DHQ: Digital Humanities Quarterly

2016
Volume 10 Number 4

Attention Ecology: Trend Circulation and the Virality Threshold

Nicholas M Van Horn <nvanhorn_at_capital_dot_edu>, Capital University
Aaron Beveridge <aaronbev79_at_gmail_dot_com>, University of Florida
Sean Morey <smorey_at_utk_dot_edu>, University of Tennessee, Knoxville

Abstract

This article demonstrates the use of data mining methodologies for the study and research of social media in the digital humanities. Drawing from recent convergences in writing, rhetoric, and DH research, this article investigates how trends operate within complex networks. Through a study of trend data mined from Twitter, this article suggests the possibility of identifying a *virality threshold* for Twitter trends, and the possibility that such a threshold has broader implications for attention ecology research in the digital humanities. This article builds on the theories of Jacques Derrida, Richard Lanham, and Sidney Dobrin to suggest new theories and methodologies for understanding how attention operates within complex media ecologies at a macroscopic level. While many various theories and methods have investigated writing, rhetoric, and digital media at the microscopic level, this article contends that a complimentary macroscopic approach is needed to further investigate how attention functions for network culture.

The phrase “this is trending” is so common, so deceptively innocent, that it often passes from speaker to listener without 1
an opportunity to interrogate its validity. Nearly every broadcast of cable news, sports journalism, or entertainment media relies on this phrase to justify the attention paid to a particular event or topic of interest. For many forms of digital rhetoric, having “this is trending” attached to a digital artifact is, in and of itself, the desired end goal for producing those artifacts — millions, and possibly, billions of visits, views, likes, or shares. This phrase quantifies attention and repackages it as a statement of value. Indeed, “this is trending” is no mere economics of attention. This phrase suggests that the object it describes has, during the process of attention accretion, become more than its accounting — the trending object acquires value. If attention somehow creates value for trending objects, then this process becomes even more powerful for digital artifacts that hyper-circulate at a viral level.

Such questions of circulation and virality are of interest within writing studies, but writing-related fields are nascent in the 2
use, approach, and methods related to the digital humanities. Recent conversations in writing studies, rhetoric, and composition have attempted to predict and understand how these fields might pursue research questions made possible by digital humanities tools, methods, and resources, or, at least, better understand where these fields intersect with the digital humanities. For instance, in their collection *Rhetoric and the Digital Humanities*, Jim Ridolfo and William Hart-Davidson write that although writing studies has a rich tradition in performing analysis with digital tools, building new tools, and securing National Endowments for the Humanities (NEH) Office of Digital Humanities (ODH) Digital Humanities Start-Up grants, there has been little understanding of how these fields intersect: “Despite the lengthy history of the term *digital humanities* outside rhetoric and writing studies, inside the field of rhetoric and writing studies there has been, up until recently, very limited mention of it” [Ridolfo and Hart-Davidson 2015, 2]. In addition, many DHers within English departments might be unfamiliar with how writing studies has made strides toward this field. Ridolfo and Hart-Davidson argue that this is, in part, due to the field’s concern with the teaching of writing within composition and professional writing classes, this concern mostly focusing on how such classes might adopt computers and other technology as part of its pedagogy. However, as writing studies moves away from researching writing primarily in terms of students and pedagogy, toward researching writing without subjects, digital humanities tools provide even more opportunity for looking at large data sets of writing that are not tied to questions of student agency or

learning outcomes.

Despite this turn away from writing subjects, the concept of audience remains central to research in rhetoric and writing studies. But in the era of Big Data and information overflow, it is increasingly difficult to identify an audience. While recent work with the canon of delivery, such as Collin Brooke's work in *Lingua Fracta* [Brooke 2009] or Jim Ridolfo and Dànielle Nicole DeVoss's concept of "rhetorical velocity" [Ridolfo and DeVoss 2009] attempt to account for audience reception via electronic delivery and the digital circulation of multimodal writing, these theories often focus on audience-as-user, how audience members, as individuals, receive and then re-circulate writing toward other audiences. Such theories are useful in thinking through how an audience remixes and re-delivers writing, but do not address the larger scales across which digital writing now circulates. 3

In his work *Postcomposition*, Sidney I. Dobrin argues that writing studies needs to push its investigations into these larger scales, into complex understandings of hyper-circulation and writing-as-system [Dobrin 2011]. Dobrin's argument parallels those made about literature studies, lately by Matthew Jockers, who proposes a methodology of macroanalysis to account for how patterns or trends emerge across large data sets of literary texts [Jockers 2013]. The macroscopic perspective has often been understood in terms of "distance" and "artificial constructs" within the humanities — what Franco Moretti has termed "distance reading." As Moretti explains, "abstraction is not an end in itself, but a way to widen the domain of the literary historian, and enrich its internal problematic" [Moretti 2007, 1]. 4

Toward a macroanalysis of writing, this article performs a macroanalysis of Twitter in order to extend a theory of hyper-circulation that can help explain how macro-audiences engage with digital writing technologies. This article builds on Dobrin's network theory by looking at how the concept of attention provides a macroscopic understanding of audience in complex networks of mediation. If many of the attempts underway to rewrite rhetorical concepts, such as delivery, focus on the microscopic invention and delivery of digital texts (the agency of individual users/writers), then attention ecology is the attempt to understand the macroscopic, aggregate trends occurring within social media as a way of identifying an audience in complex networks of mediation. Ultimately, the authors identify the term *virality threshold* as a concept that may be useful for future researchers interested in macroscopic analyses of writing and audience. 5

Certainly, the use of contagion metaphors (such as virality) to understand the spread of ideas and information among social groups is not new. Likewise, the threshold concept is also not a new addition to the discourse of idea circulation. Prior to the Internet and the subsequent invention of contemporary social networks, the terms *contagion* and *threshold* were core concepts for Diffusion of Innovations (DoI) theory. Like the digital humanities, DoI was necessarily interdisciplinary as it sought to understand how new ideas were spread and adopted within various areas such as rural sociology, communication, education, marketing, general sociology, anthropology, economics, political science, agriculture, psychology, statistics, engineering, and others [Rogers 1983, 53]. Everett M. Rogers' *Diffusion of Innovations*, originally written in 1962, is credited with collating a diverse array of ideas from various disciplines to better understand "the process by which an innovation is communicated through certain channels over time among the members of a social system" [Rogers 1983, 6]. Rogers' initial description of DoI theory focuses on the relationship between the concept of adoption and the idea of the threshold. Certainly, there are similarities between Rogers' use of the term adoption and this article's use of the term attention. For example, Rogers uses the example of people considering whether or not to join a riot or protest. In this example, an individual's threshold is determined by the number of other people a person needs to observe taking part in the protest before they too will join. The concept of threshold, for DoI, thus refers to the amount of other adoptions a particular person needs to observe before they will adopt an innovation or action themselves. Someone who may be described as an early adopter will have a low threshold for the number of people they need to observe before choosing to take part in the protest. However, a person who needs to observe a large amount of people taking part in the protest would be described as having a high threshold according to Rogers — and therefore, they would be a defined as a late adopter. 6

DoI theory remains active and relevant today for the study of social networks. Thomas W. Valente's *Network Models of the Diffusion of Innovations* applies DoI theory to network analysis, and recent scholarship that applies DoI to the study of Twitter and other social networks builds on his work. Valente defines network analysis as "a technique used to analyze the pattern of interpersonal communication in a social system by determining who talks to whom" [Valente 7

1995, 2]. While this article draws upon the interdisciplinary history and terminology made available by DoI theory, it also diverges from DoI theory in two key ways. First, recent theoretical work in writing studies puts pressure on network theories that are non-ecological — meaning, network theories that focus on individual users within individual networks, rather than focusing on the complex interactions of multiple networks. Therefore, this article replaces the user-as-node (person to person transmission of ideas within an individual network) with the location-as-node model (the circulation of ideas among and within multiple locations — among and within multiple networks). In contrast to user-as-node approaches, this article looks neither at users (or the adoption of ideas or innovations by users) nor does this article focus on any specific trend. Instead, this study considers only that a trend occurred, and that it was circulated among locations.

As a result of the focus on locations rather than users (or subjects, or agents), the term *adoption* from DoI theory is replaced with the term *attention* in this study — this is the second divergence from DoI theory. We argue that one way of analyzing how hyper-circulation occurs is by searching for indicators of attention and the degree to which such attention escalates. Using data science methodologies, we introduce a preliminary operational definition of "virality threshold" as one concept for beginning to think about macroscopic attention for writing studies and digital humanities research.[1] If attention is to provide an effective framework for thinking about networked writing and macroscopic audiences, then scholars must begin to carefully investigate the terminology used within such a framework. The term "viral" is often used to indicate that a text, topic, or digital artifact has hyper-circulated among multiple networks of mediation and gained a significant amount of attention. But what is the difference between something that is merely trending and something that has become viral? Why might such determinations be critical to writing studies and digital humanities research? 8

Trends themselves have become the basis for news reports — CNN, ESPN, Fox, MSNBC — all of these media networks now use the phrase "this is trending" as the basis for discussing a topic or bringing something to viewers' attention. This phenomenon is not an altogether new occurrence. Viewer polls — asking viewers to give an opinion or respond to a questionnaire — have provided a similar form of feedback for television and print journalism. Exit polls in elections also provide similar feedback, because voters provide immediate feedback about their choices or opinions while the voting process is still underway. The immediate question raised by trend feedback is whether trends are popular (and sometimes viral) because of the content of the trends, or whether some trends become popular because they gain enough momentum and inertia to continue acquiring attention as circulation increases. Does trend feedback become a veritable snowball effect that may, at times, lead to virality? The obvious answer to this question is yes, trend feedback is no doubt occurring. People care about what is popular or trendy merely because those things are popular and trendy — even if viewers disagree with the actual content of their feeds and networks, the question of whether content merits trending in the first place seems to now be secondary to the question of whether content has gained attention. This is why operational definitions for terms like *trend*, *virality*, and *attention* are crucial for understanding and critiquing network culture. 9

A Differential Typology of Iteration: Theory

Thomas H. Davenport and John C. Beck — in their management studies book, *The Attention Economy* — use the phrase "human bandwidth" to discuss the limited attention humans have in the face of an overflowing feed of text, image, and video — all of which are constantly vying for a meaningful portion of the available "bandwidth" of "human" attention [Davenport and Beck 2002]. This term, *attention*, has become useful to rhetoric and writing studies as well. Richard A. Lanham understands human attention in similar economic terms to Davenport and Beck's, but rather than seeing attention as an issue of human management, Lanham's attention framework operates with the assumption that the "devices that regulate attention are stylistic devices" [Lanham 2007]. Lanham's *The Economics of Attention* explains that while information is widely available, human attention is a scarce resource to be carefully allocated through broad rhetorical applications of *style* — one of the five canons of rhetoric. Attention is a crucial term for research into digital media because an attentive audience can no longer be assumed — the mind numbing noise of information overflow is the new norm. Constantly updated social network feeds, 24 hour news networks, movies, blogs, ebooks, digital media, print media, ubiquitous screens, and always available interfaces — all fighting for attention — try to shock, entertain, or scare users into viewing for as long as possible. 10

In fact, attention itself acquires attention — the “tipping point” as it has commonly been called of late — when humans notice that a significant number of other humans are paying attention to something specific. In DoI theory, this is called “critical mass.” Critical mass, in innovation diffusion theory, is achieved when “the minimum number of participants needed to sustain a collective activity” has been reached. However, the problem with critical mass, as Thomas W. Valente notes, is that often critical mass is achieved shortly after a disproportionately small number of people within a social system have adopted an innovation. Like a school of fish avoiding a predator when only a fraction of the fish actually see the predator with their own eyes, the rest of the group moves in coordinated association with the group as a whole. The changes in direction are chaotic on a micro level, but it is a patterned and organized flow of movement from a macro perspective. 11

Sidney I. Dobrin argues in *Postcomposition* that purely rhetorical approaches to complex media ecologies — in this case, the ecology of attention — may misidentify “pauses in the system” or “delays in fluctuation” as more traditional causal/rhetorical relationships between parts and wholes [Dobrin 2011, 141–142]. As Dobrin argues, “the whole of writing can never be explained by way of its parts but instead through the properties of its interconnectedness” [Dobrin 2011, 144]. Therefore, the question of trend content, as the rhetorical question of exigence, is only one aspect of a much broader ecology of attention, and focusing on this question in isolation fails to account for broader ecological and technological issues pertaining to digital media. That is not to say that the questions regarding trend content ought to be ignored or drastically reduced in their importance, but limiting analyses of mediation to relationships between human attention and “content” will no doubt overlook key cybernetic and non-human technological factors that may account for a significant portion of the networked forces affecting what becomes a trend. As Dobrin argues, the human-as-node model of network research cannot currently explain the complex fluctuating-flow of writing that occurs during “hyper-circulation” and “network saturation” [Dobrin 2011, 181–186]. Attention ecology, therefore, works to understand how complex networks of mediation restructure the reception of compositions through hyperactivity and continuously-fluctuating flow. 12

When attention acquires or redirects attention, this functions as a feedback loop within all the various interfaces and mediums vying for attention. In DoI theory, one explanation for this phenomenon is known as “weak ties.” As Valente explains, 13

> weak ties facilitates critical mass because it permits innovation to be spread to otherwise unconnected subgroups. Size, centralization, and characteristics of early adopters may be immaterial in their effect on critical mass if weak ties are not present in the network to insure that subgroups share information with one another. [Valente 1995, 85]

Dobrin’s *Postcomposition*, however, resists the weak ties theory in order to “abandon” the notion that total-connectedness through weak ties of all subgroups, within a particular network or among multiple networks, explains attention feedback and hyper-circulation [Dobrin 2011, 152]. As Dobrin explains, 14

> we may be able to achieve a more ecologically based view of writing by examining the spaces of strange loops as dynamic spaces in which systems establish internal order and as locations in which interrelations take hold. Given that strange loops contribute to the openness of autopoietic systems, we must concede that it is in the strange loop space where interrelations with other systems are most likely to occur. [Dobrin 2011, 166]

There are many parallels between Dobrin’s description of strange loops and the function of weak ties theory in DoI, but the crucial difference is the shift in focus from the subject (user-as-node) to location and place. 15

Certainly, it is easy to understand attention at a personal or experiential level, but attempting to explain how attention may be regulated, redirected, or acquired within complex systems of mediation requires a more careful and well-tested definition. The primary reason this study does not limit attention to “humans” (whatever that may mean) is because a composition or inscription may jump from a human network to a nonhuman network (not that these are ever totally inseparable) due to the reiteration of bot readers/writers and the activity of weak artificial “intelligences” already at work among the many available networks of mediation. The more intensely studied canons of rhetoric — invention, 16 arrangement, and style — may have very little to do with what gains attention, and certainly even less of an effect on what “goes viral.” Many of our art, literature, and film studies colleagues often lament the artistic quality and cultural benefit of what becomes popular in film, music, literature, and art. Part of this can be explained by massive marketing campaigns and corporate agendas regarding media and entertainment choices available to consumers, but a large amount of what becomes popular on a “grassroots” or “viral” level of attention is just as difficult for critics to applaud as quality art or meaningful cultural contribution. Although Lanham’s preferred rhetorical canon is *style*, in an attention ecology, temporal and spatial notions associated with the canon of *delivery* are necessarily pushed to the forefront.[2]

At the most basic level, operational definitions depend on identifying methods of measurement for specific phenomenon. The phenomenon we attempt to measure in this article is the *iteration* of writing in social media as a way of understanding how attention is redirected or acquired within complex networks. Writing in social media is not limited to the modality of unicode text that appears on computer screens, but writing is understood in the broader sense of any digital modality (text, image, audio, video, etc.) that may be used to make meaning. The basic unit of analysis for this article — iteration — is taken from Jacques Derrida’s work in *Signature Event Context* [Derrida 1988]. In *Signature*, Derrida uses the concept of iteration — the movement of writing through time and space — to undercut simplistic notions of communication derived from speech act theory. For Derrida, iteration is “the spacing” that (re)moves writing from its original determinate “context,” and this movement or “force of rupture” is what “constitutes writing.” Writing moves through time (iterates) because of its ability to repeat and reproduce an “alterity” of its prior form in “absence” of any original intent of meaning or content [Derrida 1988, 9]. This removal, or spacing, occurs because writing reiterates its collection of signifiers at a later point in time. Social networks provide the possibility of locating iterations of writing across time and space. Of course, this possibility is not unique to social networks, but the availability of temporal and location-based data, provided by social media networks, makes this study more feasible through expedient access to large amounts of networked writing.[3] 17

This article builds on what Derrida calls a “typology” of iteration: “Rather than oppose citation or iteration to the noniteration of an event, one ought to construct a differential typology of forms of iteration, assuming that such a project is tenable and can result in an exhaustive program, a question I hold in abeyance here” [Derrida 1988, 18]. Raúl Sánchez, in *The Function of Theory in Composition Studies*, argues that the “exhaustive program” described by Derrida suggests the possibility for some kind of future “empirical investigation, resulting in a mapping or organizing of the various ‘forms of iteration’ one might find” [Sánchez 2006, 36]. A typology is a classification or categorization according to general types. A differential typology integrates the various classes or categories of which it is comprised. Therefore, the differential typology utilized in this study is as follows: iterative writing is writing that moves, but this movement is a derivative of two other basic components: time and space. While time and space are relative to one another, time is iteration’s constant. In other words, even if writing does not move spatially (from one space, place, or location to another), it is always reiterating temporally (it is always moving through time). 18

Movement through time is a necessary condition of writing, while movement through space is not, although writing, in practice, usually moves spatially as well. Even prior to the Gutenberg press, monks passed exemplars from monastery to monastery for other scribes to copy in order to expand the available texts in their libraries. The Internet and digital technologies have not substantially changed the functional integration of writing and movement — writing has never existed in a temporal or spatial vacuum. However, spatial movement has not always been an *inherent* component of the various technologies available for communication. In digital and electronic media, movement and circulation among users and their respective locations are often the driving force behind the development of new writing networks. Codex technologies (often, paper books) require something else to move them from one space to another, and historically this role was often provided by publishers, book stores, and libraries. Technologies like Twitter, on the other hand, were developed to move and circulate writing among a network of users — Facebook, Instagram, YouTube, and Vine were all built with this core functionality in mind. This intentional circulation to encourage more circulation is what Jim Ridolfo and Dànielle Nicole DeVoss call “rhetorical velocity,” where new combinations of digital delivery and multimodal production form the basis for understanding writing as it relates to movement and circulation [Ridolfo and DeVoss 2009]. Ridolfo and DeVoss focus on microscopic deliveries, what we are here calling user-as-node or the individual choices made by single authors, and this is categorically different from the macroscopic analyses of delivery and circulation we attempt 19 with this article. Microscopic user-as-node analyses are no doubt still crucial to ongoing research in rhetoric and writing studies, and to the digital humanities more broadly conceived. However, as humanities researchers continue to push their investigations of writing and circulation into theories of complex networks and network culture, ongoing work is needed to expand digital research methodologies beyond the scope provided by rhetorical and pedagogical methods.

Therefore, this article seeks to investigate one of the lingering questions regarding media ecology and complex networks from Dobrin's *Postcomposition*: 20

> For a number of reasons, including the desire not to make "scientific" the ecology of ecocomposition postcomposition beyond perhaps a grammatological science, I am not going to distinguish between micro and macro approaches in establishing complex ecological theories of writing, though there may be potential for doing so in future theoretical work. [Dobrin 2011, 142]

This "potential" that Dobrin describes for distinguishing between "micro and macro" approaches to "complex ecological theories of writing" is the basis for the following differential typology of iteration. Thus, we define the microscopic approach in complex networks as attempting to understand individual users within specific networks. This is what we refer to as the user-as-node concept of network research. This reductive concept understands network behavior as affecting the choices of the user within the network — often leading to reductive cause/effect relationships between users and the information that flows within a single network. 21

Alternately, the macroscopic approach attempts to understand aggregate activities of groups of users across an interactive ecology of multiple interactive networks. We see this as the environment that allows for Dobrin's hyper-circulation and network saturation. We know that most users, if not all, interact within multiple networks of mediation. A single user, for example, will watch television while writing on Twitter and posting images to Instagram with their mobile phone, and monitoring Facebook on their laptop — one user, three screens, and four networks. If we analyzed this particular *user* only within Twitter, for example, then that user's tweet about a particular television show may appear to be motivated (caused) by another Twitter user who previously tweeted about the same show, when, in the case of this example, the tweet about the television show may have been motivated instead by a Facebook post or by the show itself. Or, maybe the user received a text message from a friend who was watching the show at the same time (a fifth network). The problem with the user-as-node in a single network model is that it often ignores this broader media ecology of multiple interactive networks, and it too easily allows for simplistic cause/effect claims about circulation and determinate reductions of agency. Accounting for users as nodes within complex ecologies may be theoretically possible, but it is likely a pragmatic impossibility when considering that any user's actual position within multiple interactive networks may be too indeterminate and ever-shifting to predict a future position. 22

The preface to the ten volume *Encyclopedia of Complexity and Systems Science* defines complex systems as "systems that comprise many interacting parts with the ability to generate a new quality of collective behavior through self-organization, e.g. the spontaneous formation of temporal, spatial or functional structures" [Meyers 2009, iii]. Individual social networks, like Twitter for example, certainly fulfill the requirements for being defined as complex networks: "They are therefore adaptive as they evolve and may contain self-driving feedback loops" [Meyers 2009, iii]. However, the user-as-node within Twitter (within a single complex network) does not account for the cybernetic post-subject within a complex writing *ecology*. That is, the user-as-node model typically focuses on a single user in a single network and cannot adequately account for the information flow between networks. Complex ecologies of writing must work to understand, as Dobrin explains, that "the whole of writing can never be explained by way of its parts but instead through the properties of its interconnectedness" [Dobrin 2011, 144]. This does not mean that we do not look at parts and wholes, but we do so with the theoretical understanding that complex networks are so dynamic and ever-changing that there is no practical way to ever recreate a user's cybernetic position or to capture an all-encompassing macroscopic ecology. And even if we could recreate or map that positioning, it would only represent that single moment in time — a "pause in flow" as Dobrin calls it — and it would likely tell us very little about that user's future position and attention. As the *Encyclopedia of Complexity* explains, "complex systems are much more than a sum of their parts. Complex systems are often characterized as having extreme sensitivity to initial conditions as well as emergent behavior that are not readily predictable or even completely deterministic" [Meyers 2009, iii]. 23

Therefore, if identifying the flow of information for individual users within individual networks represents the microscopic perspective for media ecology, the macroscopic perspective looks at how users are aggregated or moved by large trends of information flow. In other words, the microscopic sees the user as directing the flow (what is often thought of as agency), and the macroscopic sees groups of users as following trends or hyper-circulation (being redirected by the flow). Certainly, for writing studies and digital humanities scholars who are interested in affecting audiences within a singular network, Diffusion of Innovation theory applications for Twitter research (also applicable to other individual networks) continue to produce useful models for understanding how ideas circulate among specific social groups within an individual network (see, for example, [Lee, Kwak, Park, and Moon 2010], [Wu, Hofman, Mason, and Watts 2011], [Guille and Hacid 2012]. However, the work forwarded in this article looks for ways to begin moving beyond such microscopic studies, and start moving toward macroscopic social network research in its own right (rather than merely using macroscopic analyses to reinforce micro/user-as-node research). 24

To return to our earlier metaphor, while fish in a school all act individually based on other fish around them, we can also learn about their movements by watching the school as a whole. However, in terms of complex networks, no grand unifying theory yet exists to account for the integrated operations between micro and macro levels of research. We propose no grand unifying theory in this article, but we also do not see the two as conflicting or contradictory. Much effort and research in rhetoric and writing studies has been given to questions of agency and individual users' (writers) ability to direct the delivery of writing — the theory of "rhetorical velocity" is a great example of this. We see no inherent problem with microscopic research — it is important pedagogical work, and it ought to continue. Rather, our research attempts to make the macroscopic (large scale, aggregate) perspective of writing useful for media ecology and writing studies as well: ours is an inclusive agenda, and an attempt to expand the available methods for studying networked writing. 25

The type of iteration at the core of our research is what is commonly called a "trend." As mentioned above, trends identify large scale, aggregate information flow. Trends occur within social media technologies when users repeat similar words, or hashtags (#), over the same period of time, or when large groupings of similar words or phrases iterate together. The period of time and systems used to identify trends are always arbitrary. There is no positivist basis or claims of capturing social reality forwarded by this research. Data science, in many of its variations, may be understood as postmodern in the ways it conducts its research — the scopes constructed are always folding back on themselves, always limited by the self-identifying archives they create, and always self-modified to tentatively answer exploratory questions. As Cary Wolfe explains in *What is Posthumanism?*, building on the work of Niklas Luhmann, the combination of systems theory and Derridean theories of writing provide the basis for "reconstruction" [Wolfe 2009, 8]. While Wolfe's work with critical animal studies takes a far different trajectory than does our own research, the same underlying theory is applicable to this study. The reconstruction of trend flow in this article — of macroscopic information flow — is limited by the scope that allows or recreates its very possibility. The acknowledgement and study of such limitations is what makes data science — the operational methodology of measurement deployed in this article — amenable for humanities scholars who have long theorized the generative possibilities of the archive. 26

Writing Data: Methods

In order to conduct a macro-analysis, we decided to focus on micro-blogging, specifically the social media site Twitter. Although each tweet is only 140 characters, the total number of tweets provide the opportunity to test a massive data set. In addition, since tweets are often shared, linked, and tagged, the larger Twitterverse provides a rich source of information in which to understand how writing circulates as writing. 27

Although no history of Twitter is "long," much research has been done in order to determine how Twitter may be used for a variety of purposes and a variety of users. For example, most research on Twitter falls into at least six methods: content analysis, topic modeling, clustering, sentiment analysis, opinion mining, and audience exposure analysis. Content analysis analyzes Twitter for a variety of purposes, but mainly to research what users tweet and for what purposes. Topic modeling looks for specific topics or conversations which are then used to model other phenomena or events. For example, Michael J. Paul and Mark Dredze have analyzed tweets related to public health and shown that such data can provide quantitative correlations with public health data, suggesting that Twitter can be applied to public 28 health research [Paul and Dredze 2011]. Clustering provides a methodology for showing relationships or natural groupings between Twitter users, hashtags, or specific terms: who retweets or follows whom, how hashtags relate to each other, etc., in order to show how information circulates on small scales and how users relate to each other. Sentiment analysis [Bifet and Frank 2010] attempts to determine the feelings of users toward a particular story or subject, often toward ends such as brand management. Relatedly, opinion mining attempts to search Twitter for the opinions of users on particular topics, which may range across a continuum on the level of emotion such opinions contain (in other words, some opinions may be more measured or more reactionary). While such research can be useful, it does little to show macroscopic trends across a network (or between networks) as a whole. Coming closest to this study, audience exposure attempts to determine the number of users who have seen an individual tweet. While the number of times a tweet is retweeted is relatively easy to find, the number of actual users who have viewed the tweet is not. Tom Emerson, Rishab Ghosh, and Eddie Smith offer one method through their investigation into a tweet by the Bin Laden Live Tweeter [Emerson et al. 2012], but, again, such a method still only accounts for a single tweet, a single user-as-node, and does not attempt data analysis on the scale required to best understand hyper-circulation and how the structure of the network as a whole (and its connections to other networks) influence how hyper-circulation occurs.

While there are many websites that assist in accessing social network data, most of the applications and websites that 29
provide access to archived or streaming data charge a fee, provide limited or already processed data, and usually are focused entirely on brand management and marketing research. This happens primarily for two reasons: (1) it is common practice for social networks and search engines to license their data for resale, providing an additional source of revenue, and (2) other companies are willing to pay for this data because it provides a specialized form of brand management and marketing research that cannot be acquired elsewhere. While some of the current web applications that provide access to social network data have started to “grant” free access to a limited number of academic institutions [Krikorian 2014], depending on the types of research that may be completed with the given data, the built-in commerce/marketing focus of these web archives limits the scope of research questions that may be asked. Therefore, while the cost of access to data continues to be a significant gate-keeper for social network research, the larger issues are the questions raised by the types of restrictions placed on data because of a systematic preference for marketing and commerce related research.

Because of these limitations, we opted for an open-source academic research software called MassMine, as it accesses 30
and collects data from social network APIs.[4] Many of the top social networks, such as Twitter and Facebook, provide an API (application programming interface) for application and software development. Access to APIs is generally provided free of charge in order to encourage software developers to write mobile apps and to provide the infrastructure for companies to create web applications and supplementary services that are not provided by the networks themselves. For example, the photo sharing application/network Instagram initially gained momentum as an API application that provided the ability to easily take, edit, and post photos to other social networks. Now, Instagram is its own viable social network and is owned by Facebook, but Instagram’s initial function of interacting with user data through access to an API allowed it to develop its primary photo sharing feature. APIs are useful in that they can be re-purposed for social media research, using this access to collect and analyze data on the network, its users, and the content circulated therein.

Despite this accessibility, APIs are not beyond reproach for they subject the researcher to their own limitations. We 31
believe this trade-off results in a net benefit for the academic investigator as the limitations are based more so on bandwidth restrictions and less on prescribed end-uses for the data. Network access through APIs is typically limited by the constraints of big data itself. Put simply, existing network infrastructure cannot support unfettered access to social network databases. Indeed, API functionality for large-scale services like Twitter is often rolled out incrementally as internal engineers solve the problems of big data one step ahead of users. The types of limitations created by an API are different than the access issues, costs, and research limitations of data types that are common to existing commerce and marketing tools. In other words, APIs do not generally limit the kinds of data or what you can do with the data, nor do APIs charge for data, but they do limit how often data access may occur (because of bandwidth restrictions) and how much (or how far back in time) data may be acquired each time an API is accessed. Anyone with the right programming skills may access publicly available data through APIs, thus eliminating the immediate cost

barriers involved in using APIs. However, the bandwidth restrictions do create access issues. For example, historical data beyond a certain point in time may not be accessible with many social network APIs, but if a researcher knows what they want to research ahead of time (i.e., they have pre-established a well-formed research question), then it is possible to mine data on a regular basis and create a large historical archive over a period of time. MassMine, as an open-source research tool created for academics, allows such pre-determined archiving of data by providing full API data access to users unaccustomed to programming low-level software solutions.

The data collected for this study was pulled from Twitter's API over seventy-four days. Analysis and visualizations were conducted using both MassMine and the open source R statistical computing language. During the period of time in which we were pulling data from Twitter, we collected data on over 17,343 unique trends, including the Christmas, New Year's, and Valentine's Day holidays, the deaths of numerous celebrities and culturally significant figures, and Super Bowl Forty-Eight. Our measures of interest in the rise and fall of trends across geographic regions emphasized the need for temporal precision. MassMine was set to retrieve and archive the top ten trends every five minutes from nine different locations across the US, and an additional top ten trends for the entire US as a whole. The choice of five minutes represents the maximum refresh rate of Twitter's trend data. 32

Geolocation of information on our data was made possible via data filtering, specified as Yahoo Where On Earth IDentifier (WOEID) codes, which allowed us to track the location of trends across space and time. Twitter's API allows for trends to be identified as either (1) keywords specified as literal #hashtags (for example: #superbowl) that are defined by users, or (2) trends that are defined according to the most common phrases appearing in user statuses. We decided to pull trends based on both, rather than limiting our analyses to prescribed #hashtag trends. This approach ensured that we identified organically-arising trends, in addition to more explicit #hashtag events. Moreover, sometimes a singular trend may use multiple #hashtags. For example, there were numerous #hashtags surrounding the death of Nelson Mandela (e.g., "Nelson Mandela"; "#MandelaMemorial"; "#NelsonMandela"; "Mandela"; "#MandelaFuneral"), but Mandela's death could be interpreted as a singular trending topic. To be sure, both kinds of data were collected and jointly analyzed. 33

Trends provide one particular way of understanding attention, but we want to be careful not to equate trends with attention. In fact, one of our current limitations is that we are only looking at trends as they appear on Twitter. Ideally, attention ecology would look at data from multiple social networks, search trends from search engines, and other information that may help to identify attention — such as data mining the tickers that appear at the bottom of twenty-four hour news networks as a way of identifying what "news" is given attention at various times of the day, looking at ratings numbers and movie box office returns to get an idea of how entertainment events may be accounted for, and certainly finding ways to understand how print media and oral culture (word-of-mouth) continue to be a significant aspect of attention ecology. There are potentially infinite ways of reconstructing an ecological understanding of how attention may function with complex networks of mediation, but when we account for place and time the problem becomes finite. It is in this way that Twitter trends provide a useful starting point for understanding how to attach the concept of attention to time and place. As more tools become available to humanities researchers studying digital media and complex networks, a broader macroscopic understanding of attention ecology will emerge. 34

Circulation Analytics: Results

Our interpretation of the analysis is deliberately narrow for two reasons: first, we are only accounting for trends within one particular network. Even though Twitter has almost thirty-seven million active monthly users in the US, it is unlikely that a significant number of those users restrict their use to Twitter and do not acquire/circulate similar information through other networks. Second, the temporal resolution of Twitter trends is necessarily fixed at some minimum span (currently, five minutes), as defining trends tacitly implies aggregating content across time. Further, within a unit of analysis repetition of content is tabulated, with the ten most active topics selected as "trending" by Twitter's ranking algorithm. This means that a trend only becomes available for data collection through API access once it has hit the top ten in a particular location. Conceivably, it is reasonable to assume that many trends reside at the eleventh or lower ranked position in numerous geographical locations before being revealed in the top ten of a particular location. Therefore it is problematic to understand the first appearance of a trend in Twitter's system as showing where a trend 35 begins or which locations seem to be the most influential in "causing" trends to "spread" to other locations. Another difficulty can arise due to the magnification of the scope of analysis. That is, trends that are initially below the level of detection at city-level could, in aggregate, reveal their circulation at the national level before appearing on any individual location's top ten. This can occur, for example, when network attention is broad and diffuse across a wide geographical region. Indeed, an analysis of our own data revealed that over 29% of the trends that occurred in all ten locations were identified at the national level first.

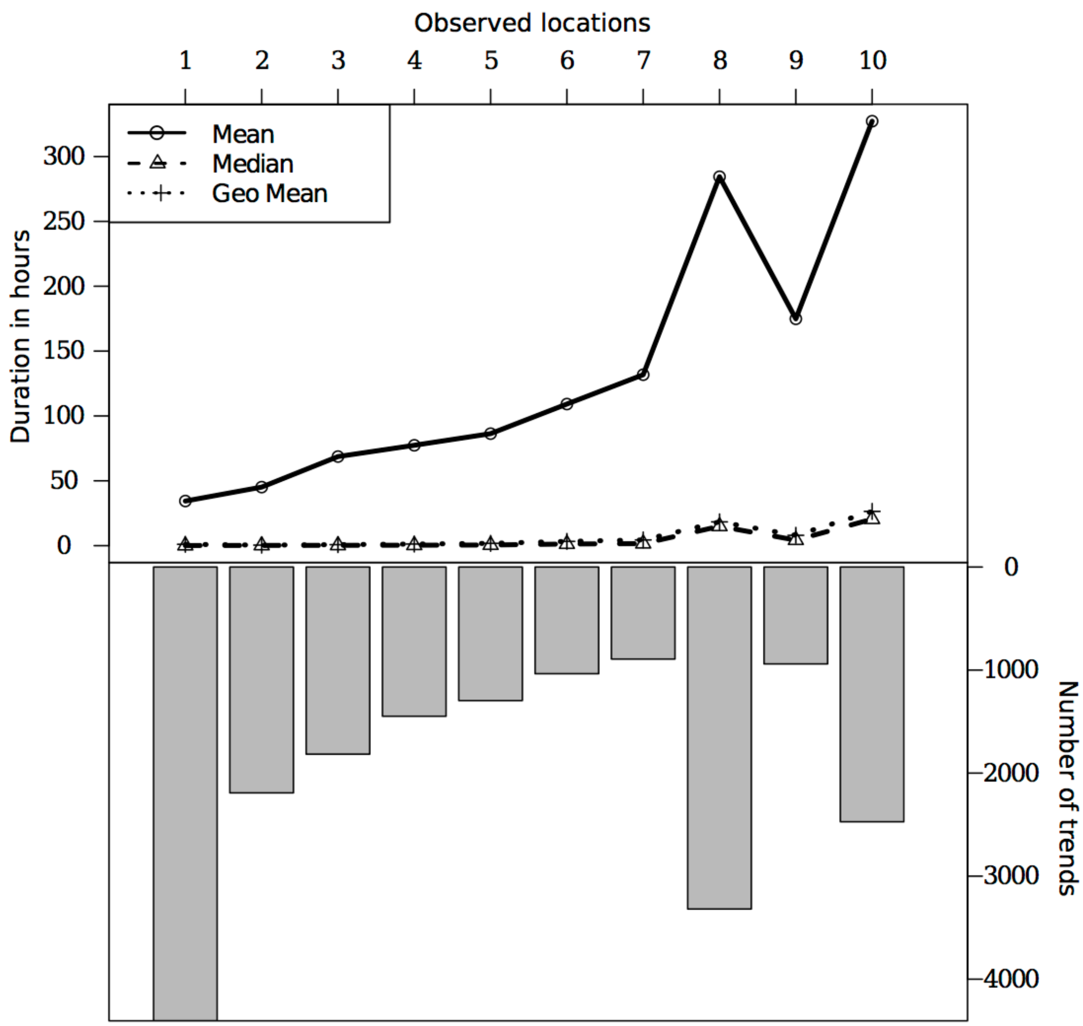

Figure 1. Average duration of trends (in hours) as a function of the number of locations. The number of trends included in each estimate is depicted under each point on the curve.

Our collection efforts constituted samples of 10 trends for each of ten geographical locations: Chicago, IL; Columbus, OH; Denver, CO; Houston, TX; Jacksonville, FL; Los Angeles, CA; New York, NY; Seattle, WA; Washington D.C.; and the United States. These locations were chosen because they were large cities with good geographical and cultural spread. The sample trends in these locations were identified every five minutes, for just over seventy-four days. At 288 samples per day, this amounts to 288×74×10×10≈2,131,200 records. Because trends are a spatial and temporal phenomenon, there is a great deal of redundancy across regions and time spans of analysis. Table 1 describes the breakdown of unique trends as identified across all samples. In total, 17,343 unique trends were observed across all locations, with decreasingly fewer topics reaching trending levels at multiple locations. For example, 14.3% of trends (2472) appeared in all ten locations. This 14.3% of trends is important because these are the trends that appeared in the top ten trends of every single location and the top ten trends of the broader US location, and thus can be more confidently compared to one another.

# of observed locations	# of trends	% of total trends
1	17343	100.0
2	12939	74.6
3	10747	62.0
4	8931	51.5
5	7483	43.1
6	6187	35.7
7	5152	29.7
8	4259	24.6
9	3412	19.7
10	2472	14.3

Table 1. Frequency of trends identified across one or more locations.

Table 2 summarizes trend duration at the group level, with group defined according to the number of locations at which individual trends were observed. For example, the median duration of trends that appeared at only one location across their entire lifespan was 0.08 hours (4.8 minutes), while trends that were identified at all ten locations had a duration of 20.43 hours. While this result might conform to reasonable expectations, it is not at all logically necessary. All measures of central tendency suggest an overall increase in the longevity of trending information corresponding to its geographic spread. However, the effect of spread is small in all estimates, with the exception of the mean, which is subject to large variations contingent on uncharacteristically large or small values. Figure 1, which shows average duration above the number of values used in each calculation, illustrates this. Fluctuations in duration estimates for positions 8-10 move in tandem with corresponding changes in the underlying counts, particularly for the mean. Still, an increase in the lifespan of a trending topic commensurate with its reiteration is fundamental to discussions and theories of virality. How then, can observations regarding the behavior of trends — topics that, by definition, are already popular within a network — lead to further understanding of the mechanisms and conditions behind the rise and fall of viral trends? The insensitive nature of our measure prohibits us from inferring the origin of trends, both spatially and temporally. As discussed above, trends only reach a level of identification after they've achieved top ten level popularity. 37

# of observed locations	Median	Mean	Geometric Mean	Max
1	0.08	34.38	1.04	1759
2	0.18	45.12	0.40	1713
3	0.25	68.69	0.78	1731
4	0.34	77.37	1.24	1761
5	0.50	86.33	1.76	1654
6	1.26	109.18	3.29	1681
7	1.56	131.79	4.41	1683
8	14.93	284.44	18.34	1726
9	4.28	174.87	7.97	1726
10	20.43	327.25	26.40	1776

Table 2. Group-level duration estimates (in hours) for trends occurring in one or more locations.

Despite the limitations inherent in trend analysis, we believe that it is possible to approach an analysis of trend circulation with those trends already identified as such. The risk arises from including sparse trends, defined as topics that appear in only a subset of measured locations. Such data cause systematically biased estimates of trend location, duration, and genesis when leveraging averaged, group-level investigations of this sort. For example, a trending topic 38 that persists for a long time only at the national level, but never at city-based levels, can grossly underestimate its apparent geographical spread, while simultaneously skewing the apparent duration of similarly grouped trends with much shorter lifespans. Likewise, a social phenomenon linked to a given city (e.g., trends unique to a state fair) that lasts for an extended period may exaggerate its apparent value in the greater national or social context.

To avoid such pitfalls, we conducted an analysis on the subset of trend topics that appeared, at some point in their lifespan, in all measured locations — nine cities and at the national level. This ensured that trends were not systematically binned for analysis based on geospatial restriction alone. Put clearly, in the following analyses, all trends appear at each level of analysis. This reduces the risk of spurious trend profiles due to uncharacteristically short or long duration, and/or narrow or diffuse movements during recirculation and reiteration. Duration distributions for these "ubiquitous" trends were notably shifted toward a greater length of time, largely due to the removal of single-instance trends that populated (11% of) the full data set. Median and mean duration among ubiquitous trends was 20.43 hours and 327.25 hours, respectively, compared to the corresponding duration of the full data set (median = 0.50 hours; mean = 111.75 hours). 39

Figure 2 provides an overview of group-level ubiquitous trend duration as a function of what we call geotemporal sequencing. Each curve corresponds to one of the ten measured locations, with national-level data depicted in red. The horizontal axis represents the temporal sequence, or order in which trends appeared at a given location. For example, the first point on the red curve depicts the average lifespan duration of trends that appeared first at the national level before appearing at any other location. By contrast, trends contained in the final point of the red curve were observed at every other measured location before they appeared at the national level. Critically, by restricting this analysis to so-called ubiquitous trends, which appeared at all locations by definition, each curve in Figure 2 is composed of the exact same data. The only difference is the order in which the data were grouped for analysis. This approach ensures that our comparisons reflect underlying patterns in recirculation/reiteration rather than artifacts of the particular data chosen for analysis. 40

Several interesting patterns are revealed by the geotemporal sequencing results. Most notably, the red curve (national-level trends) reveals a marked increase in duration contingent upon when a trend appears at the national level. This result is remarkably consistent across the entire geotemporal sequence. That is, trends that begin their journey at the national level have the shortest overall lifespan — shorter than trends beginning at every other city-level location (as evidenced by the fact that all gray curves begin higher than the red curve). Continuing along the red curve, a clear increase in duration is observed, one that remains higher than all other locations after the fifth position in the curve. Increase in mean duration as a function of order was found to have a significant linear trend. The coefficient of order was estimated to be approximately twenty-six hours ($p < .0002$). Thus, in our data, each additional shift to the right in the national-level curve reflected an increase in staying-power of just over a day's length. Put simply, the longer it takes a trend to reach the national level, the longer that trend's overall lifespan is likely to last. 41

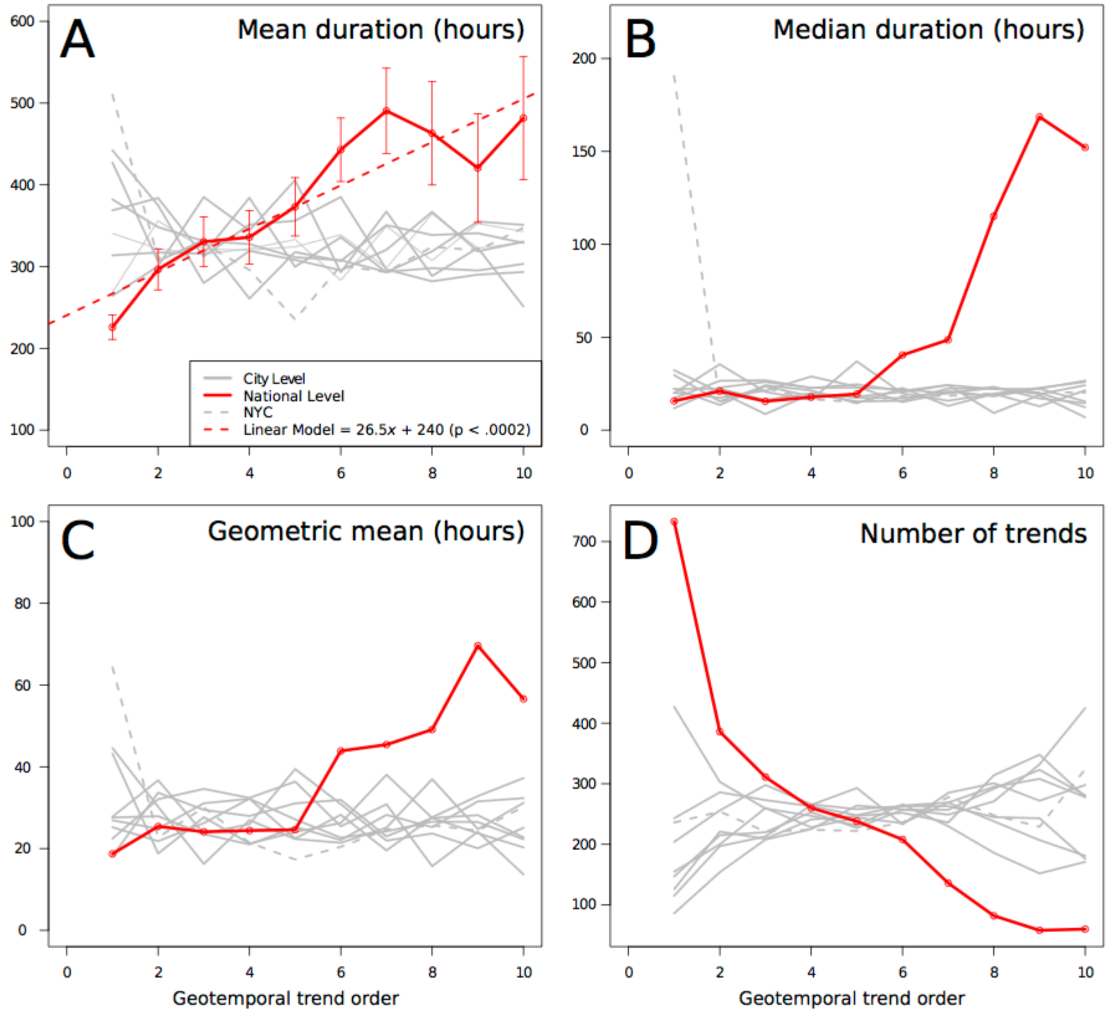

Figure 2. Geotemporal analysis of trend duration.

This result mirrors the pattern observed in Table 2, and once again compliments what common sense might predict. However, as previously mentioned, this result is not a logical consequence of the analysis. Unlike Table 2, our geotemporal analysis reveals why this is so. Consider the nine gray curves in Figure 2. Each of the gray curves is composed of exactly the same data as the red curve. Furthermore, points near the end of each curve reflect the averages of trends that arrived late to each location. Applying similar logic to these data, we might also suspect that average duration should increase commensurately with how long it takes to reach each location. In other words, we would expect each curve to have a slope > 0. Instead, we find no systematic change in duration across ordering for any city-level location (as evidenced by the relatively flat curves). The only possible exception to this finding involves New York City (NYC). Trends in our data that appeared first in NYC tended toward much longer lifespans. A likely explanation for this is the disproportionately large population of NYC, which establishes it as a cultural, technological, and social juggernaut relative to our other locations. Regardless, the remainder of the curve associated with NYC is flat and indistinguishable from our other cities, and lies in stark contrast to the pattern observed at the national level. 42

If each curve represents the same data, why do we observe such a striking difference in the functional form of trend lifespans when our investigative lens shifts from local city-level perspectives to a national one? The key to this result lies in the effect of geotemporal order. By design, each trend included in our analysis appears at all ten locations. However, it is entirely possible — likely, even — that trends quickly spread and gain attention at multiple locations, only to quickly diminish in popularity shortly thereafter. Indeed, this is the predominate life-cycle of trends observed in our data. Most topics quickly recede into obscurity — roughly 50% of the topics we measured disappeared in less than 21 hours. Because all trends identified as such by Twitter reflect topics receiving attention in social media, the question of interest becomes what separates popular sentiment from so-called viral trends? 43

Given the sheer number of active users and their tweets, virality can no longer be defined according to raw counts. Rather, the hyper-circulatory nature of viral information should reflect the underlying social attention itself, which in turn can lead to reiteration and recirculation. A metric of attention would ideally factor in activities and influences across multiple social networks, news media, popular culture outlets such as television, podcasts, and radio, as well as the impossible-to-measure aspects of interpersonal interactions between users and their “real-life” friends and family. Despite the untenable nature of such an analysis, we believe the geotemporal analysis provided here gives us a glimpse into the nature of such hyper-circulating information. This glimpse is reflected in the behavior of trends grouped according to their appearance at the national level. Notice again that trends that appear first at the national level do not particularly fair well from a sustainability standpoint, relative to any other location. This observation holds for trends that appear at the national level in position one through five of their life-cycle. However, a tipping point occurs around position six, after which trends generally persist for much greater periods of time. It is worth noting that this striking increase holds for all metrics — mean, median, and geometric mean — strongly indicating that this effect is not an artifact of measurement alone (e.g., the effect is not due to random outliers). The effect is correlated with the number of trends observed at each position in the geotemporal curve of the national-level analysis (see Panel D of Figure 2). This reduction in the number of trends should not be viewed as the cause of the increased durations observed, but rather as a by-product of the mechanisms mediating the lifespan pattern. Put simply, fewer trends are likely to achieve “viral” status, so we should expect increasingly fewer observations as we move rightward along the red curve. 44

We are careful to not draw cause and effect conclusions, particularly given the limited scope of our analysis. A future investigation might include monitoring a greater number of city-level locations, as well as foreign locales. Nonetheless, we speculate that this uptick in trend sustainability might reflect an underlying change in the nature of the attention driving the mechanisms of reiteration and recirculation. We refer to this critical mass as the “virality threshold.” At this point it is not possible to say with certainty what “causes” topics to surpass the virality threshold, if such a notion is even meaningful. One difference between “normal” trends and viral trends is apparent though: trends that quickly rise to detectable levels at the national level tend to be volatile in nature — quick to rise, and quick to fall. By contrast, topics that bubble up slowly, growing and spreading across the country in systemic fashion are much more likely to persist. The next logical question then becomes what characterizes such topics in today’s world of ubiquitous technology and information overload? We can only speculate, but we believe that identifying such information will require simultaneous analysis of both inter- and intra-network connectivity. For example, it is impossible to fully understand hyper-circulation activity on Twitter without considering chatter on Facebook, the news, television, etc. A look at Table 3 provides a tantalizing, albeit anecdotal confirmation of our suspicions. In the left-most column we report the twenty shortest-lived trends that made their first appearance at the national level. The twenty trends listed in the second column represent the topics with the longest lifespans, taken from those trends that made their tenth and final appearance in our data at the national level. By contrast to the shortest-lived trends, long-lived trends tend to center around celebrities and corporate identities more likely to be circulating on more “controlled” networks, such as television and news media. A more careful investigation into the mechanisms dictating the nature of the virality threshold is left to future research. 45

20 Shortest-lived trends	20 Longest-lived trends
Hanna and Emily	Jordan Hill
ASC	#WCW
Dumb Hoers	Jordan Crawford
#AskAngie	Jimmy Kimmel
Evil Ryu	#coupon
William James	Florida
Wifey Material Or Nah	#MCM
#NFLIndoors	Tyler
AIM	Holtby
Leonard Fournette	World Cup
Ron Baker	#PLL
#BartTo40K	Freaky Friday
50% of Americans	Kendrick
Pickle Eggs	Keenan Allen
If She Eat	Jamaal Charles
#SOYJOY	Vegas
Tekele Cotton	Tom Hanks
Shane Larkin	WCW
South Carolina State University	#gameday
Powerball	Steve Harvey

Table 3. Forty trends observed with the shortest and longest lifespans. The shortest-lived trends were taken from those topics that appeared at the nation level as their first location. The longest-lived trends were those identified at the national level as their tenth location. These correspond to the shortest and longest trends depicted in the first and last points on the red curve in Figure 2.

Future Writing Studies: Final Thoughts

As noted above, the results from this study suggest that the longer a trend takes to reach a national audience, the longer that trend is likely to last. Long-lasting trends tend to be slow and steady, systematically circulating from city to city until finally reaching the national level, while trends that first appear at the national level are more volatile and chaotic. To use a metaphor often found in politics, trends that initiate from a “grass roots” level (city-level rather than national level) have a greater potential to be sustained over time. Or, to use another metaphor, a bottom-up approach lasts longer than top-down, if we think of the “top” as national conversations driven by celebrities, large-scale institutions, mass media, etc. 46

If “rhetorical velocity” can be a practical tool of digital rhetorics, and if velocity is defined in physics as “speed with direction,” then this study suggests that the goal that any user might want to achieve through digital delivery is a direction with slow and steady speed. As much as digital technologies are lauded (or cursed) for their ability to create near instantaneous ubiquity, such high speed can be followed quickly by relative obscurity. 47

Such findings might also suggest that, like performance-based modes of delivery through a human body, digital delivery requires a rhythm: it cannot simply rise and crescendo forever, but works best when alternating between upbeats and downbeats, perhaps picking up the pace from time to time, but slowly building toward a chorus where multiple voices join in and tweet along. However, just as not all trends are the same, not all “viruses” are the same, either. If attention is in fact becoming a justification, in and of itself, then we need to begin to better differentiate between kinds and types of attention — or, another way to put it, we need to develop a differential typology of iteration. This article suggests one type of differentiation by using the virality threshold to tell the difference between normal trends, and trends that have 48 become viral.

Specific to this study, there's also a geographical component at play for how trends might circulate — another possible form of differentiation. Although this study didn't look at the individual content of tweets beyond a trend analysis, nor why one tweet might circulate from one location to another (only that they circulated), future research might investigate the reasons why one trend might move from one location to another. Animal studies describes one form of animal navigation as "path integration," where an animal integrates continuous path cues throughout its travels and uses them to navigate back to a starting point. A theory of "trend integration" might be developed in order to track the movement of trends and determine the various cues that help spread a trend from one location to the next. However, given that animals use a variety of senses to achieve path integration (inertial cues, proprioception, motor efference, optic flow, and sensitivity to the earth's magnetic field), and given that social network activity occurs in a complex system, any practical attempt to evaluate "trend integration" would require additional development in the tools and techniques for studying social network trends. 49

Such research would require an integration of the macroscopic (what we have proposed here) with the microscopic. To return once again to our school of fish, an ichthyologist can look at how fish behave individually, as a school, or how these two levels operate together to create a whole, organized movement. Our contention is that this kind of analysis *could* be done with large-scale writing systems, but not at this time, for too little research has looked at the macroscopic perspective to understand exactly how a whole school of tweets moves. Even an individual fish (tweeter) cannot be abstracted into a generalized higher-level operator (trend) without some sense of the larger scale. We cannot integrate micro to macro until we first understand them separately. 50

As for our introduction of the term *virality threshold*, we think this concept, and others like it, are the first steps toward critical frameworks for investigating how attention functions in network culture. If attention itself, more so than the content of any trend, is the driving force of hyper-circulation and network saturation, then we need to develop operational methodologies for determining whether or not something actually is "viral." If a digital object's status as "trending" or "viral" is the basis for developing narratives in news media and entertainment journalism, then it is important to make sure other trends and viral topics are not being ignored — especially ones attached to social movements and cultural critique. Eunsong Kim, in the "Politics of Trending", reminds readers that "Trending is visibility granted by a closed, private organization and their proprietary algorithms" [Kim 2015]. Twitter's system for locating and identifying trends, for example, is a black box; Twitter does not explain how they quantify trends. Therefore, it may be even more necessary to develop methodologies for showing that something is *not* trending or viral — even though news media, advertising, and marketing campaigns may try to start a trend by pretending it has already gained momentum. If our research accomplishes nothing else, we hope it encourages critical responses to the phrases "this is trending" or "this is viral" with questions like: How do you know it was a trend or viral topic? Within how *many* networks did this actually trend or become viral? Is this an organic/grassroots trend, or has it been manufactured? 51

While the virality threshold opens the possibility of differentiating between popular trends and those cultural artifacts that become viral in nature, much future work is needed to determine how such differentiations may be investigated operationally across multiple networks. Also, the question remains unanswered as to whether virality is simply a case of attention garnered, or does the content have a significant effect on which trends reach and surpass the virality threshold? We think there is a tipping point when the content becomes less important than the attention itself, but it is not clear whether that will ever be predictable. This question is a new spin on the age-old form/content debate, but in this case the form is a given typology of iteration and the content is whatever text, topic, or digital/cultural artifact motivated a given trend in the first place. This is not a classic chicken/egg problem, because without a doubt a cultural artifact must be delivered to a network before it can gain attention. Rather, the questions raised here are how does the content gain attention *after* its initial delivery, how much of a role in attention-gathering does content play, and/or does sizable attention itself aggregate attention and play a far more significant role in what eventually becomes viral? We offer no answers to these questions, but we hope this study shows the possibility of using operational research methodologies to understand macroscopic attention. Certainly, these questions will never have a final answer — as networked culture and media ecologies are always dynamic, shifting, and changing over time — but this does not mean that writing studies and the digital humanities cannot expand the available methodologies for investigating these 52 questions and provide tentatively effective descriptions for how attention functions.

Notes

[1] While Valente does reference macroscopic methods, the framework is only used to address the ways that network saturation and critical mass effect potential changes in the adoption threshold of individuals within a network. In other words, it only serves to reinforce DoI's microscopic analyses (user-as-node/adoption), rather than standing on its own as a separate framework for analyzing networks from the top-down or at a distance.

[2] Of course, the canons of rhetoric are highly interconnected, and elements of style may certainly contribute to delivery. However, we make such distinctions here to differentiate from Lanham's focus on style.

[3] Traditionally, such data was not always available to scholars, which is one reason the authors used MassMine, which allowed the authors to archive social media data independent of the social media platforms.

[4] For more information about MassMine (or to download it), visit MassMine.org.

Works Cited

Bifet and Frank 2010 Bifet, A. and Frank, E. (2010). “Sentiment Knowledge Discovery in Twitter Streaming Data”. *Discovery Science*. Eds. Bernhard Pfahringer, Geoff Holmes, and Achim Hoffmann. New York: Springer. 1-15.

Brooke 2009 Brooke, C. G. (2009). *Lingua Fracta: Toward a Rhetoric of New Media*. New York: Hampton Press.

Davenport and Beck 2002 Davenport, T. H. and Beck, J. C. (2002). *The Attention Economy: Understanding the New Currency of Business*. Boston: Harvard Business Review Press.

Derrida 1988 Derrida, J. (1988). *Limited Inc*. Evanston: Northwestern University Press.

Dobrin 2011 Dobrin, S. I. (2011). *Postcomposition*. Carbondale: Southern Illinois University Press.

Emerson et al. 2012 Emerson, T., Ghosh, R., and Smith, E. (2012). “Case Study: Predicting Exposure of Social Messages: The Bin Laden Live Tweeter”. *Practical Text Mining and Statistical Analysis for Non-Structured Text Data Applications*. Eds. Gary Miner, John Elder, Andrew Fast, Thomas Hill, Robert Nisbet, and Dursun Delen. Waltham: Academic Press, 2012. 797-802.

Guille and Hacid 2012 Guille, A., and Hacid, H. (2012). “A Predictive Model for the Temporal Dynamics of Information Diffusion in Online Social Networks”. In *Proceedings of the 21st International Conference on World Wide Web* (pp. 1145–1152). New York, NY, USA: ACM. http://doi.org/10.1145/2187980.2188254

Jockers 2013 Jockers, M. L. (2013). *Macroanalysis: Digital Methods and Literary History*. Champaign: Illinois University Press.

Kim 2015 Kim, E. “The Politics of Trending”. *Model View Culture*. March 19, 2015. https://modelviewculture.com/pieces/the-politics-of-trending.

Krikorian 2014 Krikorian, R. “Introducing Twitter Grants”. *Twitter Blog*. February 5, 2014. https://blog.twitter.com/2014/introducing-twitter-data-grants.

Lanham 2007 Lanham, R. A. (2007). *The Economics of Attention: Style and Substance in the Age of Information*. Chicago: University of Chicago Press.

Lee, Kwak, Park, and Moon 2010 Lee, C., Kwak, H., Park, H., and Moon, S. (2010). “Finding Influentials Based on the Temporal Order of Information Adoption in Twitter”. In *Proceedings of the 19th International Conference on World Wide Web* (pp. 1137–1138). New York, NY, USA: ACM. http://doi.org/10.1145/1772690.1772842

Meyers 2009 Meyers, R. A. (Ed.) (2009). *Encyclopedia of Complexity and Systems Science, 2009 edition*. New York: Springer.

Moretti 2007 Moretti, F. (2007). *Graphs, Maps, Trees: Abstract Models for Literary History*. New York: Verso.

Paul and Dredze 2011 Paul, M. J., and Dredze, M. (2011). “You Are What You Tweet: Analyzing Twitter for Public Health”. *ICWSM*, July: 265-272.

Ridolfo and DeVoss 2009 Ridolfo, J. and DeVoss, D. N. (2009). “Composing for Recomposition: Rhetorical Velocity and

Delivery". *Kairos: A Journal of Rhetoric, Technology, and Pedagogy*, 13 (2).

Ridolfo and Hart-Davidson 2015 Ridolfo, J. and Hart-Davidson, W. (2015). "Introduction". *Rhetoric and the Digital Humanities*. Eds. Jim Ridolfo and William Hard-Davidson. Chicago: University of Chicago Press. 1-12.

Rogers 1983 Rogers, E. M. (1983). *Diffusion of Innovations*. New York: Free Press.

Sánchez 2006 Sánchez, R. (2006). *The Function of Theory in Composition Studies*. Albany: State University of New York Press.

Valente 1995 Valente, T. W. (1995). *Network Models of the Diffusion of Innovations*. Cresskill, N.J: Hampton Press.

Wolfe 2009 Wolfe, C. (2009). *What is Posthumanism?* Minneapolis: University of Minnesota Press.

Wu, Hofman, Mason, and Watts 2011 Wu, S., Hofman, J. M., Mason, W. A., and Watts, D. J. (2011). "Who Says What to Whom on Twitter". In *Proceedings of the 20th International Conference on World Wide Web* (pp. 705–714). New York, NY, USA: ACM. http://doi.org/10.1145/1963405.1963504