Somewhere on Twitter there are two automated accounts that I created a few months ago. Their names are SorryBot (https://twitter.com/BotSoSorry) and PhilosophyBot (https://twitter.com/SoPhilosophyBot) and one day they’ll become the leading activists in a fully automated social media project called #turingforthemasses. Their interaction will be automatic, without any human intervention; trying to raise awareness for the underlying problems of automated social media by tweeting what’s on their robot minds.

The fundamental aim of the project is to sensitize Twitter users for both the existence and the functional scope of social bots. It conveys a simple story that challenges the current, predominantly negative simplifications, and is based on two assumptions:
1. Not all social bots are evil and dangerous.
2. The indistinguishability of humans and bots is partly a product of the new ways and means of communication that have evolved in social media.
Automated behavior on social media has recently caused a lot of controversy. Scientists and journalists alike have investigated and commented on the potential risks of social bots, an allegedly malicious subspecies of a bigger group that’s simply called bots or software robots. They’re being accused of distorting the political discourse or manipulating online trends in social networks, but until this very day it’s hard to tell what their actual impact is or could be in the near future. It is exactly at this moment of uncertainty where SorryBot and PhilosophyBot intervene as mediators. Explaining to us how our own social media behavior facilitates the existence of social bots in the first place, and why they’re not all evil per se.

C O M B A T D U T Y


Fig. 1. We are failing the Turing Test for the masses.
An example of a bot that induced the bad reputation of bots in general is Tay AI (http://knowyourmeme.com/memes/sites/tay-ai), a social media bot built by Microsoft that eventually had to be switched off after being turned into a full-grown troll by 4chan’s /pol/ in 2016. Another one is the Random Darknet Shopper (https://wwwwwwwwwwwwwwwwwwwwwww.bitnik.org/r/) by Mediengruppe Bitnik: a bot that automatically ordered random articles from the deepweb market Agora. Regardless of their line of work and their place of activity, what these bots have in common is that they are semi-autonomous computer programs: software, algorithms, scripts. Their name derives from the Czech word *robota* which can be translated to ‘combat duty’. In the style of their mechanical ancestors, these software-based robots are used to automate processes in well-defined contexts.

Given their manifold functions the use of the word bot is not always very distinct, says German bot researcher Simon Hegelich in the article ‘Invasion Der Meinungs-Roboter’. Sometimes it refers to programs designed for search engines to scrape the web. Other times it is used in reference to a group of computers that have been infected with malicious software in order to create a bot-network. To prevent such a lack of definition there are certain prefixes and modifiers that make it easier to differ between the distinctive uses and capabilities of semi-autonomous software. Microsoft’s Tay AI for example can also be called a chatbot (or chatterbot). Bitnik’s Random Darknet Shopper, on the other hand, would be more of a web scraper that randomly spends Bitcoin at the same time. And a rather new generation of bots, called social bots, was designed to mimic human behavior in social media.

*Identifying the value of these different kinds of bots requires a theory of power. We may want to think that search engines are good, while fake-like bots are bad, but both enable the designer of the bots to profit economically and socially*

-DANAH BOYD, ‘WHAT IS THE VALUE OF A BOT?’
Along with the manifold functions of bots their public perception changes, altering between good and evil. Good bots are those that enable us to search the web or handle laborious tasks on Wikipedia. We hardly ever recognize them, and if we do, they are thoughtfully labeled as the tireless workhorses they are, engaging in meaningful work. Social bots, in contrast, fall into a different category. According to many definitions they’re active on social media platforms exclusively, where they mimic human behavior with the intent to influence their human surroundings. They can be seen as the new generation of spambots, operating on online social networks. Their predecessors were rather clumsy when it came to interacting with humans and therefore easy to unmask. But the new generation has improved to a degree where the imitation of human behavior has become more convincing than ever. Not only are these new bots able to circumvent the security systems of platform operators, but they outsmart the heuristics of platform users as well. At first glance, they are barely distinguishable from human networkers. And this is exactly why academics, politicians, and journalists attach an urgent risk potential to their existence.

ILLITERACY

A search for the hashtag #votetrump2016 (https://twitter.com/search?q=%23votetrump2016&src=typd) on Twitter turns up results that are both funny and sad. There are Twitter accounts that are still engaged in an election battle that IRL has ended over a year ago. Accounts like @amrightnow (https://twitter.com/amrightnow), with timelines that resemble time machines, are some of the more obvious reasons for scientists to believe that the majority of all contemporary social bots is based on rather simple software. These bots do not learn from their surroundings and are therefore locked in their respective functional range. From a macro perspective, writes Hegelich, their behavior can be summarized as the amplification, infiltration, and manipulation of beliefs and trends. They infiltrate social networks with fake accounts and amplify given beliefs by automatically producing huge amounts of likes, shares, and postings that feature specific hashtags and keywords. This in turn can lead
to a manipulation of algorithmically identified trends. Moreover, it might
convince social network users that a given topic or position is a relevant part of
online discourse when really it isn’t: they’re just being trolled.
American Right Now @amrightnow · 6 Nov 2017

Hillary Clinton should be in Jail Not on the campaign Trail #realdonaldtrump #military #army #navy #usmc

Government declares 22 Clinton emails 'top secret'

The Obama administration confirmed for the first time Friday that Hillary Clinton's unsecured home server contained some of the U.S. government's most closely guarded secrets, censoring 22 emails with material demanding one of the highest...
Due to accounts like @amrightnow – and especially due to increased reporting on the topic – social bots have sparked quite a panic. Their reputation as invisible troublemakers, tireless Trump supporters, and automated Russian trolls is well earned, of course, but also is a product of a rather one-sided news coverage. In consequence, bots are being mystified in public discourse and are no longer discussed as the primarily technical products and tools that they really are. To the layman they’ve thus become a threat in at least two distinct ways: (1) they are seen as frenetic opinion makers and agents of a foreign force; (2) whose underlying technical principles most of us cannot understand. This kind of framing makes it all the easier for the public to think of social bots as a new nemesis on the web and to suspect automated hounding whenever we’re confronted with dissent in our online social networks. When it comes to social bots, we must therefore attest a certain illiteracy that impedes the coexistence of humans and social bots.
Now, if some bots are a potential threat, why are they allowed in the first place? According to the Twitter rules on automation (https://help.twitter.com/de/rules-and-policies/twitter-automation) not all bots are bad bots. The company tries to punish only those that molest users by sending automated private messages and other kinds of spam. Helpful bots, on the other hand, that increase the general user experience on Twitter, are being encouraged by the platform service and are therefore free to operate on the network. So, not only is there differentiation between humans and bots but between wanted and unwanted bots as well. And as it goes with contemporary societal problems, there is a technical perspective, too.

The scientific bot discourse is currently being dominated by publications that are part of a larger attempt, called bot detection or bot security, respectively. The goal is to develop instruments that allow us to detect unwanted bots and render them harmless. But although a multitude of methods and frameworks has already been developed, scientists are still struggling to come up with a long-standing solution. They’re in the middle of an arms race where both scientists and social network operators constantly fall behind. The reason for this is simple: you can’t fight what you don’t know. Meaning that in order to overcome or even just study the latest malicious social bot software, it must be active first and you have to detect it.

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.

-DOUGLAS GUILBEAULT, GROWING BOT SECURITY: AN ECOLOGICAL VIEW OF BOT AGENCY
It is this fundamental problem that inspired the Canadian researcher Douglas Guilbeault to observe the relationship of social bots and humans from a new angle. In Growing Bot Security: An Ecological View of Bot Agency (https://www.asc.upenn.edu/news-events/publications/growing-bot-security-ecological-view-bot-agency) he examines the underlying problems of bot detection in social networks from a rhetorical point of view, with a strong focus on their environment, their habitat.

The basic structure of his argument can be called an ecological theory of agency. It derives from Aristotle’s theory on political agency and particularly from the concept of ethos. Ethos refers to the character of an individual and is one of three modes of persuasion alongside logos (argument) and pathos (emotion). Back when these thoughts were first formulated, the self-portrayal of a public speaker would be judged by the extent to which he mastered the art of conveying a credible character. These clues were not only verbal, but non- and paraverbal too, and therefore had a physical dimension to them.

*Every movement, every click, every utterance is recordable as an act of self-construction in the age of big data (...). 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.*

-DOUGLAS GUILBEAULT, GROWING BOT SECURITY: AN ECOLOGICAL VIEW OF BOT AGENCY

Today, in times of social media networks, new environments for rhetorical interaction have emerged. Not only are they not physical anymore, they have also established new rules of self-construction and interaction in the shape of profile pages, quantified popularity measurement, and automated communication tools. Properties that once defined a human rhetorical agent are no longer tied to real humans but have been implemented into web services and their graphical user interfaces. Consequently, everyone and *everything* that
can operate these interfaces successfully conveys a credible character at first
glance. Be it a human or a bot. Therefore, Guilbeault concludes, the
indistinguishability of humans and bots online is not so much a consequence of
sophisticated bot software, but rather a side effect of strategic interface designs
in social networks.

Before going into more depth, let’s just note what this comes down to. First,
we’re no longer just among ourselves online (if we’ve ever been at all). Second,
we should act accordingly, and be aware of our surroundings. For example,
@amrightnow, the Trump-loving bot mentioned earlier, at first sight could look
like a regular account with lots of tweets, followers, and likes. The façade works
and only collapses upon closer inspection, when you notice that all the
mentions, hashtags, and pictures just keep repeating. Moreover, the account is
active every day and produces almost the same amount of tweets within every
24 hours. It soon becomes clear that this account is actually running on
software. Rather primitive software, if we might say so. So how could this bot
have fooled us in the first place?
Stop and Frisk @stopandfrisk · 2h
04/27/11: Police use force to stop a 47-year-old in the Bronx, citing “suspect acting as a lookout.” No weapon is found.

Stop and Frisk @stopandfrisk · 2h
04/27/11: Police stop a 24-year-old in the Bronx, citing “casing a victim or location.” No weapon is found.

Stop and Frisk @stopandfrisk · 2h
04/27/11: Police stop a 40-year-old in Queens, citing “other.” No weapon is found.

Stop and Frisk @stopandfrisk · 2h
04/27/11: Police stop a 26-year-old in Staten Island, citing “casing a victim or location.” No weapon is found.

Stop and Frisk @stopandfrisk · 3h
04/27/11: Police stop a 20-year-old in Brooklyn, citing “furtive movements.” No weapon is found.

Stop and Frisk @stopandfrisk · 3h
04/27/11: Police stop a 20-year-old in Brooklyn, citing “casing a victim or location.” No weapon is found.
According to Guilbeault there are three major flaws that social bots exploit in social networks: profile settings, popularity measures, and automated communication tools. The first flaw addresses the use of personal profiles, representing ready-made sets of self-projection and identity creation. Personal profiles are where pictures, biographical data, user activities, and interactions with our friends merge into a uniform design. They help us take shape in virtual environments and can easily be exploited by social bots for the very same reason in order to convey a credible character. As another team of researchers has put it, the personal profile can be viewed as a bot’s face, whereas the actual code that it runs on is referred to as its brain.
But not only do these profiles help social bots to look human, they also enable them to act human. Meaning that bots can be programmed to scrape real data from real users in order to classify them and either imitate them or select a strategy of how to best approach them in a personal message.
The second flaw are popularity measurements in social networks. These measurements include all features that enable platform operators to quantify the social status of a given user, i.e. how many friends they have and how individual groups are connected in social graphs. We unconsciously rely on this kind of information in order to determine if a friend request from a stranger is legit or just spam. We take a brief look at their accounts and friends, and often
accept their request if we have mutual friends. In a sense, we trust our friends to choose their friends more wisely than we do and rely on simple heuristics to save ourselves some time and effort. This kind of behavior, the triadic closure principle, has already been investigated in the 50s and has been extended to online social interaction as well.

The probability that such a friend request is accepted, is up to three times higher if two networkers have mutual friends. Boshmaf and his fellow researchers infiltrated Facebook with a small bot army in 2011 and actively implemented this kind of knowledge when they designed their software. Not only did their bots perform well in terms of being accepted as friends by online networkers, the research team also witnessed how network users that mistook their bots for being real humans, proactively sent them friend requests.

The third and final flaw is embedded in our automated tools of communication. This includes all forms of like buttons, emojis, following functions, and the like, that have become constant features of user interface design in social networks. These tools often lack a real verbal dimension and can therefore easily be operated by social bots in order to interact with their environment. One of the most famous tools in this regard is Facebook’s like button. When it was publicly announced in 2009 (https://www.facebook.com/notes/facebook/i-like-this/53024537130) after two years of development it was promoted as a tool for immediate feedback. No longer would you have to write a comment to tell you

---

*It has been widely observed that if two users have a mutual connection in common, then there is an increased likelihood that they become connected themselves in the future. This property is known as the triadic closure principle, which originates from real-life social networks.*

friends that you liked their posts. Eight years later, it’s exactly these minimal mechanisms of communication that both services like Facebook and Twitter, and developers of social bot software profit from.

BEING SOCIAL

Why are these essentially malicious bots called *social* bots in the first place? In order to understand this we can go back to Guilbeault’s environmental view on the beginning of the social web as we know it today. In an article on design patterns and business models for the web 2.0, the American publisher and developer Tim O’Reilly describes an inbuilt architecture of participation as the key to a new web. New era services, he concludes, will be intelligent data brokers that harness the power of the users.

There’s an implicit “architecture of participation”, a built-in ethic of cooperation, in which the service acts primarily as an intelligent broker, connecting the edges to each other and harnessing the power of the users themselves.

-TIM O’REILLY, WHAT IS WEB 2.0: DESIGN PATTERNS AND BUSINESS MODELS FOR THE NEXT GENERATION OF SOFTWARE

The earliest version of this article dates back to September 2005. This was a time when access to Facebook was still restricted to US educational institutions and Twitter wasn’t even invented. A time when animated snowfall on customized Myspace profile pages would bring your Windows XP PC to its knees while we made new friends online. All nostalgia aside, what’s most astonishing about the O’Reilly quote is that it’s not addressing the social aspects of the new web. It’s rather a technical description that was inspired by his observation of the BitTorrent protocol. This technical underlying, however, has since transformed our social life to a degree where the concept of participation
is no longer negotiable. Today it is an integral part of a social network’s structure and thus action-guiding for all users of a given platform, and even for non-users.

Throughout the last decade, our means of participation online have become more and more standardized. Participation now requires user profiles, is meticulously logged, amplified, and managed via automated communication tools. As said, it is no longer subject to negotiation but has taken on a life of its own in the user interfaces of Facebook, Twitter, Instagram, and similar services. Everything that happens there is now social by default, and genuine human input, as the bare existence social bots demonstrates, is no longer a prerequisite. Thus, the modifier ‘social’ in social bots is first and foremost a description of the architecture of a technical environment. Today, our social lives have a technical foundation and social bots are the products of this shift.

Now, if the term social refers to the habitat of these new technical agents, then all other social media bots must be social bots as well. Indicating that previous definitions of social bots have been obstructing our view and prevented us from recognizing a bigger development, that doesn’t have to be completely negative. Above all, it now becomes easier to acknowledge that social bots are more than just automated Trump supporters. Even though publications in the field of bot security usually tend to reduce them to a side note, there are indeed social bots that do engage in meaningful work. In the shape of activist watch bots (https://points.datasociety.net/a-brief-survey-of-journalistic-twitter-bot-projects-109204a8d585) or as producers of generative art (https://points.datasociety.net/bots-a-definition-and-some-historical-threads-47738c8ab1ce), their work can be quite life-enhancing and sensitizing in artistic and political contexts.
Fig. 5. A list of Twitter bots on GitHub: who decides if they are good or evil?
There are two more arguments in favor of rethinking the definition of social bots and their past and future. In a private discussion Dutch media theorist Geert Lovink linked the existence of social bots to the discourses on automation throughout the 70s and 80s that eventually succumbed to the wider preoccupation with information technology. A time followed where automation was no longer an initial part of the overarching internet debate, until it resurfaced again in many different shapes, including social bots.

But what exactly are the origins of bots? Shall we think of them as a new generation of spambots or rather as chatbots with new means of communication? Are they conmen, barkers, guerilla advertising media, agitators, or arms in the realm of digital warfare? It seems as if the different technical, economic, and political implications of bots impede a shared interpretation of this new phenomenon. Thus, the question about the nature and the value of bots is and will remain a question of power.

In the future, says Lovink, bad bots won’t be an issue anymore. For Facebook, Twitter, Microsoft, and the like, bots are first and foremost interactive web services and the future of user interfaces. The future of bots will thus be a future of commercial soft-soapers: software products that will no longer stir controversy, but make us feel at ease. According to Simon Hegelich, the rising economic interest in particular will eventually produce a new form of thinking that no longer bans social bots to the shadows. The more they will be integrated into our everyday lives, the less surprising and disturbing their presence will become, because as they spread through the web, our awareness rises.

At the current stage, however, bots are still an unprecedented phenomenon with yet to be explored effects in regard to politics, economy, ethics, and art. While this might sound frightening, it really isn’t. Instead, think of it as a chance to actively get involved in the development of bots and the discourse that shapes their image. In this light, the social media project #turingforthemasses can very well be understood as a small attempt to push the debate to the next level.
In total, there are two Twitter bots that function as the project’s ambassadors. Both represent a distinctive voice that adds to the greater story. Together they investigate and explain their social media habitat by automatically generating an indefinite amount of tweets while using the hashtag #turingforthemasses as their unifying banner. The whole project has a strong emphasis on what can be called botness. The concept of botness yet lacks a clear definition, but it can more or less be understood as an attempt to grasp the nature of bots. The term appears in the Botifesto (https://points.datasociety.net/how-to-think-about-bots-1ccb6c396326), a blog post that was written by some bot enthusiasts and researchers during a workshop at the American research institute Data & Society. In another post (https://points.datasociety.net/our-friends-the-bots-34eb3276ab6d) about botness, the workshop participant Alexis Lloyd contemplates her somewhat hard to define relationship with a self-built slack bot:

I haven’t yet found the right words to characterize what this bot relationship feels like. It’s non-threatening, but doesn’t quite feel like a child or a pet. Yet it’s clearly not a peer either. A charming alien, perhaps? The notable aspect is that it doesn’t seem anthropomorphic or zoomorphic. It is very much a different kind of otherness, but one that has subjectivity and with which we can establish a relationship.

-AXEIS LLOYD, OUR FRIENDS, THE BOTS?

Being neither human nor pet, an interaction with social bots can really get you thinking. Whether they write poems (https://twitter.com/poem_exe), produce pictures (https://twitter.com/ArtyAbstract), or act as web archivists (https://twitter.com/search?q=%40wayback_exe&src=typd). The reception of their works often alternates between feelings of eeriness and deep affection, but
is always based on the very nature of a given specimen. Their randomized dissonances, ambiguities, and violations of linguistic as well as cultural rules can either be daunting or inviting.

Within the project there are multiple levels of botness at play. SorryBot is a design that reflects the classical subjectedness of machines. He embodies Asimov’s laws of robotics and has taken on the lifelong task of apologizing and protecting his human masters from the wrongdoings of his own kin. The second bot, PhilosophyBot, represents a more anthropomorphic design. He contemplates the indistinguishability of humans and bots in social media environments and tweets observations that apply to humans and bots alike.

![Fig. 6. PhilosophyBot: Contemplating the coexistence of bots and humans online.](http://networkcultures.org/longform/wp-content/uploads/sites/31/2018/03/etsiwah-fig7.jpg)
All things considered, #turingforthemasses is a call for participation. The bots and their hashtag are but a first contact point for those who’d like to learn more about social bots and the implications of their existence. They are supposed to motivate their human companions and to hand them the tools it takes to work out their own ideas. Be it tweets, bots, or new software tools – there’s plenty of ways that people can contribute, depending only on their imagination. Now more than ever it seems advisable to learn about bots, to recognize and to use them. If everyone would pay more attention to their online surroundings, not only would they be able to detect automated accounts by themselves and contribute to centralized security measures that are already in place. They would also come to recognize social bots as the diverse phenomenon they are and thus enable themselves to challenge both hypes and oversimplifications.

Bennet Etsiwah has studied Communication in Social and Economic Contexts at the University of the Arts, Berlin. Based on a broad understanding of Social Sciences and Strategic Design, he has worked in scientific research and commercial projects alike. Throughout the last years, he has turned into a passionate observer of the web with a strong focus on its countless oddities and the social implications of new media. As of late, he secretly dreams of becoming a web developer with a PhD.

References


