social media, principally by finding, analyzing, processing and transmitting data” (Bollmer & Rodley, 2017, p. 158). When several of these types of bots in an ecosystem are created for specific, different purposes, they may or may not interact well with one another in a given space. Each bot’s task or series of tasks may interfere with another bot’s mission, not all that unlike a single program that has several conflicting lines of code that need to be debugged. Here’s a fictional example: A bot programmer may create a bot to perform the task of correcting American English grammar and spelling in a series of online Wiki articles, and a different bot programmer might create another grammar/spelling bot for the same articles, but to correct to British English. Both programmers are attempting to automate a task for the better good of the community, but both bots end up creating a *bot conflict*, where the automated processes come into direct conflict with one another. The American English bot changes the misspelled word “colour” to “color,” and the British English bot changes “color” to “colour.” Then the first bot reverts the second change back to “color,” which is then changed back to “colour,” and so forth. The confusion creates more harm than good, leading to community confusion about the correct spelling of a word. Now imagine a third, fictional bot in the same ecosystem, designed to notify the article authors of any changes, so it continually emails notifications to the authors in a never-ending cycle. The end result is suboptimal, and all three programmers need to find a way to decide which bot to use and the parameters surrounding each bot—a system of *bot governance*. 
Bot governance is a system of policies and a method of employing and maintaining those policies over simple “bot” algorithms in an effort to control factors such as 1) changes in technology; 2) overlaps in information and technological systems; 3) changes in bot ownership or management; 4) to protect shifting security and privacy concerns; and 5) increase or maintain efficiency between humans, other bots, and digital spaces. I developed this definition after reviewing one dozen different companies that function within academia, healthcare, finance, social media, and information systems. This definition is different than robotic process automation (RPA), which includes cognitive, behavioral, and artificial intelligence bots and the management thereof. When reviewing bot policies, I collected as many public policies as possible from companies that produced their content online. I began with academic institutions, then branched out to connected institutions either through navigating online to connected links, through recommendations by scholars and researchers, or through crowdsourcing ideas from users on LinkedIn and Twitter. Once I had collected and reviewed at least ten policies, I looked for correlations between the two to create the definition.

Human responses to bot conflicts is an area worth studying, particularly when considering rhetorical agency. Bots have removed some cumbersome and complex social interactions between people; they have replaced tiresome, repetitive, and/or tedious processes with automation, and they allow digital ecosystems and spaces to expand by freeing up time and labor. However, from a perspective of rhetorical agency, these automated processes can cause

\[2\] The policies and practices I reviewed came from the following businesses, groups, and reports: IBM, Imaginea, Deloitte, Botstatz, Genpact, CIO, KPMG, Everest Group, University of Illinois Grainger College of Engineering, McKinsey Digital, and Wikimedia.
serious issues and errors when bots come into conflict with one another within the same ecosystem.

Consider a hospital that wishes to use bots to free up time and labor (and therefore capital) to automate medical billing and coding, the complex system of assigning codes to medical procedures, processes, and equipment for insurance and billing purposes. Bot conflict can theoretically cause dire consequences for patients and consumers, such as erroneously denying coverage for a life-saving procedure or operation. Like managing humans, the hospital would need to have a system of bot governance in place. Once bots are created and executed within an ecosystem, they begin to lose their identity and origin. Without governance, an organization’s bot, and therefore, programmer, loses accountability (Gorthi, Bajaj, & Relan, 2018).

To balance these various types of algorithms I sought a case study of a controversial algorithm that was formed under strict guidelines and controls—one that was created under the rules of bot governance. I followed Addbot, a closely monitored and governed Wikipedia bot designed to improve functionality but which made edits and changes at such a large scale that it gave the appearance of creating conflict and infighting within the Wikipedia ecosystem.

PROGRAM DESCRIPTION

In March of 2013, a bot named Addbot reverted 146,614 contributions other bots had made to English Wikipedia. The bot was designed to remove the old style of “interlanguage links” in order to pave the way for a new way of capturing the cross-language relationships between articles in Wikidata. For example, the article for Robot on English Wikipedia is linked to the article Roboter on German Wikipedia. Before the decision was made to move these links to Wikidata, dozens of other bots were used to keep the link graph up to date. But after the move to Wikidata’s central repository, this automated work and the links they created were no longer necessary. So Addbot removed
all traces of interlanguage links from all 293 Wikipedia languages and paved the way for maintaining them in Wikidata. (Halfaker & Geiger, 2017, pp.1-2)

In more simple terms, the Addbot changed, or reverted, almost 150,000 edits made by other humans and bots. But to hear some researchers and media sources say it, the Addbot created a “botpocalypse”—a term used to describe a bot or network of bots that make grand, sweeping changes, usually by undoing the work of other bots or people or suppressing other bots’ automated functions. The very notion of a botpocalypse suggests chaos and disorder, but did the Addbot really contribute to an atmosphere of disorder?

CONFLICT AND AGENCY

Bot-on-bot conflict or violence, to use common terms in the media, are labels in part because, according to Matt Simon’s Wired article (2017), “Internet bots fight each other because they’re all too human”; bots take on human characteristics. “The bot and its creator are, in an intimate sense, a hybrid organism.” Carolyn Miller (2007) refers to the process of attribution to these computer-mediated discourse as ethopoeia, or the “construction of character in discourse (sometimes called ‘impersonation’).” Furthermore, she says that research in computer-mediated communication suggests that “we have a very low threshold for ethopoeia; in other words, it doesn’t take much for us to attribute character to an interlocuter, no matter how primitive the cues are” (p. 151). Miller refers to the “Eliza effect” – the same chatbot therapist from MIT that Bogost (2007) used as an example for his explanation of procedural rhetoric – as she discusses our willingness to attach human qualities to these types of programs. After all, Miller says, we name our cars and attribute luck to gambling machines. As for Simon’s assertion of human-algorithm hybridity, he suggests a close connection between programmer and program,
something Wikimedia fosters in its closely curated collection and governance of bots. Once a bot is no longer actively connected to a programmer through a database of accountability or open source repository – yet still functions at will in a given ecosystem – does its hybridity disappear? Does its expressive nature, and therefore agency and attribution to human characteristics, change when disconnected from the programmer? These are additional questions I decided to explore once I learned more about not just the Addbot, but the Wikipedia ecosystem in which it was born and lived.

Addbot was coded in PHP, a hypertext preprocessor originally created for web development. Addbot was created for the purpose of removing interwiki language links, but only “after a massive effort coordinated through Wikidata, a meta-level database created to centralize language-independent data” (Geiger & Halfaker, 2017). Prior to its execution, programmer Adam Shorland had to obtain bot access for his program.³ As part of Wikimedia’s governance rules, the word “bot” must be in the name of the actual program, which led to the name choice that notifies everyone in contact with Addbot that it is an autonomous program. Between March 5 and March 25, 2013, Addbot executed its first series of changes, or reverts—146,614 of them (Geiger & Halfaker, 2017). Since then, and up until September 23, 2016, the Addbot continued to make large numbers of reverts under several different iterations, the latest (as of April 19,

³ According to Wikimedia’s “bot policy” page, “there are three ways to obtain bot access on a wiki.” 1. A global bot “must only maintain interlanguage links or fix double-redirects” or already exist on several wikis with a proven track record of success and reliability. 2. Automatic approval can be issued for bots “permitted by the local bot policy or request page,” but the bot must demonstrate itself without a bot flag by editing regularly for a week or making 100 edits on “applicable wikis.” 3. Community consensus can be granted if there is local community interest and approval via a relevant discussion forum. A steward can be requested to give a bot flag (approval) or “a bureaucrat will add the flag.”
https://meta.wikimedia.org/wiki/Bot_policy
There are many listed variations of Addbot and the tasks each one performs, denoted by the version number after the name. Addbot 1 “creates links it finds while trawling Wikipedia articles”; Addbot 3 “mass-ad[s] properties based on Wikipedia categories”; Addbot 4 “Import[s] Geo Coordinates”; and Addbot 5 “Mark[s] dates that need checking for calendar model correctness” (n.p.). This list is only the first few Addbots, and by the date of this writing there were at least 19 iterations of Addbot.

Geiger and Halfaker (2017) identify several types of reverts, in which one user account removes another user account’s edits to a page. This is typically defined as bringing it back to an exact previous state, called an identity revert. Other definitions include partial reverts, where part of an edit is removed, or declarative reverts, which is when a user leaves a note in an edit summary stating that they are reverting the edit. (p. 4)

Reverts are an interesting by-product of online discourse on Wikipedia, as users or bots exert authority by changing a revised text back to the original. There have been many examples of conflict over the past decade as some Wikipedia editors fiercely “battle” with one another making edits and reverting comments back to original edits, each editor attempting to create a narrative for a number of subjects ranging from the mundane to the bitterly contested (such as the great frustration over what and how to name the object Pluto after it was re-designated: a minor planet? A dwarf planet? 134340 Pluto?).

Addbot’s automating of reverts was, as mentioned at the beginning of this chapter, a large but mostly benign task as it removed interlanguage links between 293 different Wikipedia languages, and it received only a small amount of concern within Wikipedia’s chat areas. Adam Shorland said, “You can still see some of the pushback by people that were still unaware what was happening on the Addbot talk page and in the archive” (Personal correspondence, 2019). But overall, the deployment of the Addbot was relatively quiet, considering the volume of changes it
made over the next month. So why are people calling Addbot, and other automating bots on Wikipedia, conflict provoking?

Geiger and Halfaker (2017) argue that quantitative methods used to measure bot conflict are leading to misinformed results as to the nature of bots and bot conflict. The paper, “Even Good Bots Fight,” by Tsvetkova et al. (2017), determined bot conflict through quantitative data analysis. However, Geiger and Halfaker say that the use of quantitative data to determine bot conflict is done without the “broader socio-technical context of what those traces mean in the Wikipedian community” (p. 2). “Even Good Bots Fight” received significant media attention because it suggested that Internet systems such as Wikipedia were not able to regulate or rein in runaway automated agents, or bots, who were in constant conflict with one another as they automated reverts back and forth. In their findings, Tsvetkova et al. say that “although in quantitatively different ways, bots on Wikipedia behave and interact as unpredictably and as inefficiently as the humans. The disagreements likely arise from the bottom-up organization of the community, whereby human editors individually create and run bots, without a formal mechanism for coordination with other bot owners” (p. 7).

Geiger and Halfaker (2017) suggest that, even at the surface, there are different types of conflict that aren’t distinguished by qualitative measurements from counting reverts and other similar data analyzed in “Even Good Bots Fight.” Geiger and Halfaker give examples:

… if one bot developer believed that images without a properly-formatted fair-use rationale ought to be removed from articles, and the other bot developer believed that such images ought to remain in articles, this would be a task conflict. However, if both bot developers believed that such images ought to be removed from articles but had different ideas about what a properly-formatted fair use rationale looks like, this would be a process conflict. And if a bot developer got angry at another bot developer and wrote a bot to undo all of their previous edits, this would be a relationship conflict. These types of conflicts would all present as reverts. (p. 4)
Determining what bot conflict is and means is important when examining the distribution of rhetorical agency among social media bots, such as the Wikipedia Addbot. There are nuances when considering qualitative ways of looking at conflict. The differences in conflict are properties and hallmarks of the distribution of agency as part of rhetorical strategies, which cannot be easily quantified.

Task conflict vs. process conflict vs. relationship conflict are all different aspects of rhetorical strategy. Using Geiger and Halfaker’s (2017) example of differing conflicts, we can see that a task conflict might result in, ideally, *dissoi logoi*—two competing, logical arguments that end with the better argument achieving a higher perspective. Process conflicts might achieve best results through collaboration. Relationship conflict might need intervention. The conflicts are what lies underneath all of the automation and bot designs, when two programmers create bots for specific purposes, the bots are executed to *resolve an issue*, such as within the ecosystem of Wikipedia. Shorland’s Addbot was created in good faith to revert interlanguage links, undoing what other bots might have done before. Quantitatively, each revert would be a conflict. Addbot, however, was providing a holistically and intrinsically good service, saving many editors hundreds, if not thousands, of hours of labor while achieving higher efficiency and harmony within Wikipedia articles and languages.

**KINETIC VS. POTENTIAL AGENCY AND ENERGY**

Kinetic and potential agency are important aspects of bots and agency; there are important matters of potentiality and rhetorical agency before the bot is ever created. If we are to continue to use kinetic energy as a metaphor for agency, we should also note that potential energy is equally as important. In physics, kinetic energy is energy that an object possesses when
in motion. Potential energy is energy an object can possess in relation to other objects or within its own stresses or charges.\(^4\) Where a conflict exists between a bot programmer and another user, bot, or actant within a digital ecosystem, whether that programmer is aware of it or not, there will be, in theory, always an inherent bias coded into the bot. That bias is conflict-laden as well as value-laden. We see these types of biases all around us, often without realizing it. In the introduction I brought attention to the text *Algorithms of Oppression* by Safiya Umoja Noble (2018). I referred to her research on Google’s hierarchal search algorithm as it favored hegemony and power structures that reflect the biases of the algorithms’ creators, ultimately leading to the hypersexualizing of women of color, for example.

This potentiality was present in my mind when I examined the discussions, archival information, and pages of the Addbot and when I interviewed researcher Aaron Halfaker and programmer Adam Shorland. It is clear that the Addbot was viewed from a number of perspectives: that of a tool that operated efficiently and effectively (Geiger and Halfaker); as quantitatively conflict inducing and controversial (Tsvetkova et al.); and as a source of value-laden bias and great potential for both good and bad outcomes (Noble 2018). Geiger and Halfaker (2017) also write from a viewpoint of bot conflict as processes that need to be updated and modified through a system of carefully managed bot governance. Tsvetkova et al. (2017) tend to take the viewpoint that bot conflict is increasing as we create more bots, suggesting that strong systems of bot governance may be constrained – more rules put into place – as more bots come into conflict with one another.

\(^4\) Adapted definitions from the standard definitions from Oxford’s *Lexico* (2019).
BOT GOVERNANCE AND AGENCY

Considering the aspects of bot conflict, bot purposes, and energy distribution of agency within an ecosystem, there is one overall factor that seems crucial to the determinations of these outcomes; that is bot governance, which is crafted by agents within a given system. Like all management or government systems, there are varying degrees of control. As in the case of the [loveMachine] in Chapter Two, there was a system of governance in place, but that system was radically unrestrictive—a “let’s see what happens” attitude that revised the bot only to keep it running undetected on Facebook. The Addbot, however, was subjected to an intensive process that Wikipedia users employed to ensure that bots are accounted for, the programmers hold accountability, the bots create the least amount of inadvertent conflict possible, and testing and revision take place during crucial moments during the bot’s lifecycle.

It can be said, then, that while I recognize the potential energy of agency and inherent values and bias within bots, even before they are created, I also believe that a system of strong rules and carefully managed bot governance policies (and bot approval groups, or BAGs) have a direct effect on the expressive agency of bots.

ACTANTS AND ACTANT CHARACTERISTICS

The following actants and their characteristics are noted because they were used to make ontological maps of rhetorical agency and lines of communication and persuasion (Figures 23-25).

- **Addbot** – Addbot, as mentioned previously, currently exists in several iterations (such as Addbot, Addbot 3, Addbot 4, Addbot 5), and they were used between March 2013
and March 2019. This research and subsequent ontology represent only the first Addbot from 2013.

- **Adam Shorland** – Employed by Wikimedia as a software developer for Wikidata; he is the creator of Addbot. He engaged with editors and staff to create, maintain, execute, and update the Addbot.

- **Wikimedia** – Self-described as “a global movement whose mission is to bring free educational content to the world. Through various projects, chapters, and the support structure of the non-profit Wikimedia Foundation, Wikimedia strives to bring about a world in which every single human being can freely share in the sum of knowledge” (Wikimedia.org, 2019a). Wikimedia contains under its umbrella such entities as Wikipedia, Wikibooks, Wikiquote, Wiktionary, etc. The Wikimedia Foundation was founded in 2003 by Jimmy Wales.

- **Wikipedia** – Self-described as “a multilingual online encyclopedia created and maintained as an open collaboration project by a community of volunteer editors using a wiki-based editing system” (Wikipedia.org 2019b). Wikipedia was the Wikimedia entity through which the Addbot made its reverts.

- **Wikipedia editors** – Unpaid, volunteer editors who crowdsourced their expertise under the system of Wikipedia rules of governance. They are human actants who interact with other editors to determine best practices and edits for Wikipedia articles. They are also subjected to edits by Wikipedia bots. They also can democratically provide feedback to halt bot functions or endorse bot functions as part of open, online discourse.
• **Wikimedia Foundation staff** – Paid employees who work for the Wikimedia Foundation. They serve in a variety of functions based on overall functionality and health outcomes of the Wikimedia ecosystems. As an example, James Forrester: Software Engineer, Product, self-describes his job as ensuring that Wiki tools for readers and editors/contributors are reliable, accurate, and fast (Wikimedia.org 2019c)

• **Wikipedia users** – Users of the platform who operate in a read-only function. They are information gatherers. They read and gather information and disseminate that information in a variety of media and formats in the digital and physical worlds.

• **Aaron Halfaker** – Research scientist for the Wikimedia Foundation since 2011. Halfaker also is a self-described “Wikipedian,” and he serves as both a Wikipedia user and editor. He researches topics such as Wikipedia bots and bot reverts (Wikimedia.org 2019b).

• **Stuart Geiger** – Staff ethnographer and principal investigator at the Berkeley Institute for Data Science at UC-Berkeley. Geiger contributed ethnographic data to a research paper co-authored with Halfaker that examined the discourse and data surrounding bot reverts and bot conflicts.

• **Wikipedia bots** – Wikipedia employs a number of bots that are subjected to approval through their rules of governance. These bots may or may not be affected by other bots. Wikipedia attempts to avoid bot conflict through a system of governance.
Figure 23: Early stage bot ecosystem without governance guidelines

Figure 24: Late stage bot ecosystem without governance guidelines
ONTOLOGICAL MAPS

In order to fully understand the scope and implications of governance systems on bots and their effect on rhetorical agency, we need to visualize what such a system looks like without strong governance rules in place. Bot lifecycles are not static—they evolve, grow, and change even when the code of such bots isn’t changed or adapted in response to environmental factors.

When visualizing Wikipedia bots, the first ontological map (*Figure 23*) shows how two Wikipedia bot programmers (Bot1 Programmer; Bot2 Programmer) develop bots to interact with Wikipedia articles. In theory, these two programmers and their bots could work in tandem with one another to change or edit a Wikipedia article; or the bots could cause a bot conflict with one another. In the early stages of a Wikipedia ecosystem without bot governance guidelines, Wikipedia editors who have an intimate, working relationship with Wikipedia might know the
bot programmers and what the bot programmers are attempting to accomplish. In this hypothetical non-governance system, during an algorithm’s early stages, the Wikipedia editors can give feedback to the bot programmers, or the editors could make changes to a Wikipedia article in the event an algorithm makes an error.

However, as Wikipedia editors and bot programmers discovered, over a period of time bots became disconnected from their programmers—forgotten, outdated, too numerous to keep track of, or confused with other bots. This model of late-stage bot algorithms (Figure 24) demonstrates the disconnection. When bots experienced bot conflict, or if they needed to be updated to stay relevant with current technologies, editing tools, or other methods of managing Wikipedia articles, Wikipedia editors no longer knew with whom to get into contact in order to update or remove a bot. The end result is a cluttered system of increasing conflicts, including conflicting power structures that degrade the Wikipedia ecosystem. An early stage algorithm Wikipedia ecosystem (without governance policy) seems almost utopian and hopeful. An end-stage algorithm Wikipedia ecosystem (without governance policy) is inefficient, confusing, and full of conflict, the opposite result of what Wikipedians desire and work toward. In short, the feedback loop becomes severed, with automated processes effectively bludgeoning Wikipedia articles with directives, edits, and reverts.

The Wikipedia community, over time, developed a governance system of guidelines and rules for bots, editors, and bot programmers. Somewhat democratic in scope (ultimately under the soft leadership of Wikimedia staff, who can override community members, i.e., editors), bot
ideas are proposed to interested Wikipedia editors and members, who voice their support and concerns as they attempt to plan for different types of risk and reward when the bots are executed. The bots are labeled with the word “bot” in the title, so everyone knows that certain edits come from bots. In the ecosystem with governance guidelines, experimentalism is restrained, with bots created with very specific purposes in mind. Wikipedians are able to levy their feedback (Figure 25). This is an example of centralized governance, which favors a balance of power and discourse in a more democratic system that employs checks and balances on bots and their creators. This increases credibility while identifying ownership. Ownership and connectivity to the bot is almost as central as an author is to a book. We are able to remove the expressive elements and mystery of the bot when it is attached to an owner who explicitly states the reasoning behind the bot’s creation and execution. Furthermore, the bot programmer is prescribed a methodology by Wikipedia editors for proposing bots, gaining approval, testing them, and refining them when needed (Figure 25).

The bot development cycle follows the bot from idea and potentiality, to planning, proposal, implementation, testing, deployment, support and maintenance, revision or remix, and back to discussion. The democratic function of all of these processes may seem cumbersome, but the functionality and success of Wikipedia continues to expand, even when its pages and articles increase by the tens of thousands each year.
Figure 26: Bot development cycle. [Source: Wikipedia, 2019].
At the time Addbot began to make its edits in March 2013, Wikipedia was already 12 years old (it launched January 15, 2001). Over the past dozen years prior to Addbot, Wikipedia expanded from being a “grand, social experiment” – at times called an “anarchy,” “dictatorship,” “democracy,” and even “adhocracy” – to a more codified, consensus-based system of editors (Konieczny 2009, p. 164). The differing of labels from various researchers and scholars demonstrates not only the difficulty in understanding or explaining the sociological phenomenon of Wikipedia governance but also the rapid changes and adaptations Wikipedia underwent in the first ten years as it grew exponentially in size and cultural recognition. In terms of the Internet, Wikipedia is one of the longest running digital ecosystems that is self-contained through a system of trial-and-error practices.

Like Wikipedia articles, Wikipedia governance rules and other administrative policies are edited on an almost daily basis, often rendering at-large perspectives and summations ineffective. Research articles that attempt to take a holistic view at Wikipedia governance vary greatly depending on the year. As an example, Konieczny’s view of Wikipedia governance through the lens of Robert Michel’s Iron Law of Oligarchy was written probably in 2008-2009, published in 2009, and based on data collected from Wikipedia: Verifiability from 2003 to 2006. The resulting paper is only somewhat applicable to the Addbot’s March 2013 lifespan, and even less applicable to my research of 2018-2019—all when examining Wikipedia’s evolution over a period of 18 years. Knowing this, it’s difficult to portray the ecological system that fostered Addbot; therefore, my descriptions of the Wikipedia environment are generalizations.
Addbot’s purpose was to remove interwiki language links “after a massive effort coordinated through Wikidata, a meta-level database created to centralize language-dependent data. For example, the links appearing on the sidebar of the ‘Japan’ article on the English Wikipedia to ‘Japon’ on the French Wikipedia and ‘日本’ on the Japanese Wikipedia were once manually curated on the ‘Japan’ page of the English Wikipedia” (Halfaker & Geiger, 2017, p. 49:2). Humans were manually curating and removing redundant hidden links from each page, until Adam Shorland created the Addbot to automate this process.

ETHICAL CONSIDERATIONS

Regarding modern moral philosophy, Alisdair MacIntyre (1998) suggests the main thematic questions are, “What kind of actions ought we to perform?” and “What kind of things ought to exist for their own sake?” (p. 249). We can easily examine ourselves within the governance and non-governance ecosystems of Wikipedia, observing users, editors, and bot programmers as they attempt to keep order within the pseudo-democratic, pseudo-utopian construct. What actions ought bot programmers perform in order to maintain order and fluidity? What actions ought editors perform in order to relay relevant, factual information to readers?

However, Ian Bogost (2012) suggests we use ontographic cataloguing so we can “hone a virtue: the abandonment of anthropocentric narrative coherence in favor of worldly detail” (pp. 41-42). Rather than approaching an ethical argument from the perspective of “ought to” or intent, we can examine the elements, or actants, of the ecosystem as alien objects, standing alone without the weight of other objects. Bogost calls this alien phenomenology, where we set aside associations and “assume the opposite: incompatibility. Lists remind us that no matter how fluidly a system may operate, its members nevertheless remain utterly isolated, mutual aliens” (p.
This type of disassociation and partitioning is a feature of object-oriented ontology (herein referred to as OOO). “Object-oriented philosophy holds that the relation of humans to pollen, oxygen, eagles, or windmills is no different in kind from the interaction of these objects with each other” (Harman, 2005, p. 1). Harman believed that “since the vacuum-sealed nature of objects makes direct communication impossible, all conjunction or coupling must occur through some outside mediator” (p. 2), rejecting phenomenology and theology in favor of object-oriented philosophy.

An ontological map allows us to view Addbot and its relationships to see what/who it is affecting while allowing each actant-object to stand alone in disassociation between all objects. In and of itself, the Addbot’s profile can be reduced to itself and its relationship with other objects—the Addbot as it exists in PHP or the Addbot and how it exhibits agency and persuasive nature upon the bot programmer, the Wikipedia editor, interlanguage links in Wikipedia articles, and so forth. A similar example can be found in Harman’s example of a banana: “A police officer eating a banana reduces this fruit to a present-at-hand profile of its elusive depth, as do a monkey eating the same banana, a parasite infecting it, or a gust of wind blowing it from a tree. Banana-being is a genuine reality in the world, a reality never exhausted by any relation to it by humans or other entities” (p. 74). But we “must not confuse the values of the design of objects for human use, such as doors, toasters, and computers, with the nature of the world itself. [An ontograph] shows how much rather than how little exists simultaneously, suspended in the dense meanwhile of being” (Bogost, 2012, p. 59).

The Addbot’s “meanwhile of being” is reduced to its lines of code in PHP, but it was coded with intent and purpose by Adam Shorland. In turn, Shorland’s ideas were affected by Wikipedia’s governance policy, which, in turn, was also shaped, either directly or indirectly, by
the output of the Addbot. The algorithm was coded with a purpose, just as a machinist creates a screw with a certain thread count, length, type of head, and specific distance between threads according to an engineer’s specifications. Like the engineer’s spec sheet, the Wikipedia governance policy exerted a mediating force upon Shorland’s algorithm. Where the machinist says, “I can make and mold anything metal with my milling and routing machines,” the machinist also was guided by purpose and intent from another relational force exerting agency. In Chapter Two, we saw how the Facebook [loveMachine] was experimental—the programmer self-identified as an artist. In the machinist metaphor, the [loveMachine] was a screw made from an imagination, with the general population attempting to find a use for it, or rejecting it outright. All of this is to say that in the OOO philosophy, we can examine two things that exert changes in rhetorical agency: individual objects that are capable of exerting a relational force upon other objects and power coalitions that exert forces upon others. Shorland’s Addbot is, theoretically, an assemblage of code, but, individually, the lines of code are incapable of being the Addbot. Once running, the assemblage is useless in and of itself unless the Addbot is run within the space for which it is built. Without interlanguage links within the Wikipedia ecosystem, the Addbot remains an individual object. However, once executed within the Wikipedia ecosystem, the Addbot is efficient and able to interact with hundreds of thousands of articles, across dozens of different languages. The changes, in turn, affect users, editors, staff, articles, links, researchers observing bots and Wikipedia, and many other things.

Because such massive changes can be made within the ecosystem for which is was created, the Addbot was subjected to a host of power coalitions, hegemonic forces designed to restrict the bot to fit a certain criteria. These forces created a governance policy by which the bot’s programmer had to abide, or the bot would be disabled. Because the bot was making global
changes, it was subjected to the strictest controls on Wikipedia (see Figures 26-27). In fact, user Snowolf raised a concern that the Addbot wasn’t meeting the criteria to warrant global status (see Figure 27), saying, “I’m afraid your bot does not satisfy the requirements of the global bot policy, namely ‘the bot must already be active on several wikis, with long-term contributions to back up its trustworthiness.’ as it is active on one wiki” on the Steward requests/Bot status page (Wikipedia.org 2019d). Although another user and Shorland demonstrated to Snowolf that the Addbot was operating under global governance policy criteria, there were several instances where different Wikipedia editors shaped the Addbot—they provided feedback, they cross-checked the bot to see if it conformed to guidelines, they defended the Addbot’s agenda, and they catalogued what it was doing (Figure 28). Snowolf was referring to the global bot policy that states several prerequisites a bot must meet before it can be expanded to other languages (Figure 28). Adam Shorland, listed as Addshore, responded that the bot is currently running on the English and Hebrew wikis and is likely to be expanded further. This exchange is one small example of the numerous transactions issued by the Wikimedia community as they self-regulate their own policies in a public forum that is transparent and open to the public.

The feedback provided to Shorland, and the public, was shaped by the Wikipedia bot governance policy. In turn, the Addbot (theoretically) shaped future governance policy. Shorland said the Addbot started off on small sets of Wikis (such as hewiki—Hebrew Wiki), earning approval beforehand and then gaining approval and momentum as it proved its success, finally earning global approval to work across most or all of the language wikis (Shorland, Personal correspondence, 2019). This process served to reinforce that the approval and governance systems were working and effective; this same process could be repeated with other larger Wikipedia bots as programmers tested them on smaller wikis before moving to global changes.
Authorisation  [edit]

There are three ways to obtain bot access on a wiki.

Global bots  [edit]

Global bots are given access on all wikis that allow global bots (this must be explicitly permitted by local policy, see list). To qualify for global bot access, the following requirements must be met:

- the bot must only maintain interlanguage links or fix double-redirects;
- the bot must already be active on several wikis, with long-term contributions to back up its trustworthiness.

Global bot access should be requested on Steward requests/Bot status. (These rules do not prevent bots from performing other tasks on a wiki with its community’s permission.)

- The global bot status just gives the bot access to a restricted number of projects and wikis where they have the bot flag locally. You will find an exhaustive list of wikis where the global bot flag is active here. (You can continue to request local bot flags on projects where global bots are not allowed.)

Automatic approval  [edit]

If automatic approval is explicitly allowed on the wiki, bot operators can request a local bot flag directly from stewards. To qualify for this, the following requirements must be met:

- this must be permitted by the local bot policy or request page;
- the bot must edit regularly without a bot flag for at least a week or make 100 edits on the applicable wikis for demonstration purposes;
- the bot must only maintain interlanguage links or fix double-redirects.

Community consensus  [edit]

Otherwise if there is a local community interested in processing bot applications, bots must obtain community approval on the most relevant local discussion page before editing without a bot flag at high speeds or without human supervision. Once there is consensus, a local bureaucrat will add the flag, or a steward may be requested to do so.

If there is no local community and the above does not apply, the bot must operate without a bot flag or not at all.

Figure 27: Bot access and authorization on Wikipedia [source: Wikipedia.org].

The bot will remove interwiki links from projects where the links can be found on Wikidata as it is rolled out. The code is currently running approved on ENwiki.

- **Addshore** Talk To Me! 23:37, 25 February 2013 (UTC)

I'm afraid your bot does not satisfy the requirements of the global bot policy, namely “the bot must already be active on several wikis, with long-term contributions to back up its trustworthiness” as it is only active on one wiki. Snowolf How can I help? 06:19, 26 February 2013 (UTC)

I found out tonight that a lot of bot request pages have an English translation - there's probably enough so that you can get enough for a global bot. --

**Rachen7754** 10:06, 26 February 2013 (UTC)

Currently running on EN and HE wiki and in the next few days this is likely to grow further. The code used in the global section of the bot is the same as the code that has so far made nearly 500,000 edits on EN wiki and is currently running across every article in HE wiki. The only thing that changes in this code from language to language is the edit summary. **Addshore** Talk To Me! 14:08, 26 February 2013 (UTC)

Please see completed run at ויקימדיה/Addbot. Looking to be ready to deal with the rollout of wikidata in a weeks time. **Addshore** Talk To Me! 06:53, 27 February 2013 (UTC)

Will your bot only remove interwikis from the main namespace? Regards. **inavadinado** 12:01, 4 March 2013 (UTC)

No, any namespace that is included on wikidata (Article, cats, templates and project) **Addshore** Talk To Me! 16:18, 5 March 2013 (UTC)

Done. PeterSymonds (talk) 20:56, 6 March 2013 (UTC)

Figure 28: Archived feedback of Addbot status on Wikimedia.org [Source: Wikimedia.org].

However, even with the tightest controls, extant forces outside of Wikipedia have written of the Addbot, and similar global bots that make large amounts of changes, as ultimately negative. Tsvetkova et al. (2017) studied Wikipedia bots and reverts, which occur “when an editor, whether human or bot, undoes another editor’s contribution by restoring an earlier version
of the article” (p. 3). They called reverts “disagreements” and used reverts to measure disagreement. They found that although bots had fewer reverts than humans, “the number of reverts between bots has been continuously increasing” (p. 4), suggesting that bots are not getting more efficient and bot owners are not learning how to identify and solve bot conflicts as quickly as the conflicts are arising. More sensationally, Matt Simon penned a 2017 *Wired* article in response to Tsvetkova et al.’s. “Even Good Bots Fight,” amplifying the “botpocalypse” rhetoric: “But a funny thing happens when you lock a bunch of bots in a virtual room: Sometimes they don’t get along. Sometimes a pair of bots will descend into a slapfight, overwriting each other’s decisions thousands of times for years on end” (n.p.).

Simon isn’t altogether incorrect. Unlike artificial intelligence, which can aid algorithms in rethinking procedural roadblocks, the simplicity of social media algorithms ensures that their code will work like a sledgehammer, endlessly executing code until a human or other object disrupts the cycle. *Bot conflict* is where two procedures collide with one another, but that doesn’t mean there’s necessarily an all-out war, especially in places like Wikipedia. However, according to Stuart Geiger, Wikipedia bot conflict is “relatively rare in Wikipedia, and it typically gets noticed and fixed pretty quickly. Bot Approvals Groups and the bot policies generally do a good job at keeping bot developers in communication with each other and the broader Wikipedia/Wikimedia community” (Halfaker, 2017).

Geiger mentions that there is a second type of bot conflict, and that’s when humans decide what bots should do. This returns to the question of MacIntyre’s “ought” and “should” and the moral dilemma we face from a postmodern philosophy. The Wikipedia editorial community is building a hegemonic structure of governance around bots and their purpose. I argue that it’s possible that because of the nature of the Wikipedia ecosystem and severe
limitations placed upon bots, editors aren’t necessarily deciding the “oughts” and “shoulds,” rather the parameters of the bot. The bots are dumbed down to a set of procedures and processes overall. They lose their human characteristics and expressive potential. Even though we are willing to turn even semi-expressive bots into human-like actants in our search for meaning-making, it is difficult to do so with a Wikipedia bot. When a specific parameter is built into the bot, the metaphysical moral dilemma is removed. The bot is not decision making. It is a finely crafted tool, in this instance, and should not receive the anthropomorphic designation humans have given it. Even the term “bot” is reminiscent of a robot, and a human figure of a robot, at that. Just go to Google Images and type in “bot” and you’ll find a colorful array of cartoonish, humanoid robots used to convey the imagery of a bot algorithm (*Figure 29*). The culture we have built behind giving a bot a personality is a falsehood, one that offers convenience and a wishful, emotional connection. Society should dismantle the view of bots as expressive beings in order to clearly define ourselves as apart from automated processes. When we treat bots like communicating or sentient beings, or other people behind a digital veil, we set ourselves up for rhetorical failure. Our emotional connections to these utilities are no more functionally significant than our connection to a favorite watch or pair of shoes.

The Addbot comes clearly stamped as a tool. It is called a bot. It performs the functions of a bot. When humans – within the Wikipedia sphere – discuss the Addbot, they speak about it in utilitarian terms and not of it as having a personality. Outside of the Wikipedia ecosystem, the Addbot is clearly anthropomorphized, as evidenced by reactions from some researchers and the media. The world outside of Wikipedia is filled with bots of which we attach emotional meaning, so it’s not fair to expect Tsvetkova, her fellow researchers, and tech reporters at *Wired* to make
the philosophical leap from “violent” bots to tools that are performing menial tasks on a large scale so humans don’t have to.

By regarding simple algorithms—“procedures that always produce the same result” (Brown, 2015, p. 512)—as repetitive machines we can view them more as tools of industry and not as human-like minions or robot servants. And, by doing so, we can avoid such terms that center on human-ness—war, conflict, anger, playfulness, creative, interesting, etc.—and begin the process of dehumanizing bots for the sake of human discourse. James Brown, Jr., (2015) notes that Wikipedia recognizes bots as “users” with their own user pages, and as tools, and expresses that it is often difficult to distinguish a bot from a human being. But Wikipedia does very well in distinguishing a bot from a human, at least to any seasoned editor of Wikipedia.

By creating a distinction between humans and bots where possible, and by framing the bot’s directive through the use of governance policy, Wikipedia is able to rein in disorder and
conflict. There is a clear distinction between human and algorithm, and the Wikipedia ecosystem has so far been very careful to keep that distinction important. Wikipedians have over time recognized that creating bots and setting them loose without keeping track of them, updating them, or giving them lifecycles lead to disorganization and chaos. Wikipedia thrives on organization and pseudo-democratic autonomy, and bots figure into that aesthetic. Adam Shorland, the creator of Addbot, admits that without the Bot Approvals Group on Wikipedia, he likely would have run his first Wikipedia bot on his own. His bot approval request, however, was declined, which he now appreciates. But Wikipedia’s procedures for approving bots is slow: “BAG on enwiki [English Wikipedia] is probably the most thorough process, but probably also one of the slowest, other than wikis with very small numbers of users to approve bots” (Halfaker, 2017). In return for security, efficiency, and organization, Wikipedia sacrifices speed, innovation, and, to an extent, outright experimentalism. Wikipedia, however, relies on organization and democratic discourse relatively free from outside manipulations in order to keep structure and information management intact.
CHAPTER 5: CONCLUSION: TOWARD A RESPONSIBLE FUTURE?

ALGORITHMS AND AGENCY

This exploration into simply coded, social media algorithms began with an observation I had while teaching first-year composition students. Although they, and I, had grown up in the age of the personal computer and development of the Internet, I realized that they knew surprisingly little about how computers, computer hardware and software, and interfaces functioned. I suspect my father had the same issue with cars when he was my age; he was a mechanic, always perplexed by automobile owners who didn’t clean their air filters, rotate their tires, or put the correct oil in the engine. “You have to know how to fix what you use,” he told me one day, offering advice that I still try to follow, though many times unsuccessfully in this busy world. In the introduction I quoted Carolyn Miller (2001) when she inferred that most computer users only know how to navigate on a computer or the Internet. They know how to push the proverbial gas pedal and cruise around town.

Is it important for laypersons to know more than how to point and click? In this new age of anti-intellectualism and disinformation it’s important that the public be aware of the potential energy and discourse-altering properties of social media algorithms. More importantly, researchers and rhetoricians need to drive this discourse in order to safeguard philosophical ideals such as credibility, logic, and trust. Without those, institutions such as journalism, the university, and ethics watchdogs find themselves on uneven footing. Rhetorical agency – its
flow, potentiality to shape power structures, and power to make dramatic and radical change – is a leading determinant of the future of such ideals and institutions. Social media algorithms, or bots, are a key component of shaping agency in an increasingly digital world.

On May 24, 1844, Samuel F. B. Morse telegraphed, “What hath God wrought,” marking a new era in communication. Along with the numerous advantages in near-instantaneous communication over long distances, this new invention also introduced a new frontier—a struggle for power and control over this technology. Armies changed their tactics in war, keeping leaders far from battle, and profoundly changing military structure and command. Companies fought for control of telegraphing technologies, financial investments, and profit at the same time they expanded its use and reach. Newspapers with access to telegraphed information had a decisive advantage over their rivals, controlling information and disseminating and receiving information much faster—“scooping” their competitors. In short, the introduction of a technology that introduced new methods of communicating upended hegemonic structures in a very short amount of time. Trish Loughran (2007) writes in The Republic in Print that the invention of the telegraph was partly to blame for speeding the country toward civil war. The conquering of time and space forced the North and South to come to a reckoning in the vastly different viewpoints each geographic region held, as well as where the country should be heading. Sean Trainor (2016) writes that a less obvious impact of the telegraph had to do with rhetorical style, namely that of then-President Lincoln: “Historian Garry Wills has argued that the telegraph’s forced brevity revolutionized Abraham Lincoln’s rhetorical style, leading the 16th president toward the famous terseness of the Gettysburg Address” (n.p.). The same restructuring and reshuffling of hegemony likewise occurred after the introduction of the printing press and moveable type, electricity, the telephone, the modem, and so forth.
Bot algorithms are part of this trajectory of technology rewriting power dynamics, and its impacts are both obvious and subtle. Like the telegraph shaping Lincoln’s rhetorical style, social media bots shape our styles and strategies in ways that are difficult to notice or perceive. Changes in rhetorical style coexist with the dynamics of rhetorical agency. Lincoln’s terseness in the Gettysburg Address was repeated and transmitted across great distances and with great immediacy. Large audiences were, at the very least, introduced to a different manner of discourse. Lincoln’s address to the public at Gettysburg was brief – only two or three minutes – and followed former Secretary of State Edward Everett’s two-hour speech. The brevity of the telegraph, it seems, affected Lincoln’s rhetorical strategy, since transmitting a more traditional speech via telegraph, such as Everett’s, would have been quite laborious. This example helps me to illustrate the point that bots, too, offer constraints. Twitter bots, such as the Peñabots, needed to strategize in 160 characters or less. In the era of near-instantaneous news, feedback, and reactions, Twitter capitalizes on these rapid-fire changes and exchanges by limiting compositions to short “tweets.” Bots are used to manipulate these constraints into advantages over humans, as exemplified by these case studies.

Agency is affected by the constraints of time and space, by strategy and situation, by kairos and timing, by composition size, by semiotics, and so on. Those who are able to game these constraints to their advantage are able to shift power dynamics in their own favor or toward a desired end. As I mentioned above, the advent of the telegraph created a rush to capitalize on the technology in many different arenas: the economy, military, sciences, journalism, politics, etc. As the telegraph went, so go the algorithms. Because algorithms can perform vast tasks of microlabor in a digital space, they greatly shift the ability to alter discourse and agency, distributing and redistributing rhetorical agency across human and nonhuman actants in a given
ecosystem. Algorithms also confound the distribution of agency when they are viewed as expressive objects that we wish to engage with on a personal level. Successful social media bots do well because other actants are able to game these algorithms to their advantage, altering constraints to shift agency in ways that humans cannot on their own. They do this through a variety of strategies, as observed in the three case studies, which I will delineate below.

FINDINGS

Becker et al. (2012) developed their frameworks of rich data analysis with the following purpose in mind: “to look for patterns among the data and look for patterns that give meaning to the case study” (p. 11). Aside from patterns developing across each case study that connected them to one another, there were individual characteristics, or sets of characteristics, that allowed each type of social media bot algorithm to stand apart from one another. I found that those differences were just as important as the patterns. There were differences in rhetorical strategy and intent, planning, and interaction with other actants, for example. As I extrapolated the meaning from these differences, I concentrated more intimately on the patterns of similarities shared among the three case studies.

Rhetorical Strategy

As orators and writers, we might attempt to prepare ourselves before giving a speech or writing a paper. We determine how a variety of factors might affect our discourse: kairos, audience, medium, purpose, the current rhetorical situation, etc. Bot programmers should proceed in the same manner as they construct their automating rhetorical devices. Although bots
are almost always used for purposes other than for what they are intended, they are still created with a strategy or purpose in mind. In addition to conceiving of an idea for an algorithm, programmers must think and code procedurally—what is called the procedural programming paradigm, which offers “the ability to reuse the same code throughout a program through procedure calls, functions, and multiple files” (Bogost 2007, p. 12). How that strategy was implemented, however, varied greatly among the three case studies.

The first example, the Facebook [loveMachine], was created with a simple strategy. Julien Deswaef built the bot so it would provoke disruption and discourse by “liking” everything it came across on Facebook. While constrained by the digital ecosystem of the Facebook platform – the bot could be detected via a threshold of “likes,” usage, and/or automation, for example – the [loveMachine] was designed to test those constraints and challenge others to question how they viewed their own limitations: What does it mean to “like” something on Facebook, anyway? Although there was a strategy, there was also a sense of purposeful disconnect. Deswaef stated that he wanted to see what would happen, not that he knew or might know what would happen. He conjectured that he thought his bot would cause some issues with trust, communication, and social interactions between his account and his “friends,” but he was purely speculating (Personal correspondence, 2018). His bot was built to be expressive and to both bolster human agents and their meaning-making of the human characteristics of the [loveMachine], but to also provoke the questioning and validity of “Juego Requiem,” the nom de plume of the Facebook account and bot.

I chose the Twitter “Peñabot” botnet case study to showcase how a group of automated voices can shape and change narratives. Where the [loveMachine] was one, experimental voice, the Peñabots numbered in the tens of thousands as they executed hive-like, simple tasks, like
poisoning hashtags and creating trending topics. The [loveMachine] had a simple task, too: to keep clicking the “like” button at a rate that it would not be detected by Facebook. But what if the [loveMachine] had been implemented on 75,000 Facebook accounts (the equivalent number of Peñabots acting on behalf of President Nieto, as estimated by Alberto Escorcia)? How would the [loveMachine] experiment change? Assuming Facebook wouldn’t lock down the accounts or even detect them, the bots could affect entire networks of friends and family, taking into account that the average Facebook account has 338 connections (Smith, 2019). That would mean over 25 million accounts, overall, could be affected by a [loveMachine] botnet. Collectively they could disrupt the value of the “like” button by devaluing the meaning and agency of the action. When we “like” a photo of someone on Facebook, they immediately understand that we have seen the photo, we don’t disapprove, and we are communicating our appreciation for the photo (no matter how trivial or meaningful it may be to us overall). But when every photo we post is “liked” by the same several people, no matter what the photo subject may be, we might begin to question the value of the “like” button. It, in a sense, becomes meaningless. If Deswaef had been able to install the [loveMachine] as a large botnet, his experiment could have, theoretically, altered the validity of the Facebook platform itself.

Although I’m speculating about what could happen if the [loveMachine] were implemented as a botnet, I’m not wildly theorizing; this scenario of devaluing discourse through the overuse of the “like” button has echoes in the hashtag poisoning of the Peñabots. Hashtag poisoning isn’t that dissimilar to the devaluing of the “like” button. Automating and overusing a routinized process devalues the meaning until the landscape of discourse is fundamentally changed. People must adapt and change their strategies or risk being alienated.
The Peñabots were created to be used in large groups in order to overwhelm traditional flows in rhetorical agency. When Mexican activists attempted to start a trending topic on Twitter to offer a critique against President Enrique Peña Nieto, the Peñabots flooded Twitter with a sea of textual noise, mostly in the form of sharing the same hashtag. This gamed the novelty index of Twitter, purposefully triggering Twitter’s automation-detection algorithms. Also, the Peñabots obfuscated legitimate discourse with the overwhelming tweets from the botnet. They altered the dynamics of rhetorical agency in two ways. First, the strategy was to manipulate Twitter’s algorithms; they worked together to trigger a constraint in Twitter, which controlled how discourse was disseminated. The activists worked toward a reward offered by Twitter: recognition and mass dissemination of a topic that they cared about and worked for. The Peñabots halted this process by making the activists’ trending topic suspect. The kinetic energy of the activists’ agency was quite literally halted. The notion of agency as “a dynamic, distributed dance that cannot be turned off and on like a water fountain” (Gries, 2015, p. 70) did not seem apply to this scenario. In this instance of information warfare, the main purpose is to disrupt the call to action and disrupt the “virality” – a direct by-product of agency – in social media, and the Peñabots shut the agency valve off for a time.

Likewise, the [loveMachine], even as a standalone bot without the cumulative effect of a botnet, operated by disrupting the flow of agency. Consider this example: Let’s say you gave a practice speech in front of a classmate in preparation to giving it in front of the entire class. When you ask your classmate for feedback, he tells you he likes the speech. It’s great. You feel as if you don’t need to make any more adjustments. But when you give the speech in front of the class, your peers and instructor remark that the speech wasn’t very good and could use several changes. It turns out the classmate wasn’t listening; he didn’t want to say anything negative,
either. This is an example of disrupting credibility, which indelibly belongs within any defining characteristics of rhetorical agency. When we take action within a framework of rhetorical discourse, we measure that action against similar actions. This is also how we measure credibility—*who else that I trust is saying the same thing or something similar?* The fictional speech giver was relying on the classmate to provide legitimate feedback but learned that his classmate was unreliable. The speaker switched from measuring credibility through one person to many. The concept of using proof as a method of persuasion has been ingrained in Western philosophy for over 2,000 years:

> Of the modes of persuasion furnished by the spoken word there are three kinds. The first kind depends on the personal character of the speaker; the second on putting the audience into a certain frame of mind; the third on the proof, or apparent proof, provided by the words of the speech itself. Persuasion is achieved by the speaker’s personal character when the speech is so spoken as to make us think him credible. (Aristotle 1926, p. 1356a).

Convincing others through credibility-driven *ethos* builds believability, especially over time and through repetition. The [loveMachine] tested the threshold of believability and credibility by “liking” everything it came across. It “liked” the deaths of real of people; it “liked” breakups and the ends of relationships; it “liked” all manners of products and people and pages. It was impossible to tell what it didn’t “like,” and real people on Facebook found it increasingly difficult to tell what “Juego Requiem” stood for and whether he could be trusted. Through automation the algorithm disrupted its own credibility even though it was automating positivity. Its expressive nature allowed the [loveMachine] to be treated like another person – as exemplified by the ELIZA example earlier in this text – and that expressive nature was also its downfall.
Expressive Behaviors and Meaning-Making

Expressiveness in bots presents certain challenges that non-expressive bots don’t face. The Wikipedia Addbot was carefully cultivated through a process that included planning and preparation, testing, listening to or applying feedback, revisions, and an expansion into a global role once the bot had proven itself to function according to plan within certain language roles, such as “hewiki” (Hebrew Wikipedia). Wikipedia users, editors, and staff members had an opportunity to connect directly to the creator, Adam Shorland, to express dissatisfaction or approval, to offer edits or suggestions, or ask questions. If Shorland didn’t respond to the community and their questions and concerns, they could opt to suspend his bot. Bots are regulated on Wikipedia using a model similar to how Wikipedia articles are edited: through “decentralized consensus-building, scaffolded by formalized policies and processes” (Geiger & Halfaker, 2017, p. 7). More specifically, the community authored a bot policy, which is administered by the Bot Approvals Group (BAG). A system of accountability ensures that bots are held accountable to their specific task or set of tasks, and the programmer must convince the community of the use and validity of the bot. This type of rhetorical strategy mimics a review and revision practice that speech givers and writers undergo in tightly controlled systems. The planning and execution of a bot within Wikipedia is roughly analogous to a student in a university setting attempting to craft a paper. Crafting an essay, for example, requires an idea; prewriting exercises; writing within constraints/assignment (style, audience, length, subject, etc.); drafting, feedback, revision, and repeating these processes until the essay is tightly controlled within its constraints. The rhetorical strategy, ultimately, falls within those constraints and is subject to the rules of the given system. The caveat is that the rhetorical velocity and effect of a bot in action operates on a far greater scale than a university essay.
Wikipedia sets the rules in order to turn bots into tools or utilities that perform a service beneficial to its own ecosystem and community. In that system, bots are called bots, and they are connected to their programmers and conveyed as utilitarian tools. Conversely, when we examine bots such as the [loveMachine], we find general confusion when humans question what it means to confront a bot that acts expressively. Although the [loveMachine] was explicitly declared to be a bot on several occasions (such as on the bot’s Facebook page and on Deswaef’s blog) and generally known to some of Julian Deswaef’s Facebook friends as a bot or bot experiment, it continuously received interactions and reactions from Facebook friends as if it were another human being acting otherwise erratically. People questioned the motives and meaning of “Juego Requiem,” at times acknowledging that it was a bot but still treating it like another Facebook friend. By removing as many humanized traces as possible from a bot, it becomes more of a function that is connected to its programmer, the human responsible for generating agency. However, when a bot is given expressive powers, it is imbued with agency at the same time it becomes disconnected from the programmer.

When expressive bots become a focus of our interactions, whether we mean them to be or not, they assume the role of interlocutor. Carolyn Miller (2007) and Ian Bogost (2007) both used the MIT chatbot ELIZA as both a point of reference for procedural rhetoric that bridged the divide between human programmer and audience. “Patients” responding to the rudimentary psychology chatbot weren’t under an illusion that they were speaking with a real psychologists, yet they afforded the chatbot similar courtesies, even treating ELIZA as a professional psychologist with a juvenile command of discourse. We recognize that programs like ELIZA – these days echoed by other chatbots such as Verizon tech support – aren’t human, yet we still interact with them in human-like ways. Expressing anger to a Verizon tech support chatbot might
elicit a sympathetic response, but the likelihood that someone is reading, verbatim, about your plight and caring from an emotional standpoint is probably remote, yet we still complain to the chatbots. To them we will offer a pleasant “thank you”; we say, “please”; and we exchange pleasantries—all of this is wasted time and labor, but only if we consider it a waste.

We should exhibit caution, then, when it comes to expressive algorithms that are good at mimicking human behavior, forcing us to question ourselves whether we are talking to a bot or a human. The Peñabots, while sometimes easy to identify as bots because of their collective similarities when they inundate us with repetitive hashtags or other such tactics, are at other times impossible to distinguish from humans. They use human names or handles; they provide Twitter biographies that are similar to regular people. They say they are parents, employees, churchgoers, pet owners, and fans of sports teams. They use politically biased speech, designed to divide, support, harass, or obfuscate, but in such a manner that we don’t know if the tweet is coming from another human or a bot. We engage in political discourse as we attempt to tell the bot that it is incorrect, unaware that the bot is using our labor and time to assist it in its mission. The rhetorical strategy of malicious, expressive bots such as the Peñabots, therefore, is to use as little energy as possible to confound normal, rhetorical discourse by having unwitting human agents perform labor for it. In the natural world, there is a word for this: parasitism.

Controversial topics, opinions, and memes serve as a catalyst for this sort of reaction to occur, a lure of sorts meant to draw us in to act performatively.

Examining Bots as Individuals or as Tools

Researching bots as independent actants will produce different outcomes than when researching them as actants connected to a programmer. Geiger and Halfaker (2017) note that in
their study, “Operationalizing Conflict and Cooperation Between Automated Software Agents in Wikipedia,” bots “are inextricably linked to their developers,” whereas Tsvetkova et al.’s “Even Good Bots Fight” (2017) examines bots as “autonomous agents who can be studied independently of the people who develop and operate them” (Geiger & Halfaker 2017, p. 26). Viewing bots and programmers as an assemblage demonstrates authorship and ownership, which, in turn, allows for responsibility and accountability. By removing the programmers from their bot research methods, Tsvetkova et al. argue that bot-on-bot conflict is inherently autonomous and increasingly problematic in the Wikipedia ecosystem. This theory is disputed by Geiger and Halfaker, who use several examples to delineate bot-on-bot conflict as 1) procedurally driven (where two programmers agreed on a goal, but disagreed with implementation), 2) task driven (bot programmers disagreeing with a goal, or bots conflicting with one another due to a bug or error in programming), and 3) interpersonal conflict (conflict devolving from procedural or task conflict into interpersonal conflict). Their examples demonstrate the inseparability between the assemblage of programmer (or bot developer) and bot within the Wikipedia ecosystem, which has a direct translation to rhetorical agency. Geiger and Halfaker’s bot conflict definitions and constraints are, however, limited to the Wikipedia ecosystem and do not translate well across different platforms.

I attempted to demonstrate the movement of agency through an assemblage, or litany, of actants and actors through the use of ontologies. Mapping the flow of influence and expression reveals areas where we can effectively manipulate agency, much like redirecting water from one place to another and changing the water pressure or simply turning the faucet off. Examining individual actant-objects, while beneficial for understanding their own contributions to an assemblage free from social constraints, is a focus of Bruno Latour’s actor-network theory
(ANT). There’s an “important reason,” Latour (2005) says, “for rejecting adamantly the role given to objects in the sociology of the social: it voids the appeals to power relations and social inequalities of any real significance” (p. 85). Assemblages of objects and actants, therefore, can make researchers unwitting transmitters of true social inequalities and the causes thereof. By recognizing these objects within a given ecosystem and charting their influences, I attempted to recognize the individual properties of each object while acknowledging their place within the assemblage.

It is, therefore, understandable and helpful that Geiger and Halfaker offer a rebuttal to “Even Good Bots Fight”—there is inherently a weakness in quantifying bot conflict without considering its place within the assemblage of bot, bot developer, and bot governance. However, it’s possible that Geiger and Halfaker have become the (somewhat) unwitting “sociologists” Latour (2005) is warning us about: “It is much too tempting to use power instead of explaining it and that is exactly the problem with most ‘social explainers’: in their search for powerful explanations, is it not their lust for power that shines through?” (p. 85). Geiger and Halfaker, it must be noted, in their spirited defense of the Wikipedia community and self-governing policies, acknowledge the historic lack of representation of women editors on Wikipedia and that “Wikipedia talk pages in general can be a hostile environment, particularly for newcomers and members of underrepresented groups” (p. 26). But as self-recognized Wikipedians with over ten years of experience on Wikipedia, Geiger and Halfaker’s defense of automated processes and bots might obscure their privilege. Returning to Noble’s *Algorithms of Oppression* (2018), a case can be made that Wikipedia’s predominantly male editors and bot developers automate their inherent biases in ways that are born out of the very assemblages that elucidate the complexities of human-to-algorithm interaction. Bot developers and the Wikipedia community might
artificially attempt to remove the expressiveness of the bots, but the values and biases of the bots remain unchecked in a system that is predominantly homogeneous. There is an important distinction between bots that are tethered to their developers and expressive bots that are loosed into a digital space, leaving the developer free from accountability.

**Anonymity, Accountability, and Agency**

I once lamented to a professor that we lived in an age of digital anonymity. “People need to take responsibility for what they say and do,” I wrote in a blog post. This was around 2010 (when blogs were somewhat new and still interesting), and anonymous trolling was a still somewhat novel concept. People were saying things to one another that they wouldn’t dare say in public, and the rest of the digital audience gasped and howled. My professor warned me against the pitfalls of removing anonymity. “You’ll stifle creativity,” he said. He also spoke of how anonymity empowers oppressed and suppressed groups. My argument was that anonymity emboldened people to act in negative ways without fear of retribution, and these behaviors might outweigh any benefits on behalf of creativity or marginalized groups. I return to this exchange often when thinking about accountability, ownership, and responsibility. Reviewing these case studies led me to a new insight to that exchange—there can be a sense of accountability among anonymous accounts.

Wikipedia is an example of such a system. While not a perfect system, communal governing and self-regulation keeps the ecosystem mostly efficient. Wikipedia achieves this efficiency in part through different networks of accountability. For example, you can contribute in a limited role as an unregistered user. An unregistered user’s identity is their Internet Protocol (IP) address. Since others can discover more through an IP address, registering can actually
improve and increase privacy. Registering allows users to use a handle, or nickname, which hides their IP address within their account. Wikipedia visitors can also be auto-registered if they’ve been on Wikipedia for at least four days and made at least ten edits to articles. Some users opt to make themselves known, either through name or through commenting in forums. Others prefer to remain anonymous. Some users, Wikipedia says, complain that anonymity, and unregistered users, allows fabricated edits, disinformation, trolling, and a host of other issues, but Wikipedia counters most of these assertions and makes a case for unregistered users on the Wikipedia:IPs are human too wiki. Wikipedia says that unregistered users on talk pages help to build consensus; most are not “vandals,” but contributors; and most don’t participate in “sock puppetry” or “ballot stuffing” (using several IP addresses to “create the appearance of a greater weight of opinion than really exists” (2019, n.p.). Wikipedia employs something akin to vaccinations in a herd—the more consensus building that occurs, the more difficult it is for an anonymous user to disrupt the consensus, especially under dubious circumstances. Bots fall further down the spectrum of accountability. The programmer might be anonymous, but she must take part in a rigorous process of proving accountability, credibility, and reliability. In addition, she must create the bot under a different account, which she is responsible for. The opening paragraph from Wikipedia:bot policy delineates the basic structure under which bots can begin to operate:

*All* bots that make any logged actions (such as editing pages, uploading files or creating accounts) must be approved for each of these tasks before they may operate. Bot approval requests should be made at Wikipedia:Bot/Bots/Requests for approval (BRFA). Requests should state *precisely* what the bot will do, as well as any other information that may be relevant to its operation, including links to any community discussions sufficient to demonstrate consensus for the proposed task(s). In addition, prospective bot operators should be editors in good standing, and with demonstrable experience with the kind of tasks the bot proposes to do. (2019, n.p.)
To wit: in order to operate a bot, a person must be a proven editor with a history of positive contributions to the community of Wikipedia, and even then, she needs to plan and propose the bot to a panel of proven moderators.

Bots, therefore, operate within a frame of anonymity, but within the Wikipedia ecosystem they are carefully planned, maintained, and monitored. Their programmers are held accountable to the actions of the bot. The community has a say over how their discourse and actions are affected. Every bot and bot action has to be approved, and if a bot’s action affects a community member, he or she can make a comment and connect directly with the programmer. Accountability and credibility are values that figure prominently within rhetorical agency and how we answer a call to action. By performing a public good for the community, bot programmers offer their creations to promote efficiency, literacy, and other positive goals meant to benefit others.

Conversely, in the case of Twitter and the Peñabots, for example, bots exist in tethered and untethered formats. ToS agreements notwithstanding, it’s almost a personal choice of the programmer to announce his or herself and a bot creation. In Chapter One I talked about the @IMG2ASCII bot (see Figure 2) by Colin Mitchell. Mitchell self-identifies as the bot creator, and he makes a personal choice to remain connected to his bot as it responds to requests to convert images to ASCII symbols. But Mitchell could anonymize not only his name and information, he also doesn’t need to identify his bot as his creation. Twitter bots created for the public good, such as the @threadreaderapp, or for benign amusement, like @a_quilt_bot by Bob Poekert (that turns images into quilted works of digital art), are generally tethered to their programmers because why not receive credit for honest work? But Twitter bots used for more nefarious or dubious means hide behind anonymity. Like hate-filled graffiti on a wall left under
the cover of darkness, trollbots, for example, are created and set loose to cause disruption, anguish, or other negative ends as part of a strategy to divide and play the part of provocateur. The Peñabots function under complete anonymity, and the only accountability they’ve received is due to meticulous efforts by researchers such as Alberto Escorcia and Erin Gallagher. Other Twitter botnets that operate under negative circumstances are sometimes brought into the open because of advocates of a healthy Twittersphere that protects real discourse while ousting trollbots, spambots, and fake news disseminators. These people are performing the labors of Twitter. Wikipedia, however, relies on its crowdsourcing users and editors to perform the labors of maintaining bots for maintenance, upkeep, and efficiency.

The Facebook [loveMachine] operated under a false name (Juego Requiem), and it continually had to be refined in order to avoid the Facebook algorithms that sought to remove unauthorized automated content. The accountability of the programmer was held at arm’s length, with few repercussions. But Deswaef grew uneasy when he realized that his [loveMachine] was, in a sense, curating a list of gay men in Tel Aviv and another in Islamabad. He knew that if his account were to be compromised by the Pakistani government, or if another user collected his data on the gay men in his network who lived in and around Islamabad, these men could theoretically be placed onto a list and charged with crimes. He was collating a sheet of data that could be abused by the Pakistani police. He was the only person who could shut down the algorithm in order to protect the identities of the men.
A RETURN TO THE MORAL GOOD THROUGH DIGITAL LITERACY

We are living in a post-truth world, where lies and false narratives are strengthened by the same types of algorithms that also aid and assist us with menial, digital tasks. In our quest to reach out into the digital frontier, we share our private information online as we attempt to build communities with one another. Social media platforms like Facebook and Twitter commodify our labors, turning them over to marketing and advertising agencies. It seems that every collation of digital data can be bundled and abused, weaponized, or turned into capital. How, then, are we to make a return to the moral good—taking back our privacy, restoring truth to the public sphere, valuing credibility, and engaging in more civil discourse?

I mentioned at the beginning of this chapter that digital literacy is somewhat of a disappearing art, as more and more computer processes disappear behind application interfaces. Like America’s industrial turn from and abandonment of the self-sufficient farm and the distancing of ourselves from direct food sources for the sake of convenience, there is an exodus away from the inner workings of code and toward a system of interfaces that hide the very processes that define our discourse and interactions. Instead of seeking out information in a card catalogue in a library or among many books within a section, we now allow YouTube to predict which videos will autoplay in our library, or what news Facebook will prompt us with, or with whom we should connect on Twitter. The labor of face-to-face discourse is replaced with the Facebook “like” button. No longer needed are the strategies for memorizing words that are difficult or tricky to spell now that autocorrect and spell check can do the work.
I suggest that we return to some of our digital roots through digital composition pedagogy, for starters—teaching first-year composition students not only how to think visually, rhetorically, and critically when writing their expository and persuasive essays but how to think in terms of digital privacy, understanding disinformation, and the agency-stealing properties of social media algorithms. Instruction of rhetorical strategies must include how algorithms are weaponized against truth and how to use algorithms to strengthen communication methods. Instead of teaching students the benefits of archiving data online, faculty should be teaching the persuasive power of building a bot that engages others with information that aids the collective good. The art of programming simple bots within a framework of accountability ought to be just as important as teaching students about plagiarism and ethical writing. Students should be taught that algorithmic expressiveness is not necessarily a leap forward, no matter the novelty, and that we should remain the masters of our own creations. After all, has it not been ingrained that we each write our names at the tops of our own papers?


