Growing Bot Security:
An Ecological View of Bot Agency

DOUGLAS GUILBEAULT[1]
University of Pennsylvania, USA

Political actors are now deploying software programs called social bots that use social networking services such as Facebook or Twitter to communicate with users and manipulate their behavior, creating profound issues for Internet security. Current approaches in bot control continue to fail because social media platforms supply communication resources that allow bots to escape detection and enact influence. Bots become agents by harnessing profile settings, popularity measures, and automated conversation tools, along with vast amounts of user data that social media platforms make available. This article develops an ecological approach to thinking about bots that focuses on how social media environments propel bots into agency. This habitat-based model uses bots to expose ripe targets of intervention and innovation at the level of interface design. It also situates bots in the context of platform providers with a vested interest in interface design, revealing a range of new political problems. Most important, it invites a hybrid ethics, wherein humans and bots act together to solve problems in bot security and Internet ethics more broadly.

Keywords: Internet security, social bots, agency, ecology, algorithms, philosophy of information, ethics, power

The concept of the political platform derives from a time when politicians would stand on actual platforms and engage in the art of rhetoric. Today, politics takes place over a fundamentally different kind of platform, the platform of social networking services (SNSs), where humans are not the only agents capable of persuasion. Corporations, politicians, and even militaries are deploying software programs

called *social bots* that use SNSs to communicate with users and manipulate their behavior, posing profound issues for Internet security (Woolley, 2016). There are efforts to detect malicious bots via software systems and user vigilance, but these efforts continue to fail because of security vulnerabilities that social media platforms create. Social media platforms traffic massive amounts of user data that allow bots to target susceptible users and infiltrate networks while avoiding detection software. Social media platforms also supply automated communication tools that make it difficult for users to distinguish real and fake accounts. To protect against bots, it is vital to identify and modify the features of SNSs that increase user vulnerability. Current methods are inadequate for the task insofar as coders and users are themselves under the influence of platform vulnerabilities, indicating the need for a new approach.

This article develops an ecological approach to bot security that focuses on how social media platforms propel bots into agency. Traditionally, agency is said to rely on the general ability to construct what Aristotle (*On Rhetoric*, trans. 1991, II.i.2) called an *ethos*—a view of oneself as a certain kind of person—with intentions, beliefs, and moral responsibility. Several scholars deny bots agency on the basis of Aristotle's theories, but their analyses focus on early generations of bots that fail miserably at gaining acceptance as real users (Kennedy, 2009; Miller, 2007). Social bots, by contrast, are remarkably successful at gaining acceptance as real users, not because of changes in bot intelligence, but because of changes in their environment. This article shows how SNSs are a new kind of habitat that imposes habits of self-construction that both humans and bots equally exploit. Social bots achieve agency by manipulating profile settings, popularity measures, and automated conversation tools. An ecological approach to bots focuses on interface features of networking platforms as targets of intervention and innovation, thereby expanding the scope of bot security. It also situates bots in the context of platform providers with a vested interest in interface design, revealing a range of new political problems. Most important, it opens toward a hybrid ethics, wherein humans and bots act together to solve problems in Internet security.

The argument that follows proceeds through five stages. First, it reviews the current approaches in bot security and shows how social media platforms are a serious obstacle to their success. The second section unearths the human-centered view of agency implicit in bot security and explains its limitations in addressing bots. The third section shows how bots manipulate communities through a kind of platform persuasion, which depends on interface features that enact political power over human users. Based on these findings, the fourth section develops an ecological view of agency as depending on the norms of identity construction that online habitats impose and grow. The fifth section lays the foundations for a hybrid ethics and discusses its ecological futures.

The Limits of Bot Security

The primary approach in bot security concerns the design of bot detection software, where agency lies solely in the hands of defense coders and cybercriminals. Some detection software can classify bots that deviate strongly from normal users in terms of click rate, message frequency, and time of operation (Ratkiewicz et al., 2011; Wang, Konolige, et al., 2013). Other software systems find mild success in using network structure to detect bots (Alvisi, Clement, Epasto, Lattanzi, & Panconesi, 2013; Fire, Goldschmidt, & Elovici, 2014; Wald, Khoshgoftaar, Napolitano, & Sumner, 2013). Despite these

advancements, a recent bot detection competition showed how the leading detection algorithms are significantly limited, with each requiring error prone human supervision at several stages of analysis (Subrahmanian, Azaria, et. al. 2016). Meanwhile, bots continue to evade and invade the bot detection software of leading SNSs. The question then follows: How are bots so effective at escaping detection? One may expect that bots escape detection because of their intelligence, but this is untrue.

At present, bot designers do not really build artificial intelligence into bots. Instead, bot designers equip bots with basic scripts for adapting to the user data that social media platforms make available. Bots can harness public data to target users who are likely to connect with bots, based on their connection forming habits (Xie et al., 2012; Zangerie & Specht, 2014). Bots can also model user data and adapt to social norms, which enhances their persuasiveness (Ferrara, Varol, Davis, Menczer, & Flammini, 2014; Messias, Schmidt, Oliveira, & Benevenuto, 2013; Nanis, Pearce, & Hwang, 2011). Most daunting is the rise of botnets, which can use a central bot intelligence to coordinate hundreds of bots in data mining or denial of service attacks. Boshmaf, Muslukhov, Beznosov, and Ripeanu (2011, 2013) showed how botnets can penetrate Facebook with a success rate of more than 80%. Rodriguez-Gomez, Marcia-Fernandez, and Garcia-Teodoro (2013) further showed how botnets can evolve through a dynamic life cycle that challenges the ability to distinguish between bots and humans. Many of the major weaknesses in bot detection software stem from the vast amounts of user data that platforms traffic, something over which defense coders have little to no control.

Even more condemning for bot detection software is the fact that many users make all of their data public and connect to others without discretion, providing bots with an easy entry point into a network. Romero, Galuba, Asur, and Huberman (2011) and Chu, Gianvecchio, Wang, and Jajodia (2010) cautioned that so long as users are irresponsible in their security choices, SNSs will remain vulnerable to bot attacks. Internet illiteracy is an important factor here, but it only partly accounts for the security practices of users (Leon et al., 2012). Online platforms substantially influence the security choices of users in the direction of increased vulnerability. For example, social media platforms often display popularity measures on user profiles that incentivize the practice of forming new connections. Indeed, Steinfield, Ellison, and Lampe (2008) and Wilson, Gosling, and Graham (2012) provided evidence that popularity measures drive users to adopt behavior that maximizes their popularity. In this way, the vulnerability of users is not simply a result of Internet illiteracy. It is, crucially, a product of the cultural changes in identity production and community formation that social media platforms induce.

The second major approach in bot security involves raising user awareness, placing responsibility in the hands of human users. Most users are unaware of bots, even though bots now make up almost 50% of online traffic (Zeifman, 2015). For this reason, researchers aim to educate users on how to identify, avoid, and report bots (Davis, Varol, Ferrara, Flammini, & Menczer, 2016; Wagner, Mitter, Korner, & Strohmaier, 2012; Woolley & Howard, 2014). There are attempts to crowd source human expertise in bot detection with minor success, as the average user is quite poor at the task (Wang, Mohanlal, et al., 2013). The fact is that SNSs are an environment where the distinction between humans and bots is inherently blurry. Edwards, Edwards, Spence, and Ashleigh (2014) showed that Twitter bots often garner just as much credibility as humans, even when users know they are bots. Software programs are often perceived as agents because digital environments create what Floridi (2015) called *a proxy*

culture, where agency is assessed on the basis of minimal and newly constructed cues (Walther, 1996). Offline, humans rely on body language and speech to determine authenticity and trust (Metzger & Flanagin, 2013). But online, users rely on profile pictures, emoticons, and other automated tools that create a much lower bar for bots. SNSs also level the linguistic playing field. Turing proposed that a program can be deemed intelligent if it can perform indistinguishably from a human in conversation (Floridi, 2012). The designers of the bot Realboy noted that "passing the Turing Test is significantly easier when each message is a 140-character tweet. Tweets tend to be disconnected and poorly written . . . and each one should be believable by itself" (Coburn & Marra, 2008, para. 4). In more challenging linguistic environments like Facebook, bots can still copy and rearrange messages from the Web that nevertheless achieve significant influence (Boshmaf et al., 2013). The effort to raise user awareness faces the potentially insurmountable challenge of training users to spot bots in an environment that biases them toward perceiving bots as agents.

Current approaches in bot security are unified in their assumption that humans are the only possible source of intervention. A recurring issue with this assumption is that SNSs introduce security weaknesses that neither coders nor users are likely to overcome. The central problem with bot security is that it ignores how social media platforms limit human intervention while simultaneously empowering bots. Structural changes to social media platforms offer a viable approach to solving problems in bot security and Internet ethics more broadly. However, instituting changes in social media platforms requires understanding interfaces and bots as coevolving forms of agency in a broader media ecology. Viewing interfaces and bots as forms of agency requires a different theory of agency, with guiding intuitions that need to be made explicit. A first and necessary step in this direction is to examine the human-centered view of agency underlying bot security so as to expose its intuitions and reveal its potential for new growth.

Human-Centered Views of Agency

One of the oldest and most influential texts in political theory, Aristotle's *Politics* (trans. 1957), revolves around the claim that humans are political animals, *suis generis*. What makes humans uniquely political, according to Aristotle, are three capacities: (1) the ability to be aware of one's actions, (2) the ability to possess moral responsibility, and (3) the ability to persuade. Aristotle's criteria survive today as the foundation for various theories of agency. For philosophers, agency derives primarily from intentionality and morality (Mitcham, 2014), and for rhetoricians, agency derives primarily from the ability to persuade (Burke, 1966; Miller, 2001, 2007). Both fields generally maintain that only humans are agents with respect to Aristotle's criteria, though there are recent developments toward alternative views. As it is beyond the scope of this article to integrate these traditions, the following discussion focuses on a feature of Aristotle's theory that unifies both traditions in the context of bots, namely ethos.

Aristotle's theory of morality and rhetoric are linked in his notion of *ethos*, which refers to the character of a person. In his *Nicomachean Ethics* (trans. 1982), he argued that to develop an ethos worthy of political leadership, one must cultivate excellence (*arête*) through one's daily routines of thought and action. In his *On Rhetoric*, he added that to achieve positions of power, one must obtain excellence in the art of *ethopoeia*: the art of conveying a character that appeals to an audience (trans. 1991, II.i.2).

According to Aristotle, *ethopoeia* is equally and occasionally more persuasive than *logos* (logic) and *pathos* (emotion) because audiences ultimately assess leadership and credibility on the basis of the speaker's character. The most strategic way to construct ethos, Aristotle held, is to consider that "all people receive favorably speeches spoken in their own character and by persons like themselves" (trans. 1991, II.13.16).

When bots first entered chat forums, they were readily denied agency on the basis of Aristotle's theories. As Miller (2007) explained, *ethopoeia* requires speakers to consciously ascribe an ethos to their audience and use this ethos to anticipate the kind of character they should construct. At the time, bots failed miserably at basic forms of language, let alone self-awareness, and so Miller denied them agency. Even as bots improved in their language abilities, rhetorical scholars continued to deny them agency. In her analysis of wikibots that compose encyclopedic articles, Kennedy (2009) denied them agency because of their inability to self-identify as authors. But as Leonard (1998) predicted decades ago in *Bots: The Origins of New Species*, bots are evolving and their most recent form, the social bot, has the capacity to enact systematic and measurable influence because of critical changes in its environment.

All forms of agency depend on their environment. Aristotle's *ethopoeia* was defined for a largely oral culture, where the canonical form of rhetorical interaction consisted of full-bodied performances on the rostrum. Applying Aristotelian standards to bots assumes that the same standards for agency hold across environments. This may be a useful way to model chat bots and wikibots, whose environments rely on sophisticated standards of language that bots fail to meet. But social bots, by contrast, are remarkably successful at gaining acceptance as real users. SNSs are distinct from chat rooms and data repositories in that they provide users with a Web page for representing their identity (called a profile page), along with a number of automated communication tools built into the website's interface. Through these inbuilt resources, social bots can construct a realistic profile and rise to positions of measurable influence in online communities. Social bots do not need to fully embody Aristotelian agency over SNSs because, in a sense, neither do humans.

There is ample evidence that social media platforms are changing the norms of ethos construction. Beer (2008) explained how users are becoming dependent on social media for sustaining relationships and constructing identity. Papacharissi (2009) showed how self-expression is dictated by interface design, protocols, and default settings across three SNSs: Facebook, LinkedIn, and ASmallWorld. In a later work, Papacharissi (2012) extended these findings to Twitter. Moreover, Wilson et al. (2012) cited several experiments illustrating that SNSs have palpable effects on self-esteem and self-idealization. Various researchers also have provided broader, systemic perspectives on the large-scale changes in human culture that social media websites are contributing to, most notably Turkle (1995), Castells (2009), Papacharissi (2011), and Gillespie, Boczkowski, and Foot (2014). Together, these studies illustrate how social media platforms are channeling the present and future state of human ethos.

When examining ethos online, a trend emerges: Social media platforms not only mold human identity, but they also enable the rise of social bots. As Hwang, Pearce, and Nanis (2012) noted, "digitization drives botification: the use of technology in the realm of human activity enables the creation of software to act in lieu of humans" (p. 4). Their point is that bot influence depends on the ways in which social media platforms have molded human interaction into a mode that bots can simulate and exploit.

Social media platforms introduce rhetorical vulnerabilities into the user experience that bots manipulate. Social media providers have become aware of these rhetorical vulnerabilities and have begun to exploit them. Van Dijck (2013) quoted an interview with the CEO of LinkedIn who displays an astute awareness of "the behavioral changes taking place as a result of that infrastructure, the way in which people represent their identity, the way in which people are connecting with others, and the way in which they're sharing . . . everything" (Raice, 2011, para. 19). Social media platforms are a source of power, and social bots tap into this power by manipulating interface features that allow for a kind of platform persuasion.

Platform Persuasion

SNSs house more than a billion users collectively as the single largest consumer of the average user's online life (Perrin, 2015). Van Dijck and Poell (2015) argued that connected citizens have entered a *platform society*, in which interfaces and algorithms direct the evolution of human identity toward corporate and political ends. In the platform society, seemingly innocuous design features can have large-scale effects on culture. Van Dijck (2013) explained how, since 2008, platform providers have begun designing interfaces to steer users toward more traceable and profitable forms of identity production:

> [Social media profiles] are not a *reflection* of one's identity, as Facebook's Marc Zuckerberg wants us to believe, but are part and parcel of a power struggle between users, employers/employees and platform owners to steer online information and behavior. Interfaces are important instruments of identity formation whose steering mechanisms (algorithms, protocols, and default settings) are inscribed in deceptively simple buttons. (p. 212)

Although there is evidence that profiles reflect the personalities of users, this is not an argument against platform manipulation (Back et al., 2010). When users identify with their profiles, they accept the identity template that social media platforms enforce. Van Dijck's core insight, elaborated in her 2009 article, is that user agency is a networked concept, involving humans and algorithms as well as institutional steering activities. Bots are prime examples of the type of algorithmic agent that thrives in the platform society. Bots are able to construct an influential ethos using three platform structures—personal profiles, popularity measures, and automated communication tools—which together constitute the dominant means of platform persuasion over social networking websites.

All major SNSs provide users with a personal profile that allows them to represent their identity through a prescribed set of interface options, including profile images and categories of biographical information. Profile pages publicly display the recent social activity of users, including their posts and their interactions with others. The public accessibility of profiles changes the nature of ethos construction, giving rise to what Castells (2009) called a culture of "mass self-communication," in which users tailor their identity profile to gain the attention of their online community. Platform providers profit immensely from the mass production of personal data and have begun to design profiles that nudge users toward self-promotion. As van Dijck (2012) notes, Facebook's 2011 release of the timeline epitomizes the rise of self-promotion by forcing profiles to log and display all of a user's social activity in a uniform narrative format.

Bots similarly profit from uniform templates of ethos construction because such invariability frees them from the challenge of creating novel forms of self-expression. Profile images allow bots to easily escape the need to simulate realistic body language. They also allow bots to capitalize on rhetorical vulnerabilities associated with personal photos. Messias et al. (2013) showed how bots that download attractive photos from hotornot.com are especially effective at gaining influence. Messias et al., Hwang et al. (2012), and Elishar, Fire, Kagan, and Elovici (2012) further showed how bots can adapt their profiles and messages to users, embodying a rudimentary form of *ethopoeia*. Over social media platforms, bots do not need reflective reasoning to anticipate the ethos of interactants because their personality is already displayed on their profiles, catered to consume.

Another primary form of platform persuasion concerns the use of popularity measures. Without social media, the average person lacks the means to quantify her or his social status. Gillespie (2014) discussed how popularity measures give rise to "calculated publics," where metadata about social status inclines users to develop an ethos that increases their popularity scores. Platform providers profit immensely from calculated publics, and again so do bots. Ellison, Steinfield, and Lampe (2007) discussed how popularity measures cultivate the ideology that having more friends is an enviable social life, which supports the industry of growing and selling network data. Calculated publics also accompany a number of user vulnerabilities that empower bots. Wagner et al. (2012) explained how bots can use popularity measures to target users who are more likely to accept new friend requests: Users with high friend counts are more likely to accept new connections.

Popularity measures can also influence the authenticity or credibility a bot appears to possess. Xie et al. (2012), Metzger et al. (2013), and Westerman, Spence, and Van Der Heide (2012) found that users heuristically assess the credibility of profiles using an "innocent by association" principle in which users with high popularity are assumed to have passed a credibility test by others. Elishar et al. (2013) showed how bots can compromise even the savviest of users, Facebook developers, if the bots obtain mutual friends prior to making the request. The power of popularity measures is not unknown to abusers of bots. Forelle, Howard, Monroy-Hernández, and Savage (2015) and Woolley (2016) discussed how politicians hire thousands of bots to pad their popularity and enhance their credibility, achieving an entirely new form of bot persuasion. Bots thus expose as well as augment the rhetorical influence that popularity measures afford.

The final mode of platform persuasion that bots exploit concerns the use of automated conversation tools. Leading platforms encode various prefabricated forms of messaging into their user interfaces. Facebook provides users with a range of emotional buttons, including the "like" and "love" buttons, which allow users to add a quantifiable endorsement to the public record of appreciation for a user's updates. Park, Baek, and Cha (2014) explained how emoticons shape the emotions that users tend to express, with interesting cross-cultural differences. There is even evidence of users becoming addicted to the forms of positive reinforcement that automated communication tools supply (Dwyer & Fraser, 2016; Kuss & Griffiths, 2011; LaRose, Kim, & Peng, 2011). Twitter supplies users with automated conversation tools for receiving updates about other users (i.e., "following") and for echoing someone else's tweet back into the Twitterverse (i.e., "retweeting"). Messias et al. (2013) designed bots that achieve highly

influential positions simply by following only users who follow them back within a short period of time. Haustein et al. (2016) illustrated how similar feedback loops facilitate influence in the case of retweeting.

Bots also augment the persuasive effects of automated communication tools. The Twitter interface contains a window that informs users of the tweets currently trending in the Twitterverse. New rhetorical opportunities emerge from the ability to bias the tweets that appear in the trending window. Abokhodair, Yoo, and McDonald (2016) observed a bot strategy they called *smoke screening* (also called *hijacking the hashtag*; Chu, Widjaja, & Wang, 2012) in which bots change the associations of political hashtags by using them thousands of times in connection with distracting topics, such as food or tourism. Abokhodair et al. also discussed a related strategy in which bots stifle the organizing effects of revolutionary tweets by spamming Twitter with alternative hashtags that consume the trending pool. In this manner, bots enable mass-media strategies that far exceed the expressive capacity of individual users.

One of the first-ever bot competitions—The Web Ecology Project— demonstrates the power of platform persuasion. Software engineers from around the world competed to design the most effective bot for influencing a network of 500 unsuspecting Twitter users. The bots succeeded in substantially reshaping the network, drawing responses and interactions from users who were not directly connected previously. The organizers predicted that in the future bots will be able to "subtly shape and influence targets across much larger user networks, driving them to connect with targets, or to share opinions and shape consensus in a particular direction" (Hwang et al., 2012, p. 41). In a field test report involving real bot deployments, the Web Ecology organizers further speculated that one day bots will have the ability to determine, in their words, "the general interests of other users" (Nanis et al., 2011, p. 3). The severity of such forecasts demands discussion about bots as genuine agents, with special attention to their present and future impact on human users. Human-centered views of agency are insufficient for the task, insofar as they limit the scope of intervention to only what humans can change in their own behavior, either as coders or as users.

What Internet ethicists need is a perspective that situates users within a broader ecosystem of interactions, where interfaces and bots are dynamic sources of rhetorical influence. Articulating this perspective is a formidable challenge, but the path is not totally obscured. From Leonard's (1998) *Bots: The Origins of New Species* to the Web Ecology Project, bots have long inspired biological and, in particular, ecological imagery. Exploring this ecological imagery will crystallize the intuition that bot agency is the product of mutually transformative interactions among humans, bots, and their environment.

Toward an Ecological View of Agency

The link between ethos and ecology reaches far back in the watershed of human culture. Hyde (2004) explained how the word *ethos* takes its origins from the Greek *ethea-* (ἤθεα), meaning "accustomed place" and "home." *Ecology* derives from a related root, the Greek *oikos* (οἶκος), also meaning "home" or "place of dwelling" ("Ecology," 2016a). The pre-Socratic *ethea-* additionally referred to "custom" or "habit," where *habit* stood for both the habits of organisms as well as their habitats. In fact,

the terms *habit* and *habitat* both derive from the Latin *habitāre,* meaning "to live, inhabit, and dwell," where the essential link lies in how environments *habit*-uate organisms to particular modes of dwelling ("Habit," 2016; "Habitat," 2016). Aristotle's theory of agency internalizes these etymologies. For Aristotle, ethos arises not only from rhetorical self-expression, but also from the habits that one develops while dwelling in material and social habitats. Social media platforms embody both dimensions of ethos simultaneously: They are environments where the available modes of dwelling consist of prescribed ways of constructing the self. Every movement, every click, every utterance is recordable as an act of self-construction in the age of big data (González-Bailón, 2013; Neff & Nafus, 2016). For this reason, social media platforms are an entirely new habitat, and social bots are among the new forms of agency that social media habitats grow.

The ecological view of agency resonates with Latour's (1987) actor–network theory and its recent developments by Verbeek (2005), Bennett (2010), and Rickert (2013). What the ecological view adds is a functional distinction between organisms and environments that allows a number of mechanisms from ecological theory to inform its analyses. In this direction, Floridi (2014a) defined technological agency using three environmentally oriented criteria: *interactivity*, *adaptability*, and *autonomy*. Interactivity refers to the ability to interact with the environment; adaptability refers to the ability to optimize interactions with the environment on the basis of internally generated rules; and autonomy refers to the ability to develop new rules for behavior, in response to environmental conditions. Social bots satisfy these criteria in full. They interact with their environment via sensors, they adapt to user data, and they evolve behaviors independent of their designers. The ecological approach reveals that, in the case of social bots, Floridi's mechanisms facilitate agency because they allow bots to construct a viable ethos in particular habitats. Floridi identified general mechanisms that characterize agency across environments. The ecological approach applies Floridi's mechanisms to specific environments with respect to the norms of ethos that constitute a particular habitat. To discover the norms of ethos in a particular habitat, the ecological approach does not rely on human inference alone. It invites a hybrid approach in which bots serve to unveil and explore the skeletal grooves of agency built into networking platforms, thus exposing ripe sites of intervention and innovation.

In pursuing a hybrid methodology, it is useful to learn from one of the first ecological hybrid sciences: *cybernetics.* Pickering (2006) explained how cyberneticians designed technologies as *ontology engines* to both demonstrate their philosophy and spread it through society. As an example, he referred to Ashby's (1948) *homeostat*, which models the brain and its environment with identical circuits that regulate each other through electrical signals. The homeostat refers to the biological process of *homeostasis*, in which ecologies sustain a dynamic organization through regulatory mechanisms. The founder of cybernetics, Norbert Weiner (1950), argued that social systems also rely on homeostasis, in which communication technologies are the means of ecological regulation. In his *Steps toward an Ecology of Mind*, Bateson (1972) explained that, for cyberneticians, the self is an information ecology that is modified as it extends through communication media. Because technology is inseparable from the evolution of self, Bateson maintained that cybernetic technologies are "of great importance, not only theoretical, but also ethical" (p. 466). Pask was a cybernetician who actively designed technology to augment the growth of self, as revealed by his efforts to construct "an environment with which the inhabitant cooperates" (Pask, 1969, p. 496). Today, these views are advanced by Floridi (2011, 2014b) who argues that information

technologies shape the ontological and ethical questions that occupy a culture. In his view, humans are information organisms that hang in the delicate balance of a vast information ecology, filled with countless other species of information organisms, including bots. To ignore the role of social media platforms in growing human-bot ecologies is not unlike ignoring the health of natural environments in growing the future of the human species. As a step in this direction, I recommend that Internet ethicists view social bots as an ontology engine for the ecological view of agency, not only because this view provides explanatory coherence, but also because it reveals new solutions in the domain of hybrid ethics.

Hybrid Ethics

Bots are among the new technologies that, in the words of Howard (2015), have the potential to either lock us up or set us free. To ensure that bots facilitate the latter, it is necessary to identify and modify the features of online platforms that weaken the security of users. Human-centered approaches are inadequate for the task because both coders and users are largely blind to the ideological forces that enliven platforms. Platform persuasion depends on website architecture sinking into what Beer (2009) called the *technological unconscious* as an implicit chassis that gradually molds human psychology below conscious radar. Hayles (2005) elaborated, "Software is ideology. The interpolation of the user into the mechanic system does not require his or her conscious recognition of how he or she is being disciplined by the machine to become a certain kind of subject" (p. 61). As ontology engines, bots expose the platform features that reinforce the kind of subject online habitats are disciplining users to be, thereby indicating targets of intervention that can enhance security and augment civic opportunity.

Social bots are ideal tools for ethnographic research. Geiger (2014) illustrated that telling stories about bots captures their impact on human culture. An ecological view of bots suggests that the reverse is also possible: Using bots to tell stories about humans can reveal the cultural changes already underway. Researchers can release bots into networks and allow them to grow an ethos that represents, in microcosm, the minimal conditions for agency in a particular habitat. Examining bots from this perspective can inform environmental changes to online habitats for the purpose of strengthening user security. For instance, making popularity scores optional, private, or even nonexistent may significantly strengthen user resistance to bot attacks. It is further possible to imagine designing bots to discover latent forms of platform persuasion that Internet ethicists can pre-emptively protect against. Although environmental interventions offer a viable alternative in bot security, they have the consequence of situating bots in the context of corporations with a vested interest in interface design, creating entirely new issues for Internet ethics. These issues are all the more imminent considering recent evidence that major platform providers, including Google and Facebook, are investing billions of dollars into incorporating intelligent bots into their services ("Imperial Ambitions," 2016).

In the hands of platform providers hungry for data, bots are weapons of power. Bots have the potential to initiate a leap in the generation and traceability of lucrative user data. Bots also embody the perfect confederates for social experimentation, a controversial pastime of platform providers (Kramer, Guillory, & Hancock, 2015). However, in the hands of users, bots also have immense potential to augment political agency and civic opportunities. Bots could enable users to gather and spread political information much more effectively. Bots may also enable users to organize communities of much greater proportions

and diversity than previously possible. There already exist publicly accessible websites for designing and purchasing bots, although these services are in a nascent phase that major corporations can easily outmatch. Soon it will be possible to design bots for monitoring and optimizing communication in ways that will benefit users and social media providers. It is possible that one day bots will be able to address problems of echo chambers or misinformation (Jamieson & Cappella, 2010). It is even imaginable that bots will be able to augment economic and discursive productivity, inspired by Hidalgo's (2015) *Why Information Grows*. A crucial question thus emerges: Who, in the future, should control bots? Should users let corporations keep a leash on their digital companions, or should users prioritize open-source platforms for bot design, free from conflicts of interest? This is one of many questions that Internet ethicists will face in the future and for which a hybrid ecological approach will be essential.

The chief insight of the ecological approach is that changes to online environments are inseparable from a process of directed evolution over the future of human identity. Throughout history, scholars have both lamented and admired the unforeseen impact of communication technology on how humans understand their place in society and the universe (Gleick, 2011). Social bots may have particularly serious consequences for human identity because they do much more than provide convenient metaphors. They actually construct an identity that reflects and shapes how humans dwell in the platform society. Major philosophical challenges will arise as it becomes increasingly difficult to distinguish the role of bots as a technology from their role as an actual living extension of the self. For instance, in a world where the average user owns a fleet of bot companions, crawling the Web and interacting with others on her behalf, how will concepts of transparency and anonymity adapt? How will theories of morality and responsibility evolve? How will the psychology of selfhood change in distributed, human–bot ecologies? These are some of the open questions that authors in this volume are beginning to address. To prepare for the challenges ahead, it is crucial to view Internet security as contributing to a broader process of cultural evolution concerning human identity and its link to biological, technological, and hybrid environments. In the Anthropocene, there is an endemic lack of stewardship toward environments of all kinds, natural and digital, and researchers across fields are calling for an ecological paradigm shift concerning information systems (Floridi, 2014b; Haraway, 2015; Morton, 2010). Toward this end, the present article has explored bot agency from an ecological viewpoint in the aim of inspiring methods of platform design that help future generations of humans and bots to grow together harmoniously.

References

Abokhodair, N., Yoo, D., & McDonald, D. W. (2015). Dissecting a social botnet: Growth, content, and influence in Twitter. *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing, 2015* (pp. 839–851). New York, NY: Association for Computing Machinery. doi:10.1145/2675133.2675208

Alvisi, L., Clement, A., Epasto, A., Lattanzi, S., & Panconesi, A. (2013). Sok: The evolution of Sybil defense via social networks. In *Proceedings–34th IEEE Symposium on Security and Privacy, 2013* (pp. 382–396). Piscataway, NJ: Institute of Electrical and Electronics Engineers. doi:10.1109/SP.2013.33

Ashby, R. W. (1948). The homeostat. *Electronic Engineering, 20*, 380ff.

Back, M. D., Stopfer, J. M., Vazire, S., Gaddis, S., Schmukle, S. C., Egloff, B., & Gosling, S. D. (2010). Facebook profiles reflect actual personality, not self-idealization. *Psychological Science*, *21*(3), 372–374. doi:10.1177/0956797609360756

Bateson, G. (1972). *Steps to an ecology of mind.* New York, NY: Ballantine.

Beer, D. (2008). Social network(ing) sites . . . revisiting the story so far: A response to danah boyd and Nicole Ellison. *Journal of Computer-Mediated Communication, 13*(2), 515–529. doi:10.1111/j.1083-6101.2008.00408.x

Beer, D. (2009). Power through the algorithm? Participatory Web cultures and the technological unconsciousness. *New Media & Society, 11*(6), 985–1002. doi:10.1177/1461444809336551

Bennett, J. (2010). *Vibrant matter: A political ecology of things*. Durham, NC: Duke University Press.

Boshmaf, Y., Muslukhov, I., Beznosov, K., & Ripeanu, M. (2011). The socialbot network: When bots socialize for fame and money. In *ACSAC '11: Proceedings of the 27th Annual Computer Security Applications Conference* (pp. 93–102). New York, NY: Association for Computing Machinery. doi:10.1145/2076732.2076746

Boshmaf, Y., Muslukhov, I., Beznosov, K., & Ripeanu, M. (2013). Design and analysis of a social botnet. *Computer Networks, 57*(2), 556–578. doi:10.1016/j.comnet.2012.06.006

Burke, K. (1966). Medium as "message." In K. Burke (Ed.), *Language as symbolic action: Essays on life, literature, and method* (pp. 410–418). Berkeley, CA: University of California Press.

Castells, M. (2009). *Communication power*. New York, NY: Oxford University Press.

Chu, Z., Gianvecchio, S., Wang, S., & Jajodia, S. (2010). Who is tweeting on Twitter: Human, bot, or cyborg? In *Proceedings of the 26th Annual Computer Security Applications Conference* (pp. 21–30). New York, NY: Association for Computing Machinery. doi:10.1145/1920261.1920265

Chu, Z., Widjaja, I., & Wang, H. (2012). Detecting social spam campaigns on Twitter. In *Proceedings of the 10th International Conference on Applied Cryptography and Network Security* (pp. 455–472*)*. New York, NY: Springer*.* doi:10.1007/978-3-642-31284-7_27

Coburn, Z., & Marra, G. (2008). Realboy: Believable Twitter bots. Retrieved from http://ca.olin.edu/2008/realboy/index.html

Davis, C., Varol, E., Ferrara, E., Flammini, A., & Menczer, F. (2016). BotOrNot: A system to evaluate social bots. In *Proceedings of the 25th International Conference Companion on World Wide Web* (pp. 273–274). New York, NY: Association for Computing Machinery (International World Wide Web Conferences Steering Committee). doi:10.1145/2872518.2889302

Dijck, J. van. (2009). Users like you? Theorizing agency in user-generated content. *Media, Culture & Society, 31*(1), 41–58. doi:10.1177/0163443708098245

Dijck, J. van. (2012). Facebook and the engineering of connectivity: A multi-layered approach to social media platforms. *Convergence: The International Journal of Research into Media Technologies, 19*(2), 141–155. doi:10.1177/1354856512457548

Dijck, J. van. (2013). "You have one identity": Performing the self on Facebook and LinkedIn. *Media, Culture & Society*, *35*(2), 199–215. doi:10.1177/0163443712468605

Dijck, J. van, & Poell, T. (2015). Social media and the transformation of public space. *Social Media + Society, 1*(2), 1–5. doi:10.1177/2056305115622482

Dwyer, R., & Fraser, S. (2016). Addicting via hashtags: How is Twitter making addiction? *Contemporary Drug Problems, 43*(1), 79–97. doi:10.1177/0091450916637468

Ecology. (2016). In *Oxford English dictionary online*. Retrieved from http://www.oed.com/view/Entry/59380?redirectedFrom=ecology#eid

Edwards, C., Edwards, A., Spence, P., & Ashleigh, K. (2014). Is that a bot running the social media feed? Testing the differences in perceptions of communication quality for a human agent and a bot agent on Twitter. *Computers in Human Behavior, 33*(2014)*,* 372–376*.* doi:10.1016/j.chb.2013.08.013

Elishar, A., Fire, M., Kagan, D., & Elovici, Y. (2012). Organizational intrusion: Organization mining using socialbots. In *Social Informatics 2012 International Conference* (pp. 7–12). Piscataway, NJ: Institute of Electrical and Electronics Engineers. doi:10.1109/SocialInformatics.2012.39

Elishar, A., Fire, M., Kagan, D., & Elovici, Y. (2013). Homing socialbots: Intrusion on a specific organization's employee using socialbots. In *Conference on Advances in Social Networks Analysis and Mining* (pp. 1358–1365). Piscataway, NJ: Institute of Electrical and Electronics Engineers. doi:10.1145/2492517.2500225

Ellison, N. B., Steinfield, C., & Lampe, C. (2007). The benefits of Facebook friends: Social capital and college students' use of online social network sites. *Journal of Computer-Mediated Communication*, *12*(4), 1143–1168. doi:10.1111/j.1083-6101.2007.00367.x

Ferrara, E., Varol, O., Davis, C., Menczer, F., & Flammini, A. (2014). The rise of social bots. *Communications of the ACM, 59*(7), 96–104. doi:10.1145/2818717

Fire, M., Goldschmidt, R., & Elovici, Y. (2014). Online social networks: Threats and solutions. *IEEE Communications Surveys & Tutorials*, *16*(4), 2019–2036. doi:10.1109/COMST.2014.2321628

Floridi, L. (2011). The informational nature of personal identity. *Minds & Machines*, *21*(4), 549–566. doi:10.1007/s11023-011-9259-6

Floridi, L. (2012). Turing's three philosophical lessons and the philosophy of information. *Philosophical Transactions of the Royal Society, 370* (1971)*, 3536–3542. doi:10.1098/rsta.2011.0325

Floridi, L. (2014a). Artificial agents and their moral nature. In P. Kroes & P. P. Verbeek (Eds.), *The moral status of technical artifacts* (pp. 185–212). Dordrecht, The Netherlands: Springer Science and Business Media. doi:10.1007/978-94-007-7914-3_11

Floridi, L. (2014b). *The fourth revolution: How the infosphere is reshaping human reality*. Oxford, UK: Oxford University Press.

Floridi, L. (2015). A proxy culture. *Philosophy & Technology, 28*(4), 487–490. Dordrecht, The Netherlands: Springer Science and Business Media. doi:10.1007/s13347-015-0209-8

Forelle, M., Howard, P., Monroy-Hernández, A., & Savage, S. (2015). Political bots and the manipulation of public opinion in Venezuela. Retrieved from http://arxiv.org/abs/1507.07109

Geiger, S. (2014). Bots, bespoke code and the materiality of software platforms. *Information, Communication & Society, 17*(3), 342–356. doi:10.1080/1369118X.2013.873069

Gillespie, T. (2014). The relevance of algorithms. In T. Gillespie, P. J. Boczkowski, & K. A. Foot (Eds.), *Media technologies: Essays on communication, materiality, and society* (pp. 167–194). Cambridge, MA: MIT Press.

Gillespie, T., Boczkowski, P. J., & Foot, K. A. (Eds.). (2014). *Media technologies: Essays on communication, materiality, and society*. Cambridge, MA: MIT Press.

Gleick, J. (2011). *The information: A history, a theory, a flood*. New York, NY: Pantheon Books.

González-Bailón, S. (2013). Social science in the era of big data. *Policy & Internet*, *5*, 147–160. doi:10.1002/1944-2866.POI328

Habit. (2016). In *Oxford English dictionary online*. Retrieved from http://www.oed.com/view/Entry/82980?rskey=cHuT6h&result=3#eid

Habitat. (2016). In *Oxford English dictionary online*. Retrieved from http://www.oed.com/view/Entry/82988?redirectedFrom=habitat#eid

Haraway, D. (2015). Anthropocene, Capitalocene, Plantationocene, Chthulucene: Making kin. *Environmental Humanities, 6*(1), 159–165. doi:10.1215/22011919-3615934

Haustein, S., Bowman, T. D., Holmberg, K., Tsou, A., Sugimoto, C. R., & Larivière, V. (2016). Tweets as impact indicators: Examining the implications of automated bot accounts on Twitter. *Journal of the Association for Information Science and Technology*, *67*(1), 232–238. doi:10.1002/asi.23456

Hayles, K. (2005). Speech, writing, code: Three worldviews. In N. K. Hayles (Ed.), *My mother was a computer: Digital subjects and literary texts* (pp. 39–62). Chicago, IL: University of Chicago Press.

Hidalgo, C. (2015). *Why information grows: The evolution of order, from atoms to economies*. New York, NY: Basic Books.

Howard, P. (2015). *Pax technica: How the Internet of things may set us free or lock us up.* New Haven, CT: Yale University Press.

Hwang, T., Pearce, I., & Nanis, M. (2012, March–April). Socialbots: Voices from the front. *Social Mediator,* pp. 38–45. doi:10.1145/2090150.2091061.

Hyde, M. J. (2004). Introduction: Rhetorically, we dwell. In M. J. Hyde (Ed.), *The ethos of rhetoric* (pp. 2–8). Columbia, SC: University of South Carolina Press.

Imperial ambitions. (2016, April). *The Economist.* Retrieved from http://www.economist.com/news/leaders/21696521-mark-zuckerberg-prepares-fight-dominance-next-era-computing-imperial-ambitions

Jamieson, K. H., & Cappella, J. N. (2010). *Echo chamber: Rush Limbaugh and the conservative media establishment*. Oxford, UK: Oxford University Press.

Kennedy, K. (2009). Textual machinery: Authorial agency and bot-written texts in Wikipedia. In M. Smith & B. Warnick (Eds.), *The responsibilities of rhetoric* (pp. 303–309). Long Grove, IL: Waveland.

Kramer, A., Guillory, J., & Hancock, J. (2015). Experimental evidence of massive-scale emotional contagion through social networks. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(24), 8788–8790. doi:10.1073/pnas.1320040111

Kuss, D., & Griffiths, M. (2011). Online social networking and addiction: A review of the psychological literature. *International Journal of Environmental Research and Public Health, 8*(9), 3528–3552. doi:10.3390/ijerph8093528

LaRose, R., Kim, J., & Peng, W. (2011). Social networking: Addictive, compulsive, problematic, or just another media habit?. In Z. Papacharissi (Ed.), *A networked self: Identity, community, and culture on social network sites* (pp. 59–82). New York, NY: Routledge.

Latour, B. (1987). *Science in action: How to follow scientists and engineers through society*. Cambridge, MA: Harvard University Press.

Leon, P. G., Blasé, U., Balebako, R., Cranor, L., Shay, R., & Wang, Y. (2012). Why Johnny can't opt out: A usability evaluation of tools to limit online behavior advertising. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 589–598). New York, NY: Association for Computing Machinery. doi:10.1145/2207676.2207759

Leonard, A. (1998). *Bots: The origin of new species*. New York, NY: Penguin.

Messias, J., Schmidt, L., Oliveira, R., & Benevenuto, F. (2013). You followed my bot! Transforming robots into influential users on Twitter. *First Monday, 18*(7). Retrieved from http://firstmonday.org/ojs/index.php/fm/article/view/4217/3700#p4

Metzger, M., & Flanagin, A. (2013). Credibility and trust of information in online environments: The use of cognitive heuristics. *Journal of Pragmatics, 59*, 210–220. doi:10.1016/j.pragma.2013.07.012

Miller, C. (2001). Writing in a culture of simulation: *Ethos* online. In P. Coppock (Ed.), *The semiotics of writing: Transdisciplinary perspectives on the technology of writing* (pp. 58–83)*.* Oakville, CT: David Brown.

Miller, C. (2007). What can automation tell us about agency? *Rhetoric Society Quarterly, 37*(2), 137–157. doi:10.1080/02773940601021197

Mitcham, C. (2014). Agency in humans and in artifacts: A contested discourse. In P. Kroes & P. P. Verbeek (Eds.), *The moral status of technical artifacts* (pp. 11–29). Dordrecht, The Netherlands: Springer Science and Business Media. doi:10.1007/978-94-007-7914-3_2

Morton, T. (2010). *The ecological thought*. Cambridge, MA: Harvard University Press.

Nanis, M., Pearce, I., & Hwang, T. (2011). Pacific Social Architecting Corporation: Field test report. Retrieved from https://www.academia.edu/2169112/Pacific_social_architecting _corporation_Field_test_report

Neff, G., & Nafus, D. (2016). *Self-tracking*. Cambridge, MA: MIT Press.

Papacharissi, Z. (2009). The virtual geographies of social networks: A comparative analysis of Facebook, LinkedIn, and ASmallWorld. *New Media & Society*, *11*(1–2), 199–220. doi:10.1177/1461444808099577

Papacharissi, Z. (Ed.). (2011). *A networked self: Identity, community, and culture on social network sites*. New York, NY: Routledge.

Papacharissi, Z. (2012). Without you, I'm nothing: Performances of the self on Twitter. *International Journal of Communication, 6*. Retrieved from http://ijoc.org/index.php/ijoc/article/view/1484

Park, J., Baek, M. Y., & Cha, M. (2014). Cross-cultural comparison of nonverbal cues in emoticons on Twitter: Evidence from big data analysis. *Journal of Communication*, *63*(2), 333–354. doi:10.1111/jcom.12086

Pask, G. (1969). The architectural relevance of cybernetics. *Architectural Design, 39*(September), 494–496.

Perrin, A. (2015). Social media usage: 2005–2015. Retrieved from http://www.pewinternet.org/2015/ 10/08/social-networking-usage-2005-2015/

Pickering, A. (2006). Ontology engines. In E. Selinger (Ed.), *Postphenomenology: A critical companion to Idhe* (pp. 211–220). Albany, NY: State of University of New York Press.

Raice, S. (2011, September 19). LinkedIn CEO tries to look past valuation, take long view. *The Wall Street Journal*. Retrieved from http://online.wsj.com/article/ SB10001424053111903895904576546700522834810.html

Ratkiewicz, J., Conover, M. D., Meiss, M., Goncalves, B., Flammini, A., & Menczer, F. M. (2011). Detecting and tracking political abuse in social media. In *Proceedings of the Fifth International AAAI Conference on Weblogs and Social Media* (pp. 297–304). Palo Alto, CA: AAAI Press.

Rickert, T. (2013). *Ambient rhetoric: The attunements of rhetorical being*. Pittsburgh, PA: University of Pittsburgh Press.

Rodriguez-Gomez, R., Marcia-Fernandez, G., & Garcia-Teodoro, P. (2013). Survey and taxonomy of botnet research through life-cycle. *ACM Computing Surveys, 45*(4), 1–33. doi:10.1145/2501654.2501659

Romero, D. M., Galuba, W., Asur, S., & Huberman, B. A. (2011). Influence and passivity in social media. In *Proceedings of the European Conference on Machine Learning and Knowledge Discovery in Databases* (pp. 18–33). Berlin, Germany: Springer.

Steinfield, C. N., Ellison, N., & Lampe, C. (2008). Social capital, self-esteem, and use of online social network sites: A longitudinal analysis. *Journal of Applied Developmental Psychology, 29*(6), 435–445. doi:10.1016/j.appdev.2008.07.002

Subrahmanian, V. S., Azaria, A., Durst, S., Kagan, V., Galstyan, A., Lerman, K., ... & Menczer, F. (2016). The DARPA Twitter Bot Challenge. *Computer, 49*(6), 38-46. Piscataway, NJ: Institute of Electrical and Electronics Engineers. doi:10.1109/MC.2016.183

Turkle, S. (1995). *Identity on the screen.* New York, NY: Simon & Schuster.

Verbeek, P. P. (2005). *What things do: Philosophical reflections on technology, agency, and design.* Philadelphia, PA: University of Pennsylvania Press.

Wagner, C., Mitter, S., Korner, C., & Strohmaier, M. (2012). When social bots attack: Modeling susceptibility of users in online social networks. *WWW '12 Workshops: Making Sense of Microposts*, *2*(3), 41–48. doi:10.1.1.221.6121

Wald, R., Khoshgoftaar, T. M., Napolitano, A., & Sumner, C. (2013). Predicting susceptibility to social bots on Twitter. In *Proceedings of the 2013 IEEE 14th International Conference on Information Reuse and Integration* (pp. 6–13). Piscataway, NJ: Institute of Electrical and Electronics Engineers. doi:10.1109/IRI.2013.6642447

Walther, J. (1996). Computer-mediated communication: Impersonal, interpersonal, and hyperpersonal interaction. *Communication Research*, *23*(1), 3–43. doi:10.1177/009365096023001001

Wang, G., Konolige, T., Wilson, C., Wang, X., Zheng, H., & Zhao, B. (2013). You are how you click: Clickstream analysis for Sybil detection. In *22nd USENIX Security Symposium* (pp. 241–256). Berkeley, CA: USENIX. Retrieved from https://www.usenix.org/conference/usenixsecurity13/technical-sessions/presentation/wang

Wang, G., Mohanlal, M., Wilson, C., Wang, X., Metzger, M., Zheng, H., & Zhao, B. (2013). Social Turing tests: Crowdsourcing Sybil detection. Retrieved from http://arxiv.org/abs/1205.3856

Westerman, D. W., Spence, P. R., & Van Der Heide, B. (2012). A social network as information: The effect of system generated reports of connectedness on credibility and health care information on Twitter. *Computers in Human Behavior, 28*(1)*, 199–206. doi:10.1016/j.chb.2011.09.001

Wiener, N. (1950). *The human use of human beings: Cybernetics and society*. New York, NY: Avon Books.

Wilson, R., Gosling, S., & Graham, L. (2012). A review of Facebook research in the social sciences. *Psychological Science, 7*(3), 203–220. doi:10.1177/1745691612442904

Woolley, S. C. (2016). Automating power: Social bot interference in global politics. *First Monday*, *21*(4). Retrieved from http://firstmonday.org/ojs/index.php/fm/article/view/6161/5300 doi:10.5210/fm.v21i4.6161

Woolley, S. C., & Howard, P. (2014, December 10). Bad news bots: How civil society can combat automated online propaganda. Retrieved from http://techpresident.com/news/25374/bad-news-bots-how-civil-society-can-combat-automated-online-propaganda

Xie, Y., Yu, F., Ke, Q., Abadi, M., Gillum, E., Vitaldevaria, K., ... Morley, M. Z. (2012). Innocent by association: Early recognition of legitimate users. In *Proceedings of the 2012 ACM Conference on Computer and Communications* (pp. 353–364). New York, NY: Association for Computing Machinery. doi:10.1.1.299.129

Zangerie, E., & Specht, G. (2014). Sorry, I was hacked: A classification of compromised Twitter accounts. In *Proceedings of the 29th Annual ACM Symposium on Applied Computing* (pp. 587–593). New York, NY: Association for Computing Machinery. doi:10.1145/2554850.2554894.

Zeifman, I. (2015, December 9). Bot traffic report: Humans take back the Web, bad bots not giving any ground. Retrieved from https://www.incapsula.com/blog/bot-traffic-report-2015.html