A holistic view on the ability of automated communication to influence technological scepticism in humans

Von der Fakultät für Ingenieurwissenschaften

Abteilung Informatik und Angewandte Kognitionswissenschaft

der Universität Duisburg-Essen

zur Erlangung des akademischen Grades

Doktor rerum politicarum
(Dr. rer. pol.)

genehmigte kumulative Dissertation

von

Florian Brachten
aus
Essen

Gutachter: Prof. Dr. Stefan Stieglitz
Gutachter: Prof. Dr. Stefan Morana
Tag der mündlichen Prüfung: 29.04.2021

Acknowledgments

What a hell of a ride that was – though not in a negative way. While the operative labour of attaining a PhD is mostly an individual feat, a work of that duration and magnitude on a spiritual level is never the achievement of just one person but in many aspects a group effort. I thus want to acknowledge the people that in the end helped me go through with it.

First of all, I'd like to thank my supervisor Prof. Stefan Stieglitz who showed trust in me by giving me the position in the first place, being patient and supportive throughout the process and, as another former PhD colleague already said fittingly, provided "as much freedom as possible and as few guidelines as necessary" (Kroll, 2020, p. II).

Furthermore, I want to thank all my colleagues (former and present) at the chair. While some may see work as Jazz without the music, I always enjoyed spending time with you and am thankful for all of your support and the different adventures we embarked on. I already miss the conferences and the retreats that weld us closer together and let us see behind the academic world.

Speaking of a world beyond academia: Finally, I want to express my gratitude to my friends and family in that world that kept me grounded and put things into perspective. To my parents and my brother who probably did not always understand what exactly was going on – during my PhD studies as well as in general in life – and were always on my side regardless and unconditionally, to my friends, current and those that got lost on the way but contributed invaluably:

Thank you!

Content

Figures

Tables

Abbreviations

ACT	Automated Communication Technology
AI	Artificial Intelligence
CASA	Computers are Social Actors
DT	Design Thinking
EB	Enterprise Bot
ICT	Information and Communication Technology
IS	Information Systems
NLP	Natural Language Processing
VA	Virtual Assistant

1 Introduction

1.1 Research context and Motivation

When new technology is introduced oftentimes it is at first regarded as being merely good or merely bad with little space in between (Wartella and Robb 2008). However, it is seldom that such simplified assessment is true and either pure good or evil comes from using a technology. A reason for such simplification may rather be the difficulty to fully grasp the magnitude its usage may eventually have as too many variables shape the outcome. Examples for such influencing factors may be the sector in which a technology is implemented (e.g. public or economy), its costs, its intricacy but also general opinion climate and other time-related factors like the zeitgeist or simply, momentum. While different actors (i.e. the public, politics and businesses) try to make sense of an emerging technology, it is the task of research to examine it, try and order the technology and give the needed assessment (Hinz et al. 2019), thereby oftentimes filling the space between the mere good and bad. A thorough assessment of the implication of adapting a technology on the one hand helps to shape an understanding of realistic positive usage scenarios apart from utopic conceptions, on the other assesses scepticism towards the technology to see if and when refraining from using it is reasonable and when it is inept. In other words, it helps people to make informed decisions about where and in what form its usage is appropriate and beneficial and where caveats are appropriate.

On a very general level, when evaluating options of either engaging with or avoiding a new or unfamiliar stimulus upon first encounter, when in doubt, escaping (e.g. by rejecting it) is normally the safer choice. The possible lost benefits from not choosing an unfamiliar option are less harmful than possible gained losses from choosing to engage with it and experiencing disadvantages. Following, the more common, because safer reaction towards something unfamiliar or unknown like a new technology is oftentimes its rejection (Kerschner and Ehlers 2016). In ancient Greece, Plato expressed concerns about written text and doubted its fitness to transport knowledge (Plato 1997). Examples from modern times include concerns, listening to radio may take up too much time of children and influence them in unknown und unpreferable ways (Gruenberg 1935) as well as similar concerns regarding Comic Books, TV or Video games (Orben 2020). Famously, also the critical philosophers of Frankfurt school within the essay on culture industry stated that "*the products themselves, especially the most characteristic, the sound film, cripple [the imagination and spontaneity in the consumer of*

culture today]" (Horkheimer and Adorno 2002, p. 100). Such scepticism stems from an unfamiliarity with the technology that leads to concerns "*about the negative social, environmental and economic impacts of technology [and] that technology can undermine social cohesion, foster individualisation and isolation, eliminate jobs and erode their meaning*" (Kerschner and Ehlers 2016, p. 147). While former scepticism, e.g. regarding radio usage may seem out of place today and most people won't be concerned with this technology anymore, current concerns exist for modern technologies such as smartphones, social media platforms or advancements in fields related to the application of Artificial Intelligence (AI) (Caldwell et al. 2020).

Against these reservations, it is important to keep in mind, that concerns regarding new technology are nothing that should be frowned upon. Concerns need to be taken seriously to be able to address them and not mindlessly embrace new technology for the sake of newness. Wherever new technology exists and fosters uncertainty, "*risk assessment is needed to determine the potential dangers of a particular technology*" (Kerschner and Ehlers 2016, p. 147) as well as pointing out the benefits its use has whereby every new technology warrants an individual inspection (cf. Chhonker et al. 2018). Considering consequences emerging from technological advancements and how those impact society as well as economy, themes that IS research is concerned and especially suited to deal with are digital transformation and AI, two fields that find themselves combined in the form of automated communication technologies (ACTs) (Janssen et al. 2020, Hinz et al. 2019).

ACTs exist in different shapes like chatbots, digital assistants or social bots which (e.g. regarding AI-based digital assistants) "*provide significant opportunities, but also might become a threat*" (Maedche et al. 2019, p. 535). The authors further state that the overall topical area "*with all its challenges and opportunities, is perfectly suited for IS research*" (2019, p. 539) which is reflected in current IS literature that examines different types of ACTs and, within these types, different criteria of their usage. Phenomena that are represented in research do have common grounds but are used in very different scenarios. Considering the aforementioned digital assistants, those are examined in the context of work and how they may be applied to reduce costs or increase efficiency by supporting employees (Quarteroni 2018). Furthermore, individual aspects such as the effect of response delays in the answers of chatbots are examined (Gnewuch et al. 2018). Besides this application in a working context, occurrences in more public domains are a different field of ACTs that is also investigated. Here, bot accounts on

social media sites that post autonomously – so called social bots – are explored, e.g. regarding their potential to unbeknownst influence human users (Ferrara et al. 2016). What becomes apparent is the very different nature of the mentioned ACTs even though they share a common base. And while the technology is continuously paving its way into different areas of everyday life a scientific assessment to determine their status quo is missing.

Research has become more sensitive to the importance to understand factors that play a role in accepting and avoiding automated communication. On similar fields (i.e., for AI and smart technologies) research on the adoption exists, however, for ACTs such as chatbots and social bots it is still sparse. As the societal adoption of every new technology is coupled with individual processes distinctive for that technology (Chhonker et al. 2018) ACTs warrant individual considerations. And as especially upon the emergence of a technology it is interpreted as either being good or bad for society with no middle ground, a more holistic view is needed that takes into account different arguments (Orben 2020). Coupled with the many studies examining individual aspects of chat- or social bots, a larger view is needed to fully assess the impact of the systems. Accordingly, Janssen et al. (2020) point out that "*most scientific studies today concentrate on particular aspects of chatbots, such as the personality of cognitive chatbots, technical capabilities or their specific application purpose without providing a holistic view*" (pp. 211–212). It is thus the aim of this thesis to provide such a view for ACTs by inspecting different contexts in which they are applied to determine a status quo and deriving implications for their usage.

1.2 Research Questions

Technological advancements within the last decade (e.g. within AI, NLP, availability of the internet or computational capacity) have brought forth a new and more capable generation of bots that need yet to be fully examined (Janssen et al. 2020). To do so, first it is necessary to get an understanding of how this new generation is currently applied and how this might foster scepticism. Thus, a field is required where a) there currently is a sufficient amount of contact between users and automated communication and b) these contacts are prone to evoke scepticism by the users. The former aspect ensures that a relevant and timely phenomenon is examined whereas the latter helps to get insights into aspects weakening the trust in automated communication. A form of automated communication that matches these criteria are bot accounts on social media platforms also called social bots.

These social bots are accounts on social media sites which are not easily recognisable as artificial and automatically engage in human-like communication behaviour, e.g. by sharing links, posting messages or liking posts. One of their reported mode of operation is that they act in networks of several hundred or thousand accounts that take synchronised action which heightens the probability of sufficient contact with users (Ferrara et al. 2016). Scientific publications assume that the bot accounts try to influence the opinion of human users, a negative depiction that was picked up by media away from scientific outlets (Beuth 2017; Seiffert 2016; Sickert 2017). This in turn led to a wider dissemination of the topic and also of the concerns related to possible negative effects of bot accounts on humans. In fact, the concerns went so far that the German Bundestag instated a public technical discussion of the Committee on Education, Research and Technology Assessment that addressed the topic[1] and political parties in Germany felt compelled to issues statements, promising they wouldn't use such technology in upcoming elections[2]. For the reasons mentioned before, social bots are fitting application case to assess the current state of automated communication and how it can foster scepticism. What is necessary here is a thorough examination of social bots and their effect on scepticism towards the newly emerging form of ACTs. The first research question is thus:

RQ1*: How is automated communication in the form of social bots currently applied so that technological scepticism is fostered?*

Addressing this question will not only help to untangle the overall discussion on social bots and their possible negative impact but also give insights into how automated communication is currently applied in practice. For a social debate on scepticism towards automated communication this builds the basis for further research. After it is shown how the technology fosters scepticism, it is equally important to examine how it may be actively applied in a supportive way. As the digitalisation and with it the automation of IS is advancing (Söllner et al. 2016), people need to become familiar with it to make confident and informed use of it. This won't happen if caveats exist opposing its use. Thus, it is important to also examine the potentials automated communication holds. Doing so will help twofold:

---

[1] https://www.bundestag.de/dokumente/textarchiv/2017/kw04-pa-bildung-forschung-social-bots-488818, last accessed 2020-10-20

[2] https://www.dw.com/de/etablierte-parteien-verzichten-auf-social-bots/a-36119916, last accessed 2020-11-28

First, it will offer insights into the actual benefits automated communication can provide. Instead of examining the effect of individual features (e.g. the impact of gender of a chatbot or the formality of its answers on the usage) a thorough look at the overall capabilities of the ACTs and thus a status quo of their current state can provide implications like more realistic application scenarios. These in turn can help guide research to more practical relevant studies. It thus answers calls that IS research should "*provide scientifically grounded and practice-relevant answers to the question how the interplays between users, tasks, and technologies*" (Maedche et al. 2019, p. 536) regarding digital assistants need to be shaped. As Janssen et al. (2020) furthermore point out, the current research on chatbots is rather dispersed and while many see potential benefits from their usage, a common understanding of the ACTs and their capabilities will help to bring guidance into the research stream, enabling more determined approaches going forward. Showing concrete benefits of its usage may further help to increase acceptance and get a clearer picture what the technology is capable of. Being able to a) understand the capability of automated communication and b) point to concrete findings for supportive scenarios will strengthen the argument to engage with automated communication instead of shying away from it.

Second, it will provide a thorough perspective on ACTs as a whole as it expands the previous view on scepticism towards automated communication by a counterpoint in the form of benefits that may be achieved by their usage. While it is important to know what fosters scepticism it is equally important to know what reduces it. This way a complete picture can be presented that neither exclusively demonises the technology nor idolizes it but faithfully shows the current state of scepticism towards automated communication and how it should be evaluated. Following, the second research question is:

RQ2*: How can automated communication in the enterprise context be actively applied to support humans and decrease the potential for technological scepticism?*

The following section presents the structure of the thesis and the research papers applied to answer the questions.

1.3 Thesis structure and List of Publications

This dissertation is a 'thesis by publication' in cumulative form meaning that it consists of several published, stand-alone research papers which are combined in a synopsis to answer the

research questions. The dissertation is thus divided into two parts, the first being the synopsis and the second the published research articles.

For the synopsis, chapter two explains the research background and introduces the most important constructs regarding automated communication, chatbots, social bots and technological scepticism. Chapter three then lays out the applied research strategies and presents the research methods used before chapter four summarises and reports the results from the research papers the thesis is based upon. The fifth chapter then draws on the results and uses them to answer the research questions and discusses them in a wider context. Finally, chapter six draws a short conclusion and presents limitations and an outlook for possible future research.

The synopsis is then followed by the 9 research articles the thesis consists of and that were published on international conferences as well as in journals. Table 1 lists the papers indexed in chronological order and reports their bibliographical data (i.e., title, authors, year of publication and outlet) as well as type of outlet (CNF = Conference; JNL = Journal). These data are followed by metrics ranking the outlet (i.e. the VHB-JOURQUAL3 ranking, the CORE ranking and the Journal Impact Factor (JIF)) and the citations count of the paper from Google Scholar.

Six of the papers are conference papers, two of which are short ('research in progress') papers and four are full papers. 3 papers are journal articles, of which two are published and one is currently in the third round of the review process. The conference papers were published on the *International Conference on Information Systems (ICIS), the European Conference on Information Systems (ECIS)* , the *Pacific Asia Conference on Information Systems (PACIS),* the *Australasian Conference on Information Systems (ACIS)* and the *International Conference on Human-Computer Interaction (HCII)*, the first two being the most prestigious conferences in the Information Systems community. The journal articles were published in the *European Journal of Information Systems (EJIS)* (where it won an award as the best publication of that year) and the journal *Information Systems and e- Business Management (ISEB).* One article is under review in the third round for publication in the *International Journal of Information Management (IJIM)*. All articles were published in English language and were co-authored by two or more researchers. Co-authors were affiliated with the Queensland University of

Technology (QUT), the University of Darmstadt, the Jacobs University Bremen and colleagues from the University of Duisburg-Essen.

Table 1. List of publications ordered for relevance in relation to research questions

#		Publication	Type	VHB JQ3[1]	CORE[2]	JIF	Citations[3]
1	Title:	Do Social Bots Dream of Electric Sheep? A Categorisation of Social Media Bot Accounts	CNF	/	/	/	65
	Authors:	Stieglitz, S., **Brachten, F.**, Ross, B. & Jung, A.					
	Year:	2017					
	Venue:	Australasian Conference on Information Systems (ACIS)					
2	Title:	Defining Bots in an Enterprise Context	CNF (Short)	A	A	/	12
	Authors:	Stieglitz, S., **Brachten, F.** & Kissmer, T.					
	Year:	2018					
	Venue:	International Conference on Information Systems (ICIS)					
3	Title:	Do Social Bots (Still) Act Different to Humans? – Comparing Metrics of Social Bots with Those of Humans	CNF	C	B	/	27
	Authors:	Stieglitz S., **Brachten F**., Berthelé D., Schlaus M., Venetopoulou C. & Veutgen D.					
	Year:	2017					
	Venue:	International Conference on Human-Computer Interaction (HCII)					
4	Title:	Strategies and Influence of Social Bots in a 2017 German state election	CNF	/	/	/	29
	Authors:	**Brachten, F.**, Stieglitz, S., Hofeditz, L., Kloppenborg, K. & Reimann, A.					
	Year:	2017					
	Venue:	Australasian Conference on Information Systems (ACIS)					
5	Title:	Social Bots in a Commercial Context – A Case Study on Soundcloud	CNF (Short)	B	A	/	8
	Authors:	Ross, B., **Brachten, F**., Stieglitz, S., Wikström, P., Moon, B., Münch, F. V. & Bruns, A.					
	Year:	2018					
	Venue:	European Conference on Information Systems (ECIS)					

6	Title:	Are social bots a real threat? An agent-based model of the spiral of silence to analyse the impact of manipulative actors in social networks	JNL	A	A	2.600 (2019)	46
	Authors:	Ross, B., Pilz, L., Cabrera, B., **Brachten, F.**, Neubaum, G. & Stieglitz, S.					
	Year:	2019					
	Venue:	European Journal of Information Systems (EJIS)					
7	Title:	The Acceptance of Chatbots in an Enterprise Context – A survey study	JNL	C	/	8.210 (2019)	/
	Authors:	**Brachten, F.**, Kissmer, T. & Stieglitz, S.					
	Year:	*Under review (4th round)*					
	Venue:	International Journal of Information Management (IJIM)					
8	Title:	Virtual Moderation Assistance-Creating Design Guidelines for Virtual Assistants Supporting Creative Workshops	CNF	C	A	/	24
	Authors:	Strohmann, T., Fischer, S., Siemon, D., **Brachten, F.**, Robra-Bissantz, S., Stieglitz, S. & Lattemann, C.					
	Year:	2018					
	Venue:	Pacific Asia Conference on Information Systems (PACIS)					
9	Title:	On the Ability of Virtual Agents to decrease Cognitive Load – An experimental study.	JNL	C	C	2.161 (2019)	7
	Authors:	**Brachten, F.**, Brünker, F., Frick, N., Ross, B. & Stieglitz, S.					
	Year:	2020					
	Venue:	Information Systems and e-Business Management (ISEB)					

[1] https://vhbonline.org/vhb4you/vhb-jourqual/vhb-jourqual-3/tabellen-zum-download (accessed 2021-02-20).

[2] http://www.core.edu.au/ (accessed 2021-02-20).

[3] https://scholar.google.de/citations?user=B-3dMacAAAAJ&hl=en (accessed 2021-04-30).

2 Research Background

2.1 Automated communication

While communication traditionally takes place face-to-face, an increasing mediatization of communication can be observed and be attributed to the increasing use of information technology, meaning that social communication in general is more and more mediated by and carried out over technology (Papsdorf 2015). Especially the internet plays an important role here as it also digitizes the communication thus making it accessible to IS and thereby opening it up to automation, e.g. for usage in social media (Papsdorf 2015). Automation in general is seen as the "*use [of] machines/computers/robots to execute or help execute physical operations, computational commands or tasks*" (Nof 2009, p. 43). Transferred to communication this means that automated communication is the use of computers or (software) robots to execute the tasks of communicating with human users. The term automated communication in general encompasses a multitude of technological implementations ranging from simple automated phone systems when calling a hotline to sophisticated physical robots to interact with, e. g. in healthcare for the elderly (Lugrin et al. 2019).

Automated communication can thus be subdivided regarding the focus of the used conversational interface (i.e. the systems front-end to the user) into *spoken dialogue systems*, *voice user interfaces*, *embodied conversational agents* and *chatbots* (McTear 2017). The aforementioned automated phone systems for example correspond to speech-based *voice user interfaces*. The embodiment of automated communication examined in this thesis however are text-based chatbots as particularly those have gained much application and interest lately (Gnewuch et al. 2017; McTear 2017) and are especially relevant in current IS research (Feine et al. 2019; Janssen et al. 2020). Accordingly, when referring to *automated communication* or *automated communication technologies* (ACTs) in this thesis, specifically chatbots are meant.

Tracing back the origin of modern research on ACTs, most publications agree that it started with the chatbot ELIZA developed and described by Weizenbaum in the 1960s (Feine et al. 2019; Gnewuch et al. 2017; McTear 2017; Weizenbaum 1966). While ELIZA, a chatbot that simulated a psychotherapist, was limited by today's standards, the title of Weizenbaum's article that introduced ELIZA described the aim of this branch of research very well: *Study of Natural Language Communication between Man and Machine* (Weizenbaum 1966). Although this

example shows that research on chatbots has some tradition, new developments within the last decade brought forth a new, more capable generation of chatbots and with it a new need for and interest in research on them (Feine et al. 2019). Factors that fostered these developments on the field of ACTs were advancements in underlying technologies such as AI and NLP. Those enabled and were boosted by better speech recognition and larger amounts of processing capabilities so that new application scenarios opened up as the ACTs now were capable enough to sufficiently address more sophisticated needs of users (McTear 2017). In earlier development stages of automated communication, the latter had been disappointed as those former generations simply could not live up to the expectations placed in them (Ben Mimoun et al. 2012; Moore 2013).

One aspect that immediately becomes apparent when dealing with the current line of research on automated communication is that there is no universal accepted terminology to address specific criteria of automated communication but rather individual preferences. Terms used in this context are bots (McTear 2017), virtual assistants (Shah et al. 2016), personal assistant agents (Palanca et al. 2018), virtual agents (Quarteroni 2018), vocal social agents (Guzman 2017), personal assistants (Han and Yang 2018) or conversational agents (Gnewuch et al. 2017; Janssen et al. 2020). Furthermore, while the two latter papers both use the same term, they differ in their interpretation of it which demonstrates that even when two publications use the same expression, they do not necessarily agree on a common understanding – an aspect also acknowledged by many of these research papers on automated communication. Following, one stream of research on automated communication specifically deals with defining and classifying the individual technological entities that are examined. Considering research that deals with the actual technology, apart from classifying it on a more meta level, two streams can be differentiated: 1) Research that examines the user's perception of chatbots and 2) research that deals with the actual application and effect of chatbots.

Findings on how humans react to interactive technology like chatbots can be based on the media equation theory by Reeves and Nass (1996) which includes the "*Computers are Social Actors*" (CASA) paradigm (Nass and Moon 2000). Within a series of studies on how humans respond to different media, the authors found that "*people were polite to computers. Not only were the computers in these experiments tools for learning new information, they were social actors that people reacted to with the same polite treatment that they would give another human*" (Reeves and Nass 1996, p. 26). While these findings originate from interactions with computers in

general, they also translate to the more specific realm of chatbots. Accordingly, chatbot-related research has picked up on them and oftentimes uses them as basis for more specific investigations (Feine et al. 2019). This more specific research has brought forth insights into how humans perceive chatbots and for example found that typefaces had an impact on the perceived humanness of chatbots (Candello et al. 2017) as do response delays in the answers of chatbots (Gnewuch et al. 2018) or the formality of language used (Sah and Peng 2015). Other examples include the readiness to disclose personal information depending on the self-disclosure of chatbots (Kang and Gratch 2011) or the effect of perceived gender of a chatbot on its evaluation by human users (Brahnam and De Angeli 2012).

As mentioned earlier, the internet not only played an important part in opening up communication for automation but also offers a large field where automated communication is applied – social media (Papsdorf 2015). Of the papers concerned with the actual application and effect of chatbots a large portion thus deals with so called social bots – automated accounts on social media platforms that act in a human way by posting, liking or sharing content and which are not immediately discernible as artificial. While there are beneficial scenarios where they are applied (e.g. by answering questions or simply posting jokes (Salto Martínez and Jacques-García 2012; Veale et al. 2015)) much of the research on social bots is concerned with their application in a political context where they are thought to try and influence the opinion of human users (Ferrara et al. 2016).

Examples for their application in politics are US senate elections (Mustafaraj and Metaxas 2010), the conflict between Russia and the Ukraine (Hegelich and Janetzko 2016), the 2016 US presidential election (Bessi and Ferrara 2016), the 2016 French presidential election (Ferrara 2017) or the Syrian civil war (Abokhodair et al. 2015). The main concern here is, that these social bots try to influence human users opinions through the sheer number of the messages they post but also by the number of other social bot accounts that likewise post the same message – a behaviour labelled astroturfing (Abokhodair et al. 2015). In accordance with the CASA paradigm, studies showed that in some cases humans did not differ in their assessment of a news' credibility depending on whether their origin was a human user or a social bot (Edwards et al. 2014). Also, on an individual level, social bots were able to influence the behaviour of human Twitter users when the former directly addressed (i.e. @-mentioned) the latter – a decrease in offensive messages by the human users could be observed (Munger 2017).

The example of social bots shows how a negative view on the technology may emerge. But also, the general application of chatbots in scenarios where they are thought to support human users may be prone to initial scepticism and caveats - as every technology that is newly introduced to society is either seen as merely good or bad (Wartella and Robb 2008).

2.2 Scepticism towards technology and automated communication

On a very basic level, technology supports humans by extending their abilities to interact with the environment and to solve problems. It enables feats that otherwise are not or only with greater effort accomplishable by humans – e.g. using a screwdriver to fasten a screw or using an oven to heat a meal. On a conceptional level it can thus be regarded as an extension of the self (Belk 1988). While at first glance this may be seen as a positive impact of technology, it also fosters scepticism, as it shows that technology is used to overcome (natural) limits which in turn some believe will result in either environmental, social or economic impacts that are unfavourable (Kallis et al. 2012; Kerschner and Ehlers 2016). According to Kerschner and Ehlers (2016) three subgroups of technological scepticism can be differentiated: a) *simple scepticism*, b) *technophobes* and c) *entropy pessimists*. While all of these subgroups at their core have reservations before using technology, their reasons differ as does the severity of reservation and the abstractness of possible usage consequences. Simple scepticism states that technology needs to prove itself and show that the negative consequences of its usage don't outweigh the positive. The second group argues that all technology is used to overcome limits which in turn leads to negative social or environmental impacts and finally the entropy pessimistic group states that all technology in the end consumes more energy than it produces so there is no possible sustainable application. While these views are somewhat conceptional to different degrees, one common aspect which is also more concrete is a possible negative impact on employment and society through a loss of jobs due to technical advancements. This fear is at the core of and amongst the oldest and best-established reservations against technology in general and its advancements per se (Frey and Osborne 2017).

Regarding interactions with technology, of which communication is one aspect, findings on the so-called uncanny valley describe an unease of humans when interacting with autonomous actors such as (physical) robots. According to the hypothesis, humans' affinity towards artificial actors such as robots increases with the human-likeness of these actors but takes a sudden dip (the so called uncanny valley) before perfect humanness is reached (Mori et al. 2012). While the representation of the relation in form of the aforementioned curve is disputed, the overall

relation between human-likeness as one factor and its likeability in general holds true (Rosenthal-von der Pütten and Krämer 2014). Considering possible reasons, the authors amongst others state "*evolutionary-biological oriented explanations addressing the fact that humans' sensitivity to perceptual mis-matches regarding fellow humans follows evolutionary exigencies like pathogen avoidance and mate selection"* (p.422).

Other concepts that encompass a sceptic view towards digital technology are technostress or the digital detox movement. Here, technology is seen as something overwhelming for a user, especially due to its complexity and ubiquity of modern information and communication technologies (ICTs) (Srivastava et al. 2015). Technostress describes stress due to increasing use of information technology particularly at the workplace while digital detox describes a more general trend towards a conscious (limited in time) abandonment of digital technology to detoxify the mind (Sutton 2017). As both terms (stress as well as toxicity) imply, technology here is seen as something that strains human users which is why, in the case of digital detox, those retreat from using it. However, it is important to note that for both concepts not technology itself is problematic but its omnipresence which leads to greater complexity. It is precisely this complexity however, that poses a problem when it comes to increasing automation as the latter goes along with the former (Lee and See 2004).

While automation of systems is thought to in the end simplify the life of its users, automated systems become more complicated and less transparent. This lack of transparency then leads to a lack of knowledge about these complex systems which in turn may lead to negative public attitudes towards technology in general and automation in particular (Kerschner and Ehlers 2016). In line with the previous findings that one of the oldest reservations towards technology is the fear of job loss, many people when asked to define automation, state (wrongfully as the author points out) that it "*partially or fully replace[s] human work*" (Nof 2009, p. 43). As social influence is an important aspect in convincing users to use a system (Dwivedi et al. 2019), the overall public opinion is not unimportant when attitude towards a new technology is build which indicates the importance to fully understand the public opinion formation on automation and automated communication.

In line, findings from research on AI tools in the enterprise context, a field related to that of automated communication, show that mainly social hurdles such as a lack of employee acceptance, again due to the fear of job, loss are among the main challenges for the

implementation of these tools (Schlögl et al. 2019). However, as ACTs are projected to become "*a key element in the future of work*" (Maedche et al. 2019, p. 535) familiarity with them is crucial to be able to compete on the working market. This familiarity does not only originate in using such systems at the workplace but also from private usage of akin technology and a general open attitude towards it. Furthermore, a fear of job loss from employees leads to a lack of acceptance which may in turn lower the job satisfaction if systems are implemented regardless (Schlögl et al. 2019). Following it is also in the interest of enterprises to not alienate their employees when introducing these systems.

Addressing factors that may lead to the active avoidance of specifically automated communication technology, Christy et al. (2019) state that a lack of sensitivity or capabilities in ACTs may lead to frustration with the users which then refrain from using them. However, as the authors point out, automated communication has *"the potential to significantly influence the speed and efficiency of information systems if […] adopted successfully by the target population"* (p. 32) which is why understanding factors that lead to the avoidance but also adoption of these systems is crucial.

3 Research Design

In the following the research strategy and the applied research methods are described. As this is a cumulative dissertation, it is first explained how the individual research papers are connected to each other and to the research questions. Afterwards the applied methods are summarised.

3.1 Research Strategy

Figure 1depicts the relationship of the papers and their meaning regarding the aim of the thesis. Papers P1 and P2 build the foundation from which the questions are answered. They categorise and define terminologies on which the latter papers draw and help to initially delimit different domains from each other.

Four papers (P3, P4, P5 and P6) are then used to answer RQ1. These papers all deal with the phenomenon of Social Bots and examine different facets, i.e. their application in a global political event to point out differences in their behaviour compared to that of human users (P3), their application in a regional political event (P4) and the application in a commercial context (P5). Furthermore, P6 deals with the potential impact Social Bots can have on the opinion forming process.

Albeit to a lesser extent, P6 is also related to RQ2 as it examines the possible impact of automated communication. Three papers mainly address RQ2 and deal with just that – effectivity of automated communication for active usage by humans. Besides their overall acceptance in the enterprise context (P7), the papers show their potential to be applied in support of creative workshops (P8) and their ability to decrease the cognitive load of human users (P9).

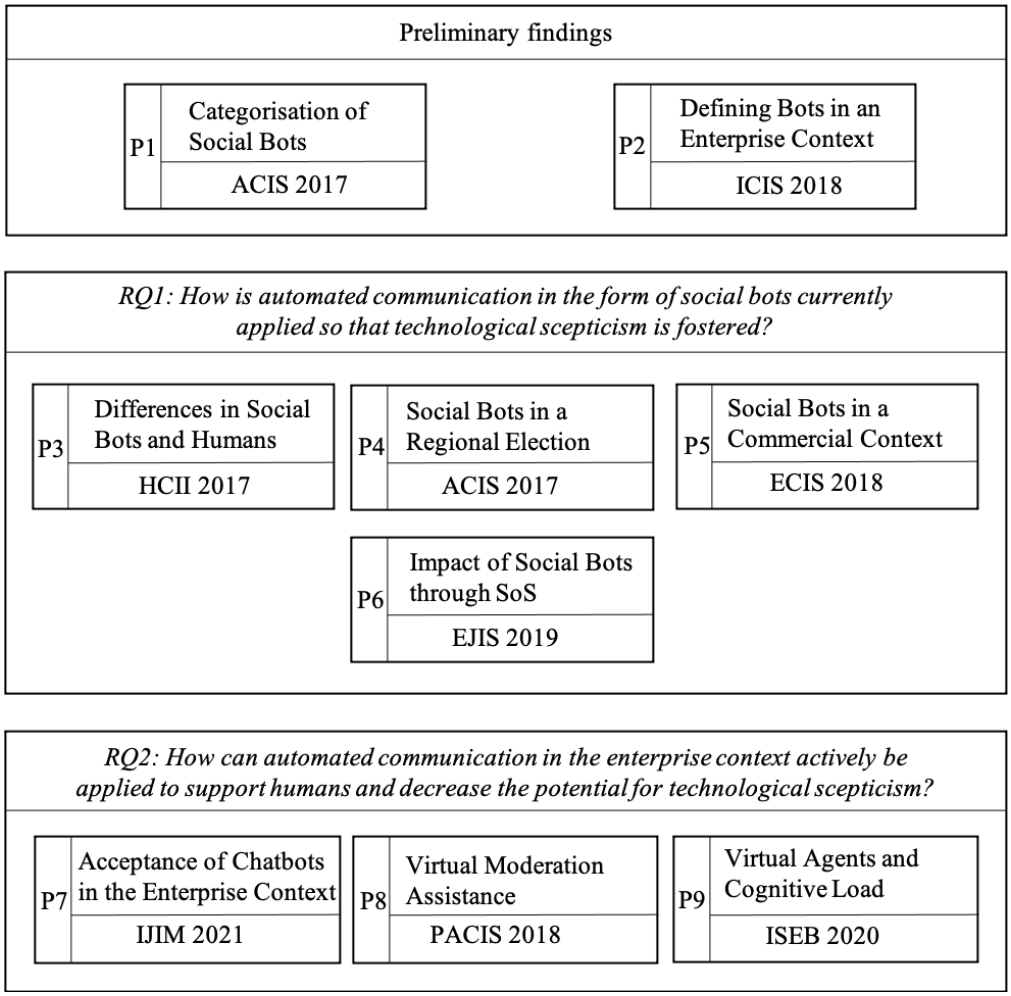

Figure 1. Overview and classification of the research paper

3.2 Applied Research Methods

Table 2 summarizes the research methods applied throughout the papers used in this dissertation. The papers drew on data collected from academic literature databases, interviews, surveys, experiments, simulations and the APIs of social media platforms. Analyses were of quantitative (i.e. statistical analysis and summaries) as well as qualitative (i.e. content analyses and literature reviews) nature.

As literature reviews help "to understand a domain, to uncover gaps in research and develop an agenda for [further] research" (vom Brocke et al. 2015, p. 208) this method is suited to serve as foundation for further research. Accordingly, paper P1 contains a systematic literature review regarding publications on social media bot accounts to derive a concrete understanding of the term. P2 also relies on a literature review to define bots in an enterprise context. It furthermore also uses quantitative data from an online survey to validate the definition.

To observe social bots in the field, i.e. on social media platforms, data from these platforms needed to be collected and analysed. As social media analytics "*deals with methods of analysing social media data*" (Stieglitz et al. 2018, p. 157), a social media analytics approach was used in papers P3, P4 and P5. They dealt with behaviour of social bots and drew on actual usage data from social media platforms to analyse their behaviour. Paper P6 took a different approach and used a simulation via an agent-based model, as this paper examined an aspect of social bots not directly observable in practice. Paper P7 studied the acceptance of chatbots by humans and used an online survey to gather the data and a regression model to analyse those whereas P8 used expert interviews to get an assessment on the fitness of bots for certain tasks grounded in experience. Paper P9 again used quantitative data from a survey combined with performance data from an experiment which were evaluated via statistical methods.

Table 2. Overview of Applied Research Designs

Paper	Research approach	Data collection method	Data analysis method
P1	Literature review	Academic literature databases	Qualitative summary
P2	Literature review and Survey study	Academic literature databases and Online-Survey	Qualitative & Quantitative summary
P3	Social Media Analytics	Twitter-API (Social Media Platform)	Quantitative summary
P4	Social Media Analytics	Twitter-API (Social Media Platform)	Quantitative summary
P5	Social Media Analytics	SoundCloud-API (Social Media Platform)	Quantitative summary & Network Analysis
P6	Virtual experiment	Simulation of an agent-based model	Quantitative summary & Network Analysis
P7	Survey study	Online-Survey	Regression model
P8	Qualitative research	Interviews	Content analysis & qualitative summary
P9	Laboratory experiment	Survey in combination with performance data from a laboratory experiment	Statistical analysis

4 Research Results

This chapter presents the findings from the individual research papers. It is subdivided according to the structure laid out in chapter 3.1 with each section being related to an overarching research goal: Section 4.1 will first present preliminary results regarding the classification and categorisation of bots in different contexts. This builds the basis for sections 4.2 and 4.3 which present the results regarding automated communication in the form of social bots and automated communication actively applied by users for their benefit respectively. In the following, Table 3 summarises the findings from each paper.

Table 3. Summary of the individual research papers

Articles	Summary
P1	This paper addresses the need to better understand the intentions of bots on social media and to develop a shared understanding of how 'social' bots differ from other types of bots. A systematic literature review on all relevant papers that examined social media bot accounts lead to the development of a categorising scheme which groups bot accounts by two dimensions – Imitation of human behaviour and Intent.
P2	The authors of the paper argue that it is important to systematically differentiate between interactive systems such as chatbots and virtual assistants (VA) that are used in a private context and those used in an enterprise context as those require different assessments in terminology as well as in research approaches. Through reviewing literature, a working definition for enterprise bots was extracted upon which the authors propose a research design to examine the acceptance of the former in an enterprise context and present preliminary results from a pre-study.
P3	To assess the differences between human users and social bot accounts, the authors in this paper gathered a set of data from Twitter, identified 771 bot accounts and compared their metrics with those of 693 human users. Amongst others, results showed that bot accounts have less followers than human accounts, post less retweets, use less @-characters and use a higher number of links in their tweets compared to human accounts.
P4	As social bots previously were mostly examined in large international political events, this study researched their occurrence in a regional election in Germany. The authors gathered a dataset from Twitter consisting of 182,995 tweets and searched for bot accounts. Only a small number of 0.2% (61) of all accounts in the dataset could be identified as bots. Language barriers, the smaller scope of the election and characteristics of the national party system were discussed as possible reasons.

P5	To depart from the political context as well as the platform Twitter social bots were normally researched in, this paper examined social bots in a commercial context on SoundCloud. Based on a dataset of six months of activity this paper showed that bot accounts do exist and that they write highly repetitive comments and mainly repost tracks uploaded by others and therefore do not contribute to the community.
P6	In this paper the authors examined the potential of social bots to trigger a spiral of silence in turn helping their supported opinion to seem like the majority opinion. A virtual experiment via a simulation based on an agent-based model showed that already a small amount of 2–4% bots on the overall accounts is sufficient to establish the bots' opinion as the majority opinion.
P7	This paper examined the acceptance of chatbots in an enterprise context. To get an understanding of the factors influencing employee's intention to use chatbots at the workplace, an online survey with employees at white-collar jobs was conducted. The most influential factors were intrinsic with the attitude towards using, influenced by the perceived usefulness which again was influenced by trust having the largest effect on the intention. While other factors like subjective norms did have an impact, those were smaller, and the main explicability stemmed from the former factors.
P8	In this paper the authors examined the ability of automated communication to support humans in creative processes. The authors did so by conducting expert interviews with design thinking moderators to get their take on the ability of and requirements for virtual assistants to be assistive in design thinking workshops. Combined with findings from literature, the insights from the interviews concluded in guidelines for designing virtual assistants that can support human participants.
P9	A laboratory experiment was conducted in this paper to assess the ability of virtual assistants to reduce the cognitive load of their users. 91 participants were split into two groups and had to solve a task, one group having access to information via a virtual assistant the other through a regular static FAQ. Results showed that the VA-group got a higher performance score when solving the task while having a lower cognitive load. The authors concluded that VAs are able to reduce cognitive load with beneficial results for their users.

4.1 Preliminary results on social bots and chatbots in an enterprise context

Before thorough research on the topic of automated communication and the scepticism towards it could be conducted, a common understanding and terminology of ACTs manifestations was needed to delimit different phenomena from each other and get a clearer conception of the relevant aspects. This was the goal of papers P1 and P2.

As Paper P1 argues, on the field of social bots, a common understanding of what can be understood by this term is missing as it is a relatively new field of research. In different publications the term is used to describe differently defined phenomena whereas others describe a similar behaviour of bots but use a different terminology. To get a clearer picture of research on social media bot accounts and clear up confusions regarding different types of bots, a systematic literature review was conducted.

This review, conducted in July 2017, brought forth 103 relevant papers which were then analysed by the authors. On a meta level, the identified papers clearly showed a trend of a year-by-year increase in publications on social bots indicating increasing interest and relevance of the topic (Figure 2 – the dip in 2017 is due to the date of the literature search in the middle of this year).

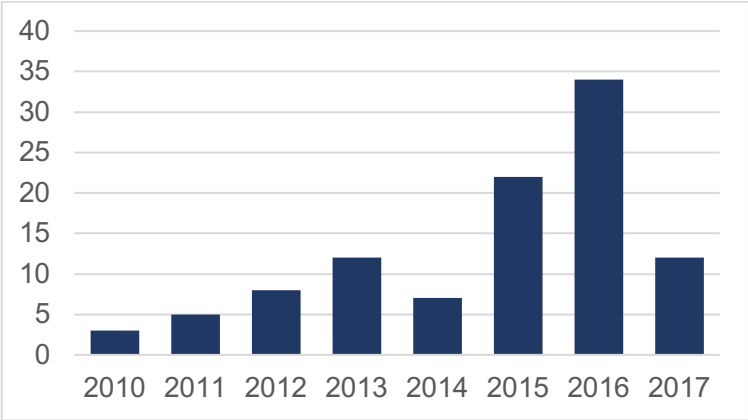

Figure 2. Number of social bots papers published in each year in P1

The identified and analysed paper led to the development of categorisation scheme for social media bot accounts which arranges such accounts regarding their i*ntent* (*malicious*, *neutral* or *benign*) and the degree to which they *imitate human behaviour* (*low to none* or *high*). The categorisation is pictured in Table 4.

Table 4. Categorisation scheme of social media bot accounts (Source: P1)

		Intent (Ferrara et al. 2016)		
		Malicious	Neutral	Benign
Imitation of human behaviour (Boshmaf et al. 2013)	High: *Social* bots	• Astroturfing bot (Ratkiewicz et al. 2011) • Social botnets in political conflicts (Abokhodair et al. 2015) • Infiltration of an organisation (Elyashar et al. 2015) • Influence bots (Subrahmanian et al. 2016) • Sybils (Alarifi et al. 2016; Goga et al. 2015) • Doppelgänger bots (Goga et al. 2015)	• Humoristic bots (Veale et al. 2015)	• Chat bots (Salto Martínez & Jacques García 2012)
	Low to none	• Spam bots (Wang 2010) • Fake accounts used for botnet command & control (Sebastian et al. 2014) • Pay bots (Subrahmanian et al. 2016)	• Nonsense bots (Wilkie et al. 2015)	• News bots (Lokot & Diakopoulos 2016) • Recruitment bots (Flores-Saviaga et al. 2016) • Public Dissemination Account (Yin et al. 2014) • Earthquake warning bots (Haustein et al. 2016) • Editing Bots, Anti-Vandalism Bots on Wikipedia (Tsvetkova et al. 2017)

P1 argues that all papers with a high degree of human imitation can be categorised as social bots, regardless of the intention of the bots. However, most publications using the term social bots are using it to refer to only those accounts with a malicious intent. Accordingly, more papers were found researching bots with a malicious intent than those with neutral or benign intent. An important take-away from P1 is that research mainly is conducted on social bots with a malicious intent which leads to more publications on that note, leaving the term social bots with a negative connotation and shaping the overall impression in an unfavourable way.

While Paper 1 addressed social media bot accounts and analysed as well as organised the terminology and characteristics, Paper 2 argued that the same was necessary for bots in an enterprise context. Furthermore, it argued that research on bots in an enterprise context was insufficient and needed to be expanded to fully understand the potential the application holds. Related to the overall lack of specific research on bots in an enterprise context, an independent definition for this kind of automated communication was also missing. P2 thus argued that not

only the labelling of different phenomena such as virtual agents or chatbots needs to be explicit but that a clear distinction between automated communication used in the private sector and that in an enterprise sector is necessary.

In a first step the authors reviewed relevant literature and found that automation implemented via natural language interfaces and conversational access that are being added to software solutions, and the familiarity with it would become more important in a professional context and people thus need to be accustomed with it. Furthermore, the literature also revealed a lack of research on how bots in the enterprise context are accepted or specifically capable of assisting human users. As the authors also found a focus on the harmful influences of growing automation (which confirms findings from P1), they concluded that a) research on the potential of automated communication in the enterprise context was necessary and b) that this research needed to be based on its own findings from the enterprise context rather than findings from public or private applications (such as social bots or smart home assistants). As a first step the paper proposes a definition for bots in an enterprise context, so called enterprise bots, and, based on this definition, to assess what influences the employee's usage intention of said enterprise bots. The proposed model was later tested in a subsequent paper (P7).

In summary, P1 and P2 found that a clear delimitation between different kind of bots is necessary (as is an individual research approach), provided guidelines of how to specifically categorise and approach social bots as well as bots in an enterprise context and found that automation and specifically automated communication in research is oftentimes depicted in a negative manner.

4.2 Social Bots

This section presents the findings from papers P3-P6 researching how automated communication in the form of social bots is currently applied and how that may lead to technological scepticism.

On the basis of the findings from P1, first, to get a broad overview on the application of social bots, the respective papers examined them in diverse scenarios: Social bots on Twitter and in the political context of an international election (P3), social bots on Twitter in the political context of a regional election (P4) and social bots on SoundCloud in a commercial context (P5).

Paper P3 looked at differences in the metrics between human users and bots accounts. While social bots do per definition act in a human-like manner, previous research indicated that some differences between them and those of human users persisted. Furthermore, knowledge on the metrics may allow to better estimate the strategy pursued by bots, adding to the knowledge of how automated communication is currently applied as well as the impact of bots, thus giving insights into their success in disseminating information and reaching human users.

The authors therefore gathered a dataset of 6.5 million tweets from the last week before the 2016 US presidential election. It contained tweets that either included the term "Hillary Clinton" or "Donald Trump", the two presidential candidates at that time. For the analysis, the authors drew a sample of 771 social bot accounts and 693 human accounts and compared them based on the hypothesis in Table 5.

Table 5. Hypotheses tested in Paper 3

Hypothesis	Status
H1a: Bot accounts have a lower number of followers than human accounts.	Supported
H1b: Bot accounts have a higher number of retweets than human accounts.	Declined
H1c: Bot accounts have a lower number of @ in Tweets than human accounts.	Supported
H1d: Bot accounts have a higher number of links than human accounts.	Supported
H2a: The number of followers of a bot account is positively related to the number of retweets of this account.	Supported
H2b: The number of followers of a human account is positively related to the number of retweets of this account.	Declined
H3: There is a positive relationship between tweeted links per day and retweets per day which differs between bot and human accounts.	Supported
H4: The more tweets are created per week, the more followers are generated.	Partially Supported

As Table 5 shows, 6 of the 8 hypotheses were supported (or at least partly supported in case of H4). The analysis showed that social bots had less followers and a lower number of @-mentions

in their tweets but used more links in their tweets compared to human users. Also, the more followers a social bot account had, the more retweets it got and the wider its tweets got disseminated. Testing of H4 also showed that bots generate one additional follower with almost every second tweet which was not true for human users. Paper P3 thus showed how bots act and that they differ from human users. These findings may on the hand help to distinguish them from each other, on the other it helps to get a clearer picture of how automated communication in the form of social bots is currently applied. The results of P3 were based on a sample from one of the largest political events in an English-speaking country with the outcomes having implications for almost the entire world. This was true for most of social bot research at the time, which is why, to get a more complete impression of social bot usage, also events with a different scope needed to be examined. Consequently, Paper P4 looked at social bot usage during a German regional election.

For Paper P4 the authors examined social bots during the election of the regional parliament for North Rhine-Westphalia, Germanys most populous federal state with approximately 18 Mio. citizens. Like in P3, the domain was again politics and the platform Twitter but expanding on that scenario, the regionality of the event observed introduced a new aspect to social bots research. Previously, most findings on social bots originated from observations of political events that attracted great interest internationally due to the possible scale of its outcomes – such as the conflict between Ukraine and Russia in 2015, the 2016 US election or the 2016 UK vote to leave the European Union. The paper thus argues that, to get a complete picture on social bot usage, also smaller events needed to be examined.

For this purpose, a Twitter dataset of one month of communication regarding the state election (in German "Landtagswahl") was collected with tweets either having to mention the term *landtagswahl*, *landtagswahlNRW* or one of the parties' Twitter accounts to be included in the dataset. In total, 182,995 tweets by 33,481 individual accounts were collected. Afterwards a multi-step approach was chosen to identify social bot accounts within the dataset, including rating by the website botometer (formerly BotOrNot) that draws on several criteria to rate the probability of an account being a bot. This approach was complemented by manual checks of the identified accounts. This procedure led to the identification of 61 social bot accounts which made up around 0.2% of all 33,481 accounts in the dataset which tweeted a total of 336 times (0.2% of all tweets, matching the share of bots on all accounts). Due to the low number of social

bots, the only identifiable strategies were traditional spamming without a political concept and the sharing of diverse news articles concerning the state election.

The paper offers two interpretations of the findings. First, the language could make a difference as most of the other results that reported larger amounts of social bot activity was conducted on events related to an English-speaking audience (as in P3 and P5). Second, the regional nature of the election may have also been of importance as former research that found larger, more significant amounts of social bots mostly examined larger events. While literature in part reported examples for application in smaller-scale political events (Mustafaraj and Metaxas 2010) or in non-English speaking contexts (Abokhodair et al. 2015), the combination of both factors may be enough to diminish social bot activity. However, as most research on social bots regarded them in the context of politics as well as on Twitter, to get more comprehensive insights, Paper P5 examined social bots in a commercial context on the Platform SoundCloud.

As already pointed out in P1 social bots needed to be examined not only in the political realm but also in other domains such as commercial contexts. While the concentration on politics is comprehensible as it is a sensitive topic, it is also conceivable that those bots are used either to defame or promote products. Accordingly, Paper P5 examined social bots on the platform SoundCloud, a social media platform for uploading, commenting on and sharing music tracks used by creators to disseminate their work. As previous research on social bots almost exclusively researched them on Twitter where other metrics are available compared to SoundCloud (cf. P3 and P4), P5 introduced the metric of comment uniqueness (CU) to identify possible social bot accounts. This measure puts the unique comments an account has written into proportion to all comments it wrote. A high score indicates that most comments are unique whereas accounts with a low score post repetitive comments with little to no variation.

The metric of CU was used on a SoundCloud dataset of six months of activity, comprising of nearly 16 m. tracks and more than 12 m. comments written by 2 m. users. Using the CU on users with at least 100 comments, a clear distinction could be made between those users with a high score (i.e. unique comments) and those with a low score, most likely being social bots. Looking at the behaviour of the latter accounts, they had less followers, posted less tracks and mainly reposted tracks uploaded by others. To get more insights into the behaviour of the two CU-groups, the authors of P5 also included a network visualisation of highly active users and their comments. As Figure 3 shows, a network of bots (red nodes) exists which frequently

(re)posts comments to the same tracks while human users (green nodes) with a low CU score are more isolated and comment on more isolated, individual tracks.

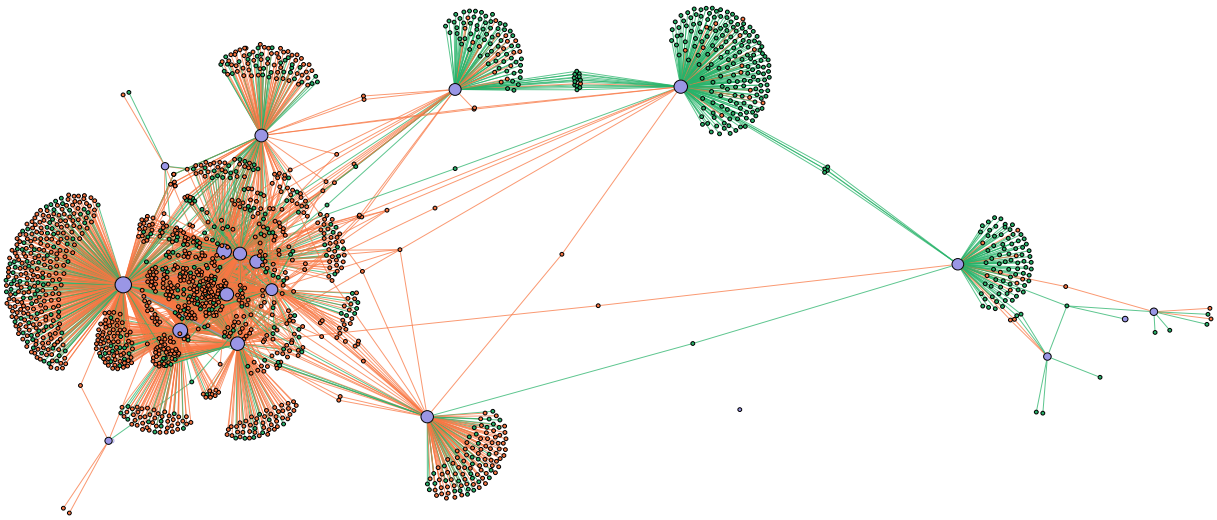

Figure 3. Network visualisation of the 25 most commented tracks on SoundCloud. The 19 purple nodes are the track authors. Users who post highly repetitive comments are coloured red, users who post diverse comments green. Edges represent comments (Source: P5).

The authors conclude that the accounts with a low CU score try to artificially boost the popularity and visibility of tracks while they don't contribute to the overall community nearly as much as real users do. This indicates that the former are mainly trying to boost visibility of certain content and in turn, sales, a behaviour normally also attributed to social bots in the political realm (except for cash sales the accounts there aim at "selling" a political view).

As in the political context, P5 demonstrated that social bots are also active in a commercial context and apart from Twitter. Besides the topical findings, while P3 and P4 each used a different approach (in part uniquely applicable to Twitter) to identify social bot accounts, P5 also demonstrated the suitability of the introduced CU score for this task. Together, P3-P5 gave a broad overview over the activity of social bots in a range of contexts and could show that a) they act different than human users in several ways, b) they don't seem to be as active in smaller, non-English speaking events and c) they are applied outside of Twitter and in a commercial context. In all cases, the bot accounts seem to pursue the goal of boosting certain contents, presumably, to increase its visibility to the human users of the networks. Comprehensibly, this behaviour may lead to a feeling of uncertainty and unease and in turn scepticism with human users towards the technology, as it is hard to spot, and its consequences are unclear. Still, as

these consequences are unclear, little is known about the actual effectivity of social bot behaviour. While most of the research assumes that social bots influence human users, e.g. by manipulating their opinion, proof of this assumption is missing. Accordingly, P5 also pointed out that "*it is not clear whether these suspected bots are successful at convincing human accounts*" (P5, p. 8) and that future research should address this research gap. This is especially important, as the assumption that bots try to manipulate opinions is one of the main concerns that fosters scepticism towards automated communication and thus needs clarification. To get a better understanding of how automated communication in the form of social bots may affect human users, Paper P6 follows the call of P5 and examines the impact of social bots on human users.

While P1, P3 and P5 demonstrated findings on the occurrence of social bots, when assessing the influence of social bots, it is unclear if and to what extend the actual opinion of human users is influenced. Furthermore, one of the main obstacles is the observability of the effect they can have on human users. While their behaviour can be observed and their strategy can be inferred from these observations, the actual influence can not. Paper P6 addresses these issues by a) using the theory of the spiral of silence to describe a possible mechanism by which social bots can influence human users and b) conducting a virtual experiment in the form of a simulation to assess this mechanism.

While many of the papers identified in P1 assume social bots effectiveness through influence on a user's opinion, in the spiral of silence not the actual opinion of a person is affected but rather the tendency to speak out if confronted with an opinion, contrasting one's own and held by an apparent majority. Paper P6 simulates this mechanism via an agent-based model. Within the simulation, the agents depict either human users of a network or social bots. These accounts differ regarding their likelihood to voice their opinion. One of two opinions is held by half of the human actors in the network meaning that opinions are equally distributed throughout the human agents in the network. Agents representing human users have a confidence-threshold which, if surpassed by too many surrounding actors expressing the opposite opinion, leads them to no longer express their opinion (but not changing it). In contrast agents representing social bots always express their opinion – regardless of their surroundings. The model tested under what circumstances the opinion held by the social bot accounts in the end was the opinion voiced by the majority of the network even though the actual opinion was distributed evenly between agents depicting humans.

The results of Paper P6 revealed that the opinion voiced by the bot accounts in the end always emerged as the majority opinion – even though the opinion between the human agents was evenly distributed. While this relation was moderated by a) the number of bots, b) the density of the network, c) where in the networks the bots attached themselves to and d) the degree to which the bots can influence human agents, the bot opinion always was the opinion voiced by the majority in the end. Transferred to the findings from P5 (pictured in Figure 3) this might be one explanation for the isolated clusters of bot accounts with hardly any human activity.

4.3 Virtual Assistants and Enterprise Bots

The following section addresses the potential of automated communication to support human users and thus to counter scepticism towards the technology. It summarises the findings from paper P7-P9.

To counter scepticism towards a new technology, it is important to not only understand the factors fostering scepticism and thus what keeps people from using it, but also what favours the usage. As P2 pointed out, findings of bots used specifically in the enterprise or working context was sparse and warranted an individual research apart from that on bots in public or private usage. Accordingly, P7-P9 took up this argumentation and examined the technology specifically in the context of task-solving and work. First, Paper P7 examined the factors influencing the intention to use bots in an enterprise context. While the workplace shifts to a digitised environment, automated communication bears great potential for increased efficiency. Still, there is little knowledge of the factors influencing the employee's intention to use automated communication in the form of bots in an enterprise context. The paper argues that enterprises should know in advance of investing into enterprise bots (EBs), whether those get accepted freely in order to not impede IT- and overall job satisfaction and what factors the intention to use them depends on to better address the needs. EBs here are defined as "*an automated user service that provides casual and conversational interactions with complex enterprise systems and processes. Enterprise Bots can e.g. answer questions or perform smaller tasks. A user can interact with it by just typing or speaking a request in natural language. It can only act in a passive way, meaning it must actively be triggered by the user.*" (P2, p. 4)
Based on previous technology acceptance research in the form of the decomposed theory of planned behaviour, Paper P7 uses an online survey to assess the factors and the overall intention to use EBs. Figure 4 depicts the results from a structural equation model.

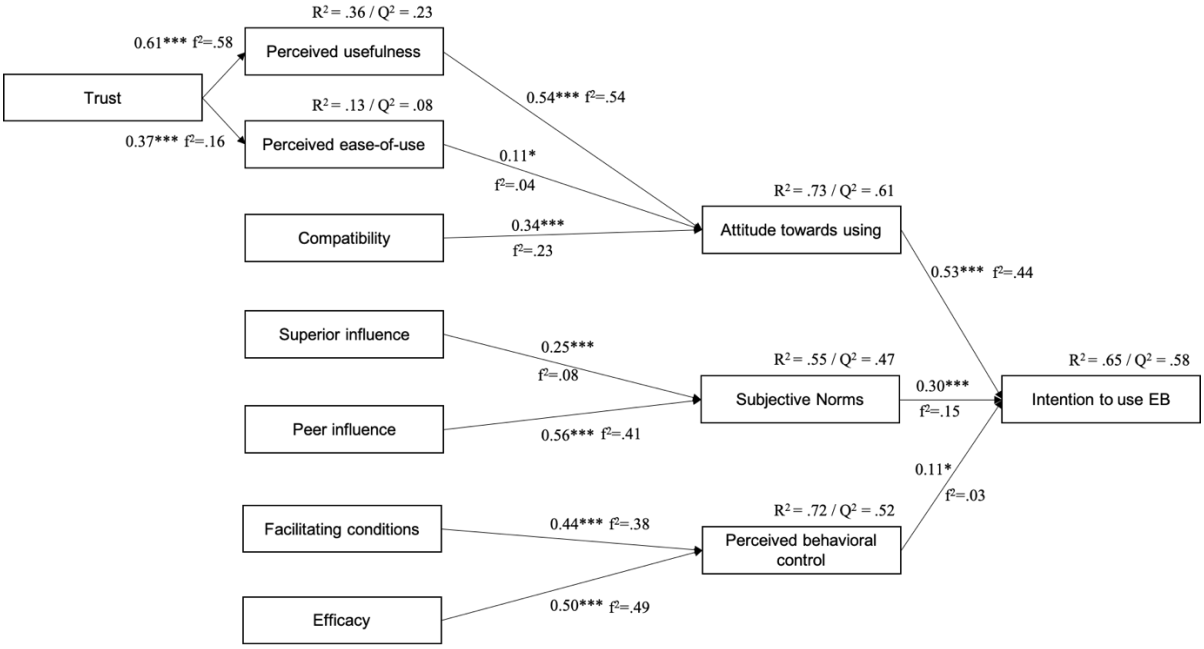

Figure 4. Results of the research model on factors influencing the intention to use enterprise bots (*p ≤ 0.05; **p≤ 0.01; ***p ≤ 0.001; Source: P7)

The main findings were that intrinsic factors had the strongest positive influence on the intention to use Enterprise Bots. External factors were also positively related to the intention but to a comparably much smaller extend. Especially the perceived usefulness of an enterprise bot had a positive influence on the attitude and in turn the intention to use them. However, while superior and especially peer influence as well as facilitating conditions and the self-efficacy affected the subjective norms and the perceived behavioural control respectively, the latter constructs small effects on the overall intention indicated the importance for employees to be convinced of the usefulness of an ACT tool for themselves in order to use it. These findings also corroborate the overall assumption, that, to decrease technological scepticism, it is important to point out the benefits of using automated communication. Picking up on this aspect, P8 and P9 research concrete scenarios in which the benefits of automated communication are applicable.

The authors of P8 examined if and how automated communication in the form of virtual assistants can be used to assist during design thinking (DT) workshops. They argue, that, to stay competitive, enterprises depend on innovative products, processes or services, one way to develop those being DT workshops. As these require on the processing of large amount of information and knowledge, virtual assistants may help to cope with this amount, in turn

reducing the workload of human moderators and leading to better results of which the company may also profit. As the goal of this paper was to assess the suitability of virtual assistants for this task, the authors decided to get the opinions of experts familiar with the DT field on to what degree VAs might be applicable in this scenario. They thus conducted 10 semi-structured expert interviews with DT experts and collected the findings into guidelines for the design of VAs supporting creative workshops. The paper presented guidelines on how to design such a system divided into the three categories 1) general conditions, 2) characteristics of a VA and 3) tasks of the VAs as moderation support. The authors concluded that based on the opinion of the experts, VAs do offer great potential to be used in the scenario of creative workshops.

While the findings from P7 and P8 were based on insights a) from *potential* users and b) from experts' opinions, P9 aimed at gathering real usage data to objectively measure and demonstrate the actual benefits the usage of VAs can have on the performance of its users. This also answers the call of P7s indication that pointing out the benefits of automated communications usage to a user will increase the intention to use these systems. P9 argues that, while the information input for knowledge-workers increases, their cognitive capacity does not. As cognitive resources used to capacity lead to a decrease in performance and an increase of mistakes by the workers, it is desirable to assist them with demanding tasks. Such assistance may come in the form of VAs which may help to reduce the cognitive load of its users.

To test this assumption, the authors conduct a laboratory experiment with participants allocated to one of two groups. Both groups were presented with the same task of solving the chronological order in which certain steps to conduct a marketing study had to be taken within a time limit of 10 minutes. While the experimental group had access to a virtual assistant that could answer questions, the control group had access to the same information, however not in an interactive form but as a traditional FAQ. The cognitive load of both groups was assessed via the NASA-TLX.

The main findings were that participants in the group with access to the VA performed significantly better (Table 6), needed a minute less to complete the task and had a lower cognitive load than those without the VA (Table 7). Furthermore, almost all (95%) of the users in the experimental group found the virtual assistant helpful. The findings indicate that VAs are a feasible tool to reduce the cognitive load of its users and thus are a promising tool to be used by employees for their own good. Combined with the findings from P8, P9 thus demonstrates

the potential automated communication holds to support human users which, as P7 found, was among the most important factors of a positive attitude towards using the technology.

Table 6. Comparison of Task-score performance (Source: P9)

	Group	N	Mean	Median	SD	SE
Task performance score	without assistant	46	13.7	14.0	7.36	1.09
	with assistant	45	19.8	20.0	7.09	1.06

Max. Score: 28 (higher is better); $t(89) = -4.01$, $p < .001$

Table 7. Comparison of perceived workload (Source: P9)

	Group	N	Mean	Median	SD	SE
NASA-TLX score	without assistant	46	10.3	11.3	4.52	0.666
	with assistant	45	7.17	6.83	3.79	0.565

Max. Score: 20 (lower is better); $t(89) = 3.55$, $p < .001$

5 Discussion and Implications

While the individual papers used in this thesis all address research questions on their own, taken together, they help to address the two principal research questions regarding the current application of automated communication that might foster scepticism in the form of social bots (RQ1) as well as the application of automated communication in the pursuit to support human users (RQ2). These two aspects are examined in the next two sections before implications for research and practice are drawn and limitations as well as possible future research are discussed.

5.1 Technological scepticism through social bots

To better understand current applications of automated communication and how these might foster scepticism towards this technology, research question 1 analysed social bots as a manifestation of automated communication. This was done based on findings from P1 which first analysed and ordered current research on the topic of social bots. One of the findings here was that the term social bots should be applied to all bot accounts on social media sites with a high degree of human imitation, regardless of their intent. However, the majority of studies using the term, and the majority of studies on social bots researched the occurrences with malicious intent. While positive applications of social bots do exist, they are not nearly as dominant in the reporting as negative scenarios which makes it more likely that the latter findings get wider dissemination and influence the overall picture of the technology. Also, possibly due to a higher news value, those are the findings that then get picked up outside of academia and thus shape the picture of the technology in public (Orben 2020). Using the insights and categorisation from P1, P3-P5 looked at the actual implementation of social bots and thus automated communication in practice. Taken together, the papers demonstrate several main aspects.

First, the papers show that automated communication in the form of social bots is currently applied in different contexts but with a common goal. While P1 found that social bots are mainly regarded in the political domain (as is the negative reporting on it), it was feasible to assume that mechanisms applied by bots in this context also have potential to be used in commercial scenarios. As especially P1 and P3 showed that social bots were used in the political domain, P5 is among the first papers that demonstrated their use in a commercial context. This is important for several reasons. For one thing, it shows the versatility of the technology. Automating communication is not only useful in a limited scenario but can be expanded to

diverse settings using similar technology. Combined with findings from P3 this also helps to abstract an overarching goal of social bots use and in turn demonstrates a way, automated communication is currently applied: To sell – be it political opinions or actual commercial goods. Regardless of where they are used, they spread a view so that people take it up. In the political context observed in P3 they offer an opinion on politics and try to get it out to the users whereas in P5 they offer songs whose popularity they try to boost. It is conceivable to assume that this technique might be also used in scenarios that have not been examined yet, e.g. to boost sales of physical consumer goods or the popularity of parties in contests where public opinion matters. An example for the latter scenario may be the Eurovision Song Contest where research showed that the outcome could in part be derived from social media data (Stieglitz et al. 2020). The latter case also demonstrates a scenario where social bots may not only disrupt opinion climate but also skew the validity of scientific findings that are based on social media data (e.g. the field of Social Media Analytics). Conclusions on human behaviour or views drawn from examining social media data might be less valuable if they originate from artificial actors instead of human users and this goes unnoticed. These possible examples demonstrate a) plausible cause for scepticism as realities might get disported by such social bot accounts where human users are not aware of or expecting artificial actors and b) the need to expand the view on social bots for further purposes and to different social media platforms apart from Twitter as these will help to broaden the general view on social bot usage and illuminate further application scenarios for automated communication. Differences between social media platforms should not be underestimated. Currently, most insights on social bots stem from research on Twitter (cf. P1) and as such the reporting on social bots is also concentrated on Twitter; however the significance of findings from Twitter differs within each country and can not be generalised. Also, the findings demonstrate that other social media platforms should not be neglected as is currently the case. One reason for this being the comparably good accessibility of the Twitter API whereas other networks are harder or not at all accessible - which is a general problem of research on social media platforms (Bruns 2019).

Due to the lack of findings from research on commercial social bot use (and in turn also the reporting on it in news media), these findings hold the potential to further incriminate the technology. Human users might not be keen to find that the visibility or popularity of a good (in P5s case music tracks) is artificially bolstered. These findings might also be applied to automated communication outside of social media sites (thus apart from social bots) and, for example to online shops like amazon. Here artificial product reviews emerge as a problem –

not only for the customers but also for the site themselves as distrust diminishes the usefulness of this major feature.

Combining the findings from P3 and P5 gives insights into another important aspect on the current application of automated communication in the form of social bots: Their overall behaviour regarding one of the core aspects of social media platforms - the idea of community. Social bots do not participate in the community of a social network as human users do. P3 and P5 show that, while social bots may be hard to distinguish from human users at first sight and from a micro perspective (i.e. on an individual level), when inspecting culminated data, differences become apparent. One of the main findings here is that social bots are not seeking to contribute to the community but seem rather bluntly determined to express their predefined messages. This makes sense considering their apparent overarching goal is to sell and not to make friends so to speak. However, one of the core benefits from using social media platforms is the community. Human users who seek out those platforms for the community-related features may get driven away in the long run should the somewhat asocial behaviour of social bot accounts prevail and thus diminish the utility of the platforms – in turn also fuelling technological scepticism. On Twitter the neglect of community by social bots becomes apparent as they mention other users more seldom, have less followers and use more links in their tweets, thus steering users away from the platform itself, on to a space the social bot accounts (respectively their operators) have more control over. For social bots on SoundCloud the same is true for the lower number of followers while the disinterest in the overall community becomes also apparent by the way smaller number of uploaded tracks and in turn the way larger number of reposted tracks. For the latter example, the network visualisation in P5 corroborates this assumption as it shows that the social bot accounts mainly stay within their own bubble and repost the same tracks while human users act much more naturally and diverse (Figure 3).

Taken together the findings that automated communication in the form of social bots aim at selling but act rather blunt in pursuing this goal and that they are not actively pursuing the community aspects of social network sites, the overall behaviour is reminiscent of brute force approaches where a goal is pursuit by sheer mass of attempts instead of delicate methods. This demonstrates a core aspect which underlies current bot application: While social bots seem to be somewhat more sophisticated in that they dispose human-like messages, their application does not seem to be worthwhile for the operators when too much effort is put into a nuanced behaviour.

This insight is supported and further refined by the findings of P4 which demonstrated that social bots are not arbitrarily applied to each and every event or case that may seem fit. While it is feasible to assume that they are a very efficient tool to disseminate messages, their usage, at least in political domains, may be limited by a combination of language and event-size criteria. Possibly, a threshold in significance of the event they are applied to needs to be met for their application to be worthwhile.

P4s findings also indicate that social bots are not sophisticated enough for application in more nuanced scenarios. Findings from former research on social bot usage in a political context (e.g. the US presidential election, the Syrian civil war, the Brexit vote or the conflict between Russia and Ukraine) mostly analysed their application in scenarios where two main parties where involved and thus tow possible sites to take. The German system examined in P4 is more diverse as several parties with different views compete. This also supports the findings from P3 and P5 that social bots operate rather blunt and need a simple goal in order to function and for their application to be worthwhile – an aspect picked up by the simulation carried out in P6.

P6 demonstrates a mechanism by which bots are able to have an effect on the opinion climate. In contrast to the notion, that the actual opinion is manipulated, the paper argues in line with the theory of the spiral of silence, that rather the tendency to express the opinion is influenced. The findings demonstrate that this is an expedient procedure which would help the accounts communicating automated to increase the visibility of their cause. P1, P3, P4 and P5 showed how automated communication is currently applied and also gave indications on how this may lead to scepticism (e.g. through a heightened reporting on malicious social bots or the opacity to the human users). Still, one of the most pressing aspects in research on social bots is that of their actual impact. The assumption that they are effective in pursuing their goals and that those automated actors influence human users in their opinion and behaviour shapes their perception in the public. This fear of being unknowingly manipulated is at the centre of scepticism towards the technology but its legitimacy had not been shown before. Paper 6 fills this gap and answers the call for the demonstration of how social bots might be able to take influence (e.g. raised in P5). While the findings are based on a simulation and thus not observed on a real-world scenario, they present a conceivable mechanism by which bots may work. Anecdotal evidence sees people abandoning e.g. Twitter as its debating culture becomes rawer and more unpleasant which may in part be fostered by bots that are not affected by a growing frustration. People

however might be bombarded with messages until they leave the platform altogether, leaving the opposite opinion the field. Against this background a growing scepticism is comprehensible.

The aspect that the findings are based on a simulation, which can always only be a simplified representation of reality, might in turn be helpful in this case as it shows that even with an unnuanced behaviour, the automated agents in the simulation always had an influence on the final opinion climate. The bluntness of social bots demonstrated in P3 and P5 might thus be sufficient and their real-world behaviour does not need to be sophisticated to be effective.

5.2 Automated communication to decrease technological scepticism

While RQ1 concerned the current application of automated communication in the form of social bots and how it can foster scepticism towards the technology, RQ2 dealt with potential application scenarios in which automated communication supports human users (i.e. in the corporate context) and how those in turn may be beneficial in decreasing a scepticism towards the technology. These aspects are addressed in the following.

Just with P1 for RQ1, P2 lay the foundation for further examination of RQ2. Through reviewing literature, it came to the conclusion, that in research a differentiation between bots in a private and an enterprise context was missing and that research on the latter was sparse. While digitalisation of companies advances and its possible benefits e.g. in efficiency gains are examined, the actual application scenarios in the form of bots are not. This finding was supported by the subsequent P7. This lack of research may be a driver for possible scepticism – the actual benefits of using automated communication is insufficiently researched and thus not apparent to the public perception – which is the opposite of what P1 found for the reporting on malicious social bots. P2 thus pointed out the need for a specific string of research on automated communication dedicated to its use in the corporate, or enterprise context to better understand its benefits. Further it demonstrated that little is known about the willingness of employees to use automated communication in an enterprise context or factors influencing this intention. Nevertheless, these aspects are important to not (further) alienate employees and to assess the general willingness of those to work assisted by automation. As the goal is to reduce reservations about the technology, findings on this area can help to understand what shapes the decision for or against using a system. P7 picked up this call and answered it by examining the factors influencing the intention to use automated communication in an enterprise context.

As P7 could show, especially factors related to notion of their own benefit from technology usage are important aspects that let employees engage with automated communication, here represented through enterprise bots. Particularly the perceived usefulness had a positive impact on the attitude towards using bots in an enterprise context which demonstrates the need to point out what benefits the use of a system has to the individual. In return this shows that a way to decrease scepticism is convincing employees of the benefits automated communication holds for them instead of simply forcing them to cope with systems. Influence of others in turn does only play a minor role in this context. Individuals thus should ideally be free to decide whether they want to use a system and should be aware of the benefits. Trust also proved to be a positive driver of the overall intention to use automated communication (cf. Yan et al. 2013). While not directly comparable, this is similar to findings for social bots that opaqueness of bot application in turn may foster scepticism. While it seems that the opposite of what P1, P3 and P5 found to foster scepticism (albeit in the public context) in turn may help to increase the likelihood to engage with bots in an enterprise setting, it is important to differentiate the two contexts and mind the different scenarios in which humans encounter the technology respectively. Still, in general users must not be deceived and need to decide for themselves if they use automated communication – which becomes more likely when they see the benefits of using it. Fortunately, as P8 and especially P9 show, these systems do have the ability to support human users.

In accordance with P7, P8 and P9 presented scenarios which demonstrate in what ways automated communication can be actively applied in support of human users. While P8 offered expert assessments and guidelines of how automated communication needs to be designed in order to support humans, P9 showed what an actual usage scenario may look like and provided measurable results that bots do indeed have a positive impact on the performance of their users. While former papers looked at individual aspects of bots (e.g. typing speed or human-likeness) and how those may affect the evaluation of the bots (Ciechanowski et al. 2019; Gnewuch et al. 2018), P8 and especially P9 give more practical insights on the usefulness of automated communication when applied to concrete scenarios. The expert's opinions from P8 on how bots might help to take over subtasks to ease their job were confirmed in P9 which demonstrated a mechanism by which automated communication can have a positive impact on human users.

The findings may help to reduce scepticism towards to automated communication as they demonstrate how automated systems may work in a cooperative mode with humans.

Automation does not need to take over and thus away jobs from humans to increase productivity of a company. Instead a cooperative mode is possible where human jobs are supported and augmented by automated communication technologies so that they can work more efficiently and possibly also make less mistakes - a concept akin to that labelled hybrid intelligence (Dellermann et al. 2019).

As especially the loss of jobs is a driver for scepticism against automation as a whole (Frey and Osborne 2017) the results from P8 and P9 point out that not only does automated communication not need to erase jobs but instead is able to support employees with taking over sub-tasks and thus easing their overall mental load when completing tasks. Together, P2, P7-P9 thus demonstrate that for users to engage with automated communication, it is important for them to see the benefits of using it, which lie in a cooperative mode where automated communication takes over sub-tasks and thus reduces the mental demands for its users to enhance their performance.

5.3 Implications for research

The thesis aimed to assess the current state of ACTs in the form of two cases – ACTs in the form of social bots and in the form of chatbots in the enterprise context. A contributes to research is the finding that ACTs don't need to be sophisticated to have an effect. Be it dissemination of an opinion by social bots or the support of knowledge workers by chatbots, there is no need for an elaborate human-like AI to achieve these tasks. Building on these findings, research on how ACTs may make a difference does not need to tend to utopic scenarios where human-like AI is shaping the communication with chatbots but will find answers in the resources common today. Considering ACTs in the form of social bots, they may indeed act rather blunt which nevertheless may even be of help in succeeding to sell their messages. This opens up a new focus in research on the phenomenon that stronger takes into account the bluntness with which goals may be pursuit. Regarding the application of ACTs in the corporate context, this means that research should thoroughly assess those systems available today, regardless of possible shortcomings and work around limitations by identifying and testing viable application scenarios. Being open to new developments in technology and having the skills to understand and use it is vital for the advancement and prospering of a society. This willingness to engage with this new technology is important to understand its benefits as well as its downsides as it helps to determine how it can be actively used for the good of their users but also in what cases it must be handled with care as it might pose a threat. This thesis

contributes to the necessary understanding regarding the technology of automated communication, what it can do for society and in what cases it may evoke unease. It demonstrates that social bots, as a current occurrence of automated communication, act in different scenarios and that a broader view on them is needed to fully grasp the magnitude of their distribution. The thesis further shows how different aspects of the social bots themselves but also of the reporting on them may evoke unease in human users and presents a mechanism by which automated communication can have an influence on the opinion climate of a discussion on a topic. This latter aspect is especially important for research on social bots as it introduces a method to theoretically estimate the potential of social bots to have an impact on users – an issue that former research only assumed but failed to address sufficiently.

Besides demonstrating how automated communication may evoke unease, the thesis also points out the necessary conviction in human users to be willing to actively engage with and apply automated communication in order to use it for their benefit which in turn decreases scepticism. It showed that the conviction of the usefulness is among the most important aspects for users to see the meaning of using bots and further presented examples of how ACTs can contribute to this conviction as they can be used to be of benefit to its human users e.g. by assisting them and thus decreasing their mental load. To better emphasize this ability, more research on this field is needed that explicitly addresses scenarios and further ways in which the application of automated communication is of benefits to human users. For this latter aspect, methodological approaches used in this thesis to assess hard to observe aspects in research on automated communication (e.g. using virtual experiments via a simulation, laboratory experiments or social media analytics) could be applied and further refined.

5.4 Implications for practice

On a practical side the findings demonstrate that automated communication can be a useful tool for companies to support employees where possible. It describes scenarios of how bots might be used to reduce mistakes and increase productivity and further gives insights that help to improve the acceptance when introducing bots (cf. P2, P7, P8 & P9). The latter aspect is e.g. critical to not waste resources on systems that do not fit employee's needs. Those and their needs should be involved in the overall development and implementation process to introduce systems that are willingly adopted by employees (cf. P7). These findings might also extend beyond the corporate sector as they demonstrate the ability of automated communication to take over tasks of their users which generally can be of use in different scenarios (P9). Wherever

cognitive resources are required to fulfil as task in the digital realm, automated communication may help to do so with positive outcomes. This in turn can shape the familiarity with and opinion of these systems in a favourable way – which again would also be beneficial to employees and the overall society.

The thesis also demonstrates the need to be aware that bot accounts do exist and act on different social media sites (cf. P1, P3 & P5). Media literacy is required to correctly estimate the source and validity of statements and also metrics such as likes or share counts. Especially the metrics can be inflated by automated accounts and might thus not mirror actual opinion climate. In in conjunction with this, the thesis demonstrated how bot accounts may have an impact on this climate (see P6). This needs to be considered, e.g. by news media in their reporting but also by researchers when conclusions are drawn based on social media data. However, the sovereignty of users in this regard should not be underestimated. They do not need to feel helpless. Knowing how bots could work also helps to better understand how their effect may be countered – e.g. by voicing one's opinion regardless of an apparent opposite majority opinion or by stating support to other users of the same opinion.

5.5 Limitations and future research

This thesis gives a broad overview on how automated communication is currently applied and was created with due diligence. However, not even in the frame of a dissertation each and every conceivable scenario can be examined. Social bots in politics as well as a commercial scenario were regarded, however, further settings are imaginable. Social bots could for example also be applied in a commercial context on Twitter and not only to promote a certain good, service or company but also to defame others. Here, an ethical evaluation of social bot usage could also be of interest. Questions may be whether it is condemnable to artificially enhance the dissemination of messages and opinions through bot usage and if this evaluation would be different for covert and open application of social bots. Furthermore, the application on platforms like Facebook, Instagram or private messaging services such as WhatsApp is possible and needs further research. Several social media platforms popular with a younger audience such as Instagram or TikTok also have richer content types like images or videos at their core. In the light of growing sophistication and availability of software to produce DeepFakes (images or videos manipulated with the help of AI by e.g. swapping faces of protagonists), social bots might also evolve and occur in these forms - which would raise additional ethical

issues that warrant examination on their own. Future research needs to take this into account when researching the phenomenon.

Regarding usage scenarios, more research is needed on the beneficial applicability of automated communication in social media. As reported in this thesis, social bots with benign intent exist but in contrast to those with malicious intent are hardly examined. Here, first research emerges that explores the feasibility of using social bots in crisis situations (e.g. natural disasters) to support emergency agencies in getting insights on the situation and disseminate their messages to help the public (Hofeditz et al. 2019). Insights from research on malicious social bots (e.g. how they disseminate their messages and may have an impact) can in turn be used to expand the reach of those bots used with benign goals. Expanding this branch of research may also be helpful to better point out the benefits of actively using automated communication – also in the form of social bots and thus shaping the general picture to be more comprehensive.

Furthermore, the results on the current application of automated communication mostly rely on quantitative data and not qualitative findings. While the former approach enabled a broader picture on current usage, the latter could dive into the actual content of the messages and also help to better understand how human users perceive a) individual social bot activity and b) the debate on bots in general. These findings may also help to better understand and address concerns of users. While the demonstrated mechanic of how bots can impact the opinion climate of a debate is conceivable, it relies on findings from a simulation which must always be a simplification of the real world. It needs thus to be demonstrated if the mechanism, plausible as it may be from the simulation, can be carried over to actual usage scenarios.

Regarding findings on positive impacts of bots, those stem from opinions of experts (P8) and a laboratory experimental setting (P9). Accordingly, they are not based on data from practice. Here it is important to transfer these findings to real world scenarios in order to assess and strengthen their external validity. Future research may observe the actual implementation of the ACTs in a company and then measure their outcome or test the implementation guidelines and refine them based on these observations.

An aspect that has not been regarded in this thesis is the factor of age. As the overall topic deals with the attitude towards new technology, older generations may have more reservations against using them than younger ones. One consequence may be that different approaches are required

to deal with scepticism in different age groups. People in retirement will have different scenarios where they interact with automated communication and also seek out different benefits from their usage. Future research thus needs to specifically take this aspect into account and additionally regard media literacy in the light of special requirements of different age groups. Especially as automated communication may be a feasible tool to stay in touch with relatives for an older generation, it is simultaneously important to understand how knowledge on the technology's opportunities but also mischief it might bring can be conveyed to this age group.

Findings on how automated communication can have beneficial consequences for human users and how it can support them, as valuable as they are, also need to reach the users in order to diminish scepticism and shape the view and opinion on the technology. While for malicious cases this deed is in part incurred by the media, the same is not true for beneficial appliances. Here, strategies on how to better communicate the advantages of automated communication to the users and which aspects to stress in this regard should be addressed. This thesis presented first findings on this latter aspect which need to be picked up and further elaborated – also against the background of possible age-related differences mentioned above.

References

Abokhodair, N., Yoo, D., and McDonald, D. W. 2015. "Dissecting a Social Botnet: Growth, Content and Influence in Twitter," in *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing*, , February, pp. 839–851. (https://doi.org/10.1145/2675133.2675208).

Alarifi, A., Alsaleh, M., and Al-Salman, A. 2016. "Twitter Turing Test: Identifying Social Machines," *Information Sciences* (372), pp. 332–346. (https://doi.org/10.1016/j.ins.2016.08.036).

Belk, R. W. 1988. "Possessions and the Extended Self," *Journal of Consumer Research* (15:2), pp. 139–168. (https://doi.org/10.1086/209154).

Ben Mimoun, M. S., Poncin, I., and Garnier, M. 2012. "Case Study—Embodied Virtual Agents: An Analysis on Reasons for Failure," *Journal of Retailing and Consumer Services* (19:6), pp. 605–612. (https://doi.org/10.1016/j.jretconser.2012.07.006).

Bessi, A., and Ferrara, E. 2016. "Social Bots Distort the 2016 U.S. Presidential Election Online Discussion," *First Monday* (21:11).

Beuth, P. 2017. "Social Bots - Furcht vor den neuen Wahlkampfmaschinen," *Zeit Online*, , January 23. (, accessed November 28, 2020).

Brahnam, S., and De Angeli, A. 2012. "Gender Affordances of Conversational Agents," *Interacting with Computers* (24:3), pp. 139–153. (https://doi.org/10.1016/j.intcom.2012.05.001).

vom Brocke, J., Simons, A., Riemer, K., Niehaves, B., Plattfaut, R., and Cleven, A. 2015. "Standing on the Shoulders of Giants: Challenges and Recommendations of Literature Search in Information Systems Research," *Communications of the Association for Information Systems* (37). (https://doi.org/10.17705/1CAIS.03709).

Bruns, A. 2019. "After the 'APIcalypse': Social Media Platforms and Their Fight against Critical Scholarly Research," *Information, Communication & Society* (22:11), Routledge, pp. 1544–1566. (https://doi.org/10.1080/1369118X.2019.1637447).

Caldwell, M., Andrews, J. T. A., Tanay, T., and Griffin, L. D. 2020. "AI-Enabled Future Crime," *Crime Science* (9:1), p. 14. (https://doi.org/10.1186/s40163-020-00123-8).

Candello, H., Pinhanez, C., and Figueiredo, F. 2017. "Typefaces and the Perception of Humanness in Natural Language Chatbots," in *Proceedings of the 2017 CHI Conference*

on Human Factors in Computing Systems, CHI '17, New York, NY, USA: Association for Computing Machinery, pp. 3476–3487. (https://doi.org/10.1145/3025453.3025919).

Chhonker, M. S., Verma, D., Kar, A. K., and Grover, P. 2018. "M-Commerce Technology Adoption: Thematic and Citation Analysis of Scholarly Research during (2008-2017)," *The Bottom Line* (31:3/4), Emerald Publishing Limited, pp. 208–233. (https://doi.org/10.1108/BL-04-2018-0020).

Christy, K. R., Jensen, J. D., Britt, B., Scherr, C. L., Jones, C., and Brown, N. R. 2019. "I Want to Talk to a Real Person: Theorising Avoidance in the Acceptance and Use of Automated Technologies," *Health Systems* (8:1), pp. 31–43. (https://doi.org/10.1080/20476965.2017.1406568).

Ciechanowski, L., Przegalinska, A., Magnuski, M., and Gloor, P. 2019. "In the Shades of the Uncanny Valley: An Experimental Study of Human–Chatbot Interaction," *Future Generation Computer Systems* (92), pp. 539–548. (https://doi.org/10.1016/j.future.2018.01.055).

Dellermann, D., Ebel, P., Söllner, M., and Leimeister, J. M. 2019. "Hybrid Intelligence," *Business & Information Systems Engineering* (61:5), pp. 637–643. (https://doi.org/10.1007/s12599-019-00595-2).

Dwivedi, Y. K., Rana, N. P., Jeyaraj, A., Clement, M., and Williams, M. D. 2019. "Re-Examining the Unified Theory of Acceptance and Use of Technology (UTAUT): Towards a Revised Theoretical Model," *Information Systems Frontiers* (21:3), pp. 719–734. (https://doi.org/10.1007/s10796-017-9774-y).

Edwards, C., Edwards, A., Spence, P. R., and Shelton, A. 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), pp. 372–376. (https://doi.org/10.1016/j.chb.2013.08.013).

Elyashar, A., Fire, M., Kagan, D., and Elovici, Y. 2015. "Guided Socialbots: Infiltrating the Social Networks of Specific Organizations' Employees," *AI Communications* (29:1), pp. 87–106. (https://doi.org/10.3233/AIC-140650).

Feine, J., Gnewuch, U., Morana, S., and Maedche, A. 2019. "A Taxonomy of Social Cues for Conversational Agents," *International Journal of Human-Computer Studies* (132), pp. 138–161. (https://doi.org/10.1016/j.ijhcs.2019.07.009).

Ferrara, E. 2017. "Disinformation and Social Bot Operations in the Run up to the 2017 French

Presidential Election," *First Monday* (22:8). (https://doi.org/10.5210/fm.v22i8.8005).

Ferrara, E., Varol, O., Davis, C., Menczer, F., and Flammini, A. 2016. "The Rise of Social Bots," *Communications of the ACM* (59:7), pp. 96–104. (https://doi.org/10.1145/2818717).

Flores-Saviaga, C., Savage, S., and Taraborelli, D. 2016. "LeadWise: Using Online Bots to Recruit and Guide Expert Volunteers," in *Proceedings of the ACM Conference on Computer Supported Cooperative Work, CSCW* (Vol. 26-Februar), pp. 257–260. (https://doi.org/10.1145/2818052.2869106).

Frey, C. B., and Osborne, M. A. 2017. "The Future of Employment: How Susceptible Are Jobs to Computerisation?," *Technological Forecasting and Social Change* (114), pp. 254–280. (https://doi.org/10.1016/j.techfore.2016.08.019).

Gnewuch, U., Morana, S., Adam, M. T. P., and Maedche, A. 2018. "Faster Is Not Always Better: Understanding the Effect of Dynamic Response Delays in Human-Chatbot Interaction," in *Proceedings of the 26th European Conference on Information Systems*, Portsmouth.

Gnewuch, U., Morana, S., and Maedche, A. 2017. "Towards Designing Cooperative and Social Conversational Agents for Customer Service," in *Proceedings of the International Conference on Information Systems (ICIS)*, Seoul, p. 14.

Goga, O., Venkatadri, G., and Gummadi, K. P. 2015. "The Doppelgänger Bot Attack: Exploring Identity Impersonation in Online Social Networks," in *Proceedings of the ACM SIGCOMM Internet Measurement Conference, IMC* (Vol. 2015-Octob), pp. 141–153. (https://doi.org/10.1145/2815675.2815699).

Gruenberg, S. M. 1935. "Radio and the Child," *The ANNALS of the American Academy of Political and Social Science* (177:1), pp. 123–128. (https://doi.org/10.1177/000271623517700118).

Guzman, A. L. 2017. "Making AI Safe for Humans: A Conversation with Siri," in *Socialbots and Their Friends. Digital Media and the Automation of Sociality*, New York: Routledge, pp. 69–85.

Han, S., and Yang, H. 2018. "Understanding Adoption of Intelligent Personal Assistants: A Parasocial Relationship Perspective," *Industrial Management & Data Systems* (118:3), Emerald Publishing Limited, pp. 618–636. (https://doi.org/10.1108/IMDS-05-2017-0214).

Haustein, S., Bowman, T. D., Holmberg, K., Tsou, A., Sugimoto, C. R., and 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), pp. 232–238. (https://doi.org/10.1002/asi.23456).

Hegelich, S., and Janetzko, D. 2016. "Are Social Bots on Twitter Political Actors? Empirical Evidence from a Ukrainian Social Botnet," in *Proceedings of the 10th International Conference on Web and Social Media*, Cologne, pp. 579–582.

Hinz, O., van der Aalst, W. M. P., and Weinhardt, C. 2019. "Blind Spots in Business and Information Systems Engineering," *Business & Information Systems Engineering* (61:2), pp. 133–135. (https://doi.org/10.1007/s12599-019-00587-2).

Hofeditz, L., Ehnis, C., Bunker, D., Brachten, F., and Stieglitz, S. 2019. "Meaningful Use of Social Bots? Possible Applications in Crisis Communication during Disasters," in *Proceedings of the Twenty-Seventh European Conference on Information Systems*, Stockholm-Uppsala, Sweden, pp. 1–16.

Horkheimer, M., and Adorno, T. W. 2002. *Dialectic of Enlightenment*, (E. Jephcott, tran.), Stanford: Stanford University Press. (http://www.sup.org/books/title/?id=1103).

Janssen, A., Passlick, J., Rodríguez Cardona, D., and Breitner, M. H. 2020. "Virtual Assistance in Any Context: A Taxonomy of Design Elements for Domain-Specific Chatbots," *Business & Information Systems Engineering* (62:3), pp. 211–225. (https://doi.org/10.1007/s12599-020-00644-1).

Kallis, G., Kerschner, C., and Martinez-Alier, J. 2012. "The Economics of Degrowth," *The Economics of Degrowth* (84), pp. 172–180. (https://doi.org/10.1016/j.ecolecon.2012.08.017).

Kang, S.-H., and Gratch, J. 2011. "People like Virtual Counselors That Highly-Disclose about Themselves.," *Studies in Health Technology and Informatics* (167), Netherlands, pp. 143–148.

Kerschner, C., and Ehlers, M.-H. 2016. "A Framework of Attitudes towards Technology in Theory and Practice," *Ecological Economics* (126), pp. 139–151. (https://doi.org/10.1016/j.ecolecon.2016.02.010).

Lee, J. D., and See, K. A. 2004. "Trust in Automation: Designing for Appropriate Reliance," *Human Factors* (46:1), pp. 50–80.

Lugrin, B., Rosenthal-von der Pütten, A., and Hahn, S. 2019. "Identifying Social Context

Factors Relevant for a Robotic Elderly Assistant," in *Social Robotics*, M. A. Salichs, S. S. Ge, E. I. Barakova, J.-J. Cabibihan, A. R. Wagner, Á. Castro-González, and H. He (eds.), Cham: Springer International Publishing, pp. 558–567.

Maedche, A., Legner, C., Benlian, A., Berger, B., Gimpel, H., Hess, T., Hinz, O., Morana, S., and Söllner, M. 2019. "AI-Based Digital Assistants: Opportunities, Threats, and Research Perspectives," *Business & Information Systems Engineering* (61:4), pp. 535–544. (https://doi.org/10.1007/s12599-019-00600-8).

McTear, M. F. 2017. "The Rise of the Conversational Interface: A New Kid on the Block?," in *Future and Emerging Trends in Language Technology. Machine Learning and Big Data* (Vol. 10341), Lecture Notes in Computer Science, J. F. Quesada, F.-J. Martín Mateos, and T. López Soto (eds.), Cham: Springer International Publishing, pp. 38–49. (https://doi.org/10.1007/978-3-319-69365-1_3).

Moore, R. K. 2013. "Spoken Language Processing: Where Do We Go from Here?," in *Your Virtual Butler: The Making-Of*, R. Trappl (ed.), Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 119–133. (https://doi.org/10.1007/978-3-642-37346-6_10).

Mori, M., MacDorman, K. F., and Kageki, N. 2012. "The Uncanny Valley [From the Field]," *IEEE Robotics & Automation Magazine* (19:2), pp. 98–100. (https://doi.org/10.1109/MRA.2012.2192811).

Munger, K. 2017. "Tweetment Effects on the Tweeted: Experimentally Reducing Racist Harassment," *Political Behavior* (39:3), pp. 629–649. (https://doi.org/10.1007/s11109-016-9373-5).

Mustafaraj, E., and Metaxas, P. T. 2010. "From Obscurity to Prominence in Minutes: Political Speech and Real-Time Search," in *WebSci10: Extending the Frontiers of Society On-Line*, pp. 1–7.

Nass, C., and Moon, Y. 2000. "Machines and Mindlessness: Social Responses to Computers," *Journal of Social Issues* (56:1), John Wiley & Sons, Ltd, pp. 81–103. (https://doi.org/10.1111/0022-4537.00153).

Nof, S. Y. 2009. "Automation: What It Means to Us Around the World," in *Springer Handbook of Automation*, S. Y. Nof (ed.), Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 13–52. (https://doi.org/10.1007/978-3-540-78831-7_3).

Orben, A. 2020. "The Sisyphean Cycle of Technology Panics," *Perspectives on Psychological Science* (15:5), p. 15.

Palanca, J., Val, E. del, Garcia-Fornes, A., Billhardt, H., Corchado, J. M., and Julián, V. 2018. "Designing a Goal-Oriented Smart-Home Environment," *Information Systems Frontiers* (20:1), pp. 125–142. (https://doi.org/10.1007/s10796-016-9670-x).

Papsdorf, C. 2015. "How the Internet Automates Communication," *Information, Communication & Society* (18:9), pp. 991–1005. (https://doi.org/10.1080/1369118X.2015.1008539).

Plato. 1997. *Complete Works*, (J. M. Cooper and D. S. Hutchinson, eds.), Indianapolis, Ind: Hackett Pub.

Quarteroni, S. 2018. "Natural Language Processing for Industry," *Informatik-Spektrum* (41:2), pp. 105–112. (https://doi.org/10.1007/s00287-018-1094-1).

Ratkiewicz, J., Conover, M. D., Meiss, M., Gonçalves, B., Flammini, A., and Menczer, F. 2011. "Detecting and Tracking Political Abuse in Social Media," in *Proceedings of the Fifth International AAAI Conference on Weblogs and Social Media*, Barcelona, Spain, pp. 297–304.

Reeves, B., and Nass, C. I. 1996. *The Media Equation: How People Treat Computers, Television, and New Media like Real People and Places.*, New York, NY, US: Cambridge University Press, pp. xiv, 305.

Rosenthal-von der Pütten, A. M., and Krämer, N. C. 2014. "How Design Characteristics of Robots Determine Evaluation and Uncanny Valley Related Responses," *Computers in Human Behavior* (36), pp. 422–439. (https://doi.org/10.1016/j.chb.2014.03.066).

Sah, Y. J., and Peng, W. 2015. "Effects of Visual and Linguistic Anthropomorphic Cues on Social Perception, Self-Awareness, and Information Disclosure in a Health Website," *Computers in Human Behavior* (45), pp. 392–401. (https://doi.org/10.1016/j.chb.2014.12.055).

Salto Martínez, R., and Jacques-García, F. A. 2012. "Development and Implementation of a Chat Bot in a Social Network," in *9th International Conference on Information Technology - New Generations*, , April, pp. 751–755. (https://doi.org/10.1109/ITNG.2012.147).

Schlögl, S., Postulka, C., Bernsteiner, R., and Ploder, C. 2019. "Artificial Intelligence Tool Penetration in Business: Adoption, Challenges and Fears," in *Knowledge Management in Organizations*, L. Uden, I.-H. Ting, and J. M. Corchado (eds.), Cham: Springer International Publishing, pp. 259–270.

Sebastian, S., Ayyappan, S., and Vinod, P. 2014. "Framework for Design of Graybot in Social Network," in *Proceedings of the 2014 International Conference on Advances in Computing, Communications and Informatics, ICACCI 2014*, pp. 2331–2336. (https://doi.org/10.1109/ICACCI.2014.6968575).

Seiffert, J. 2016. "Meinungsmache Im Netz. Wie Social Bots Falschmeldungen Verbreiten," *Deutschlandfunk*, , December 14. (https://www.deutschlandfunk.de/meinungsmache-im-netz-wie-social-bots-falschmeldungen.724.de.html?dram:article_id=373940, accessed November 28, 2020).

Shah, H., Warwick, K., Vallverdú, J., and Wu, D. 2016. "Can Machines Talk? Comparison of Eliza with Modern Dialogue Systems," *Computers in Human Behavior* (58), pp. 278–295. (https://doi.org/10.1016/j.chb.2016.01.004).

Sickert, T. 2017. "Bot Oder Kein Bot. So Erkennen Sie Meinungsroboter.," *SPIEGEL Online*, , January 19. (https://www.spiegel.de/netzwelt/web/social-bots-entlarven-so-erkennen-sie-meinungsroboter-a-1129539.html, accessed November 28, 2020).

Söllner, M., Hoffmann, A., and Leimeister, J. M. 2016. "Why Different Trust Relationships Matter for Information Systems Users," *European Journal of Information Systems* (25:3), pp. 274–287.

Srivastava, S. C., Chandra, S., and Shirish, A. 2015. "Technostress Creators and Job Outcomes: Theorising the Moderating Influence of Personality Traits," *Information Systems Journal* (25:4), pp. 355–401.

Stieglitz, S., Meske, C., Ross, B., and Mirbabaie, M. 2020. "Going Back in Time to Predict the Future - The Complex Role of the Data Collection Period in Social Media Analytics," *Information Systems Frontiers* (22:2), pp. 395–409. (https://doi.org/10.1007/s10796-018-9867-2).

Stieglitz, S., Mirbabaie, M., Ross, B., and Neuberger, C. 2018. "Social Media Analytics – Challenges in Topic Discovery, Data Collection, and Data Preparation," *International Journal of Information Management* (39), pp. 156–168. (https://doi.org/10.1016/j.ijinfomgt.2017.12.002).

Subrahmanian, V. S., Azaria, A., Durst, S., Kagan, V., Galstyan, A., Lerman, K., Zhu, L., Ferrara, E., Flammini, A., and Menczer, F. 2016. "The DARPA Twitter Bot Challenge," *Computer* (49:6), pp. 38–46. (https://doi.org/10.1109/MC.2016.183).

Sutton, T. 2017. "Disconnect to Reconnect: The Food/Technology Metaphor in Digital

Detoxing," *First Monday* (22:6). (https://doi.org/10.5210/fm.v22i6.7561).

Tsvetkova, M., García-Gavilanes, R., Floridi, L., and Yasseri, T. 2017. "Even Good Bots Fight: The Case of Wikipedia," *Plos One* (12:2), p. e0171774. (https://doi.org/10.1371/journal.pone.0171774).

Veale, T., Valitutti, A., and Li, G. 2015. "Twitter: The Best of Bot Worlds for Automated Wit," in *Distributed, Ambient, and Pervasive Interactions*, N. Streitz and P. Markopoulos (eds.), Cham: Springer International Publishing, pp. 689–699.

Wang, A. H. 2010. "Detecting Spam Bots in Online Social Networking Sites: A Machine Learning Approach," in *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* (Vol. 6166 LNCS), pp. 335–342. (https://doi.org/10.1007/978-3-642-13739-6_25).

Wartella, E., and Robb, M. 2008. "Historical and Recurring Concerns about Children's Use of the Mass Media," in *The Handbook of Children, Media, and Development*, S. L. Calvert and M. Robb (eds.), John Wiley & Sons, Ltd, pp. 5–26. (https://doi.org/10.1002/9781444302752.ch1).

Weizenbaum, J. 1966. "ELIZA—a Computer Program for the Study of Natural Language Communication between Man and Machine," *Communications of the ACM* (9:1), New York, NY, USA: Association for Computing Machinery, pp. 36–45. (https://doi.org/10.1145/365153.365168).

Wilkie, A., Michael, M., and Plummer-Fernandez, M. 2015. "Speculative Method and Twitter: Bots, Energy and Three Conceptual Characters," *Sociological Review* (63:1), pp. 79–101. (https://doi.org/10.1111/1467-954X.12168).

Yan, Z., Dong, Y., Niemi, V., and Yu, G. 2013. "Exploring Trust of Mobile Applications Based on User Behaviors: An Empirical Study," *Journal of Applied Social Psychology* (43:3), pp. 638–659. (https://doi.org/10.1111/j.1559-1816.2013.01044.x).

Yin, P., Ram, N., Lee, W.-C., Tucker, C., Khandelwal, S., and Salathé, M. 2014. "Two Sides of a Coin: Separating Personal Communication and Public Dissemination Accounts in Twitter," *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* (Vol. 8443 LNAI). (https://doi.org/10.1007/978-3-319-06608-0_14).

Appendix

Paper 1: Do Social Bots Dream of Electric Sheep? A Categorisation of Social Media Bot Accounts

Fact sheet

Type	Conference Paper
Authors	Stieglitz, Stefan; Brachten, Florian; Ross, Björn & Jung, Anna-Katharina
Year	2017
Venue	Australasian Conference on Information Systems (ACIS)
Full citation	Stieglitz, S., Brachten, F., Ross, B., and Jung, A.-K. 2017. "Do Social Bots Dream of Electric Sheep? A Categorisation of Social Media Bot Accounts," in *Proceedings of the Australasian Conference on Information Systems (ACIS),* Paper 206.

Do Social Bots Dream of Electric Sheep? A Categorisation of Social Media Bot Accounts

Stefan Stieglitz
Department of Computer Science and Applied Cognitive Science
University of Duisburg-Essen
Duisburg, Germany
Email: stefan.stieglitz@uni-due.de

Florian Brachten
Department of Computer Science and Applied Cognitive Science
University of Duisburg-Essen
Duisburg, Germany
Email: florian.brachten@uni-due.de

Björn Ross
Department of Computer Science and Applied Cognitive Science
University of Duisburg-Essen
Duisburg, Germany
Email: bjoern.ross@uni-due.de

Anna-Katharina Jung
Department of Computer Science and Applied Cognitive Science
University of Duisburg-Essen
Duisburg, Germany
Email: anna-katharina.jung@uni-due.de

Abstract

So-called 'social bots' have garnered a lot of attention lately. Previous research showed that they attempted to influence political events such as the Brexit referendum and the US presidential elections. It remains, however, somewhat unclear what exactly can be understood by the term 'social bot'. This paper addresses the need to better understand the intentions of bots on social media and to develop a shared understanding of how 'social' bots differ from other types of bots. We thus describe a systematic review of publications that researched bot accounts on social media. Based on the results of this literature review, we propose a scheme for categorising bot accounts on social media sites. Our scheme groups bot accounts by two dimensions – Imitation of human behaviour and Intent.

Keywords social bots, social media, categorisation, bots.

1 Introduction

Social media permeate society. Brands use them to influence customers' purchase intentions (Xie and Lee 2015) and political candidates use them to disseminate information to their supporters, but their unregulated nature has given rise to a flood of information of questionable credibility (Wattal et al. 2010). It has been shown that businesses such as hotels are posting manipulated content on social media to gain an unfair advantage over their competitors (Mayzlin et al. 2014). Against this backdrop, it becomes clear that social media have increasingly become interesting for people and organisations looking to influence the discussion on a certain topic. The automated dissemination of messages promises to be an efficient way to reach many people with little effort. The reasons for spreading automated content range from the dissemination of information (e.g. stock prices, weather data), spam, malware or advertisement to political intentions (Alarifi et al. 2016). Recently social bots, algorithms programmed to mimic human behaviour on social media platforms, have become increasingly attractive for people and organisations aiming to automatically distribute their messages to many recipients at very low costs. Current studies reveal that social bots are involved in online discussions about current political events, such as the armed conflict between Ukraine and Russia and the war in Syria by spamming the discussion with one-sided arguments or unrelated content to distract participants (Abokhodair et al. 2015; Hegelich and Janetzko 2016). The mere presence of automated actors in vital opinion-shaping discussions provokes the fear of manipulation and thus ethical concerns. This has led to increasing press coverage on the expected influence of social bots (e.g. Dewey 2016; Fuchs 2016; Guilbeault and Woolley 2016; Lobe 2016). The great public interest in social bots underlines the importance of a profound scientific analysis of the topic.

As the topic of social bots is still young, and as it is approached from multiple angles, the terms and definitions used to describe related phenomena are diverse. Sometimes several different terms are used to label the same concept, and sometimes a single term such as social bots is used to describe different things. This leads to a diffuse use of terms and subsequently to imprecise theoretical foundations in this area. For example, some researchers use the term "social bot" for any account on social media run by an algorithm (e.g. Forelle et al. 2015), while others use a much more restrictive definition, for example as "computer programs designed to use social networks by simulating how humans communicate and interact with each other" (Abokhodair et al. 2015, p.25). Some researchers consider a social bot a potentially harmful adversary by definition (Boshmaf et al. 2013). This confusion around terms and definitions means that there is a clear need for a structured approach to the topic of social bots.

At the same time, the prevalence of bots on social media raises interesting research questions and challenges for the IS research community. Identification techniques, communication patterns and the impact of social bots on individuals and companies are only three examples of possible research topics in IS. Our literature review shows that there has been very little IS research into these topics so far. Given that this topic is clearly highly relevant for the IS community, there is a noticeable research gap. To begin addressing this gap and provide guidance for future research, it is first necessary to clearly delineate the types and activities of bots on social media.

This paper summarises the types of bot accounts active in social media. Moreover, we discuss definitions and terms that are used in academic articles. In order to structure different types of social bots, we develop a categorisation scheme that builds upon our findings from literature. The literature review will (a) contribute to the specification of the field and will offer an excellent starting point for researchers who plan to start investigating bots on social media, (b) allow new forms of bots to be assessed faster through a categorisation system, as it offers a scheme to group and classify them and (c) provide a clear definition of social bots which demarcates them from other forms of automated actors in social media.

2 Background

This article concerns bots that run or control social media accounts. We do not consider bots that make use of social media features but do not control their own accounts, e.g. botnets that communicate by surreptitiously injecting messages into photos uploaded by the user (Nagaraja et al. 2011).

Even within this relatively narrow field of research that concentrates on bots in the context of social media, there is an enormous diversity of bots. Moreover, researchers from different backgrounds tend to approach bots from different angles using various theories and concepts. Information security

researchers view bots in an adversarial role, and demonstrate the feasibility of hypothetical attacks or devise potential defence mechanisms. For example, Pantic and Husain who research botnets that communicate by actively controlling their own social media accounts state that "Botnet software is a type of malicious software (malware) that is most often placed on a victim's computer silently" (2015, p. 172). Researchers in journalism explore how useful news reporting bots are transforming their field (Lokot and Diakopoulos 2016). Social scientists may place their own bots to explore how humans react to them (Wilkie et al. 2015). It becomes obvious that researchers in different disciplines develop their own unique perspectives and theoretical foundations.

The resulting confusion extends to the terminology, which is equally diverse. For example, a large number of the papers concerning bots on Twitter (or Twitter bots) address political goals and consider social bots in this context, e.g. during the Syrian war (Abokhodair et al. 2015), the crisis in Ukraine (Hegelich and Janetzko 2016), Venezuelan politics (Forelle et al. 2015) and regional elections in Germany (Brachten et al. 2017). While these authors are all interested in examining how bots can be identified and the extent to which they are used in practice, there seems to be disagreement in naming the researched aspects. Some authors use similar definitions of social bots: Abokhodair et al. define them as "software designed to act in ways that are similar to how a person would act in the social space" (2015, p. 840), and Hegelich and Janetzko call them "automatic programs [that] are mimicking humans" (2016, p. 579). Forelle et al. describe social bots as "computer-generated programs that post, tweet, or message of their own accord" (2015, p.1). While the first two sources both point out that those bots imitate humans, this aspect is missing in the third citation. Boshmaf et al. (2013) mention a further component: According to them, a social bot "is an automation software that controls an adversary-owned or hijacked account on a particular OSN, and has the ability to perform basic activities such as posting a message and sending a connection request" (p. 556). Here again the aspect of imitating a human user is missing, but the authors mention hijacked accounts, which were not prevalent in the other definitions. Another important difference is, again, the reference to adversaries implying that social bots are, by definition, opponents. This example demonstrates how terminological ambiguity and the lack of a shared conceptual understanding go hand in hand.

Aside from the lacking consensus on what exactly is to be understood by the term "social bots", several papers use other labels to describe phenomena which could be understood as social bots according to some of the above definitions. Igawa et al. (2016) write, that "[o]n Twitter social robots, called 'bots', pretend to be human beings in order to gain followers and replies from target users and promotes a product or agenda" (p. 73). The definition describes features of social bots (respectively goals of the bot developer) simply calling the relevant accounts "bots". The same goes for Larsson and Moe (2015) who define bots as "a piece of more or less automated computer software, programmed to mimic the behaviour of human Internet users" (p. 362). While their definition is more general as it does not limit the accounts to Twitter, it also features the aspect of human imitation.

These examples show that a broad consensus is missing but needed to precisely describe the researched aspects. In order to reach this goal, a comprehensive overview over the relevant literature is needed. The exact procedure is described in the next chapter.

3 Method

The following literature review was conducted based on the systematic process proposed by vom Brocke et al. (2009), which includes the taxonomy for literature reviews by Cooper (1988). Primarily the scope of the research was limited to bots that run or control social media accounts regardless of their intentions and methods. Information Systems (IS) researchers and general scholars have been defined as the main audience of the literature review.

To gather relevant literature, we first conducted a search in three databases: Scopus, ScienceDirect and the AIS Electronic Library (AISeL). While Scopus and ScienceDirect allow us to identify research articles on a broader level, AISeL is a source explicitly used by academics from IS. First, we searched for literature that either included the term bot or socialbot and one of the terms Twitter, Facebook, "social networks" or "social media" in the title, abstract or keywords. To get a systematic overview of recent high-quality academic publications, we limited the search to peer-reviewed articles that were released in or after 2007. Due to our own language skills and comparability we only consider articles that were written in English. To ensure the scientific quality of the publications, we only considered peer-reviewed papers from conference proceedings and scientific journals.

The search was carried out on 31 July, 2017. We found 187 entries in the Scopus database and 19 in ScienceDirect. The search in AISeL led to no results. Within the 187 Scopus entries we found three

duplicates (due to different titles but identical contents) which were excluded. We then matched the entries from Scopus and ScienceDirect to filter out duplicates found in both databases, which led to the exclusion of twelve entries.

As a next step the titles and abstracts of the remaining 191 entries were examined in more detail to assess if they were relevant for our main goal of observation. Of the 191 entries, 68 were excluded unanimously from further investigation due to irrelevance (e.g. one paper that dealt with search engine optimisation used the term bot in reference to the Google crawler and the term social media in its keywords). Of the remaining entries, 88 were included unanimously while for 35 entries the authors' evaluation differed. Those entries were re-evaluated in a group discussion. This discussion led to 16 out of the 35 entries being included, forming a sample of 104 papers. Since one paper could not be obtained, the final sample consisted of 103 papers.

Figure 1 shows the number of papers published by year on the topic of bot accounts in social media. As the figure shows, there is a fairly consistent increase in the number of papers published. The year 2014 can be described as an outlier. However, closer inspection did not reveal why the number of publications dropped in this year. For 2017, only the first seven months (January to July) are present, and the number is therefore lower than the peak in 2016. Also, the final sample consisted only of papers released in or after 2010 as no papers were identified between 2007–2009.

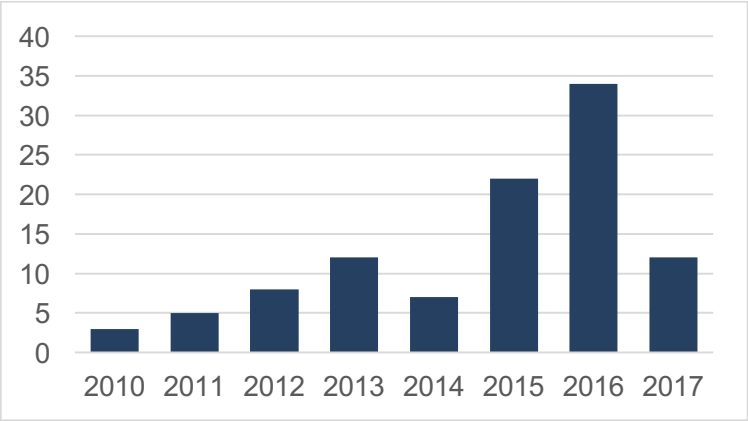

Figure 1: Number of papers published in each year in the examined sample

4 A categorisation of social media bot accounts

How can the bots active on social media be distinguished from one another? Two distinctions are commonly made in the literature.

1. First, bots are distinguished into benign and malicious bots (Ferrara et al. 2016). Benign bots aggregate content, respond automatically, and perform other useful services. Malicious bots, in contrast, are designed with a purpose to harm. Our analysis of the literature shows that the categorisation of social bots into malicious and benign ones is widely accepted.
2. The second distinction can be traced to Boshmaf et al. (2013), who wrote that the crucial difference between social bots and other types of bots is that the former "is designed to pass itself off as a human being. This is achieved by either simply mimicking the actions of a real OSN user or by simulating such a user using artificial intelligence" (p. 556).

In the following, we describe these categories in more detail, show how they were applied in related papers, and give specific examples. We also show that these distinctions are orthogonal to each other, thus giving rise to a two-dimensional categorisation of bots.

4.1 Intentions – malicious or benign?

Due to the multiple research perspectives to study (social) bots, the literature offers a wide range of definitions, terms and classifications. The most common distinction is that between benign and malicious bots (Ferrara et al. 2016). 52 out of the 103 papers used the term malicious to classify certain bot accounts or describe their behaviour (e.g. Boshmaf et al. 2016; Chu et al. 2012; Edwards et al. 2014; He et al. 2017). For example, Freitas et al. point out that "Socialbots can have many applications, with good or malicious objectives" (2015, p. 25). Bots which spread spam or falsehoods are generally defined as malicious (Bhat and Abulaish 2013; Lokot and Diakopoulos 2016; Main and Shekokhar

2015). Bots delivering useful automated information such as news or weather reports, are usually labelled *good*, *helpful* or *benign* (Alarifi et al. 2016; Chu et al. 2012). Here, 20 papers use the latter term to label such bots.

The most frequently mentioned benign bots active in social media are those used by mass media, grassroot journalists and bloggers to automatically post recently published articles or breaking news. Lokot and Diakopoulos (2016) classify these news bots based on their input and sources, outputs, algorithms and intent or function. Weather bots, sport bots, traffic bots, niche news bots and geo-specific bots are only some examples of the news bots they identified through their research. Bots can contribute positively to the recruitment of volunteers (Flores-Saviaga et al. 2016; Savage et al. 2016). Flores-Saviaga et al. (2016) deployed two bots which contacted and motivated experts to mention women who were still missing an entry on Wikipedia. Furthermore Tsvetkova et al. (2017) report that recruitment bots, as well editing bots and anti-vandalism bots are frequently used on Wikipedia. Another type of benign bots is chat bots, which can be used by enterprises to limit the need for human involvement in business-to-customer communication. In addition to that, chat bots can be used to respond to customer questions during events (Salto Martínez and Jacques García 2012).

However, the ill intents of bots are much more diverse. They include: spam, the theft of personal data and identities, the spreading of misinformation and noise during debates, the infiltration of companies and the diffusion of malware (Abokhodair et al. 2015; Bhat and Abulaish 2013; Bokobza et al. 2015; Elyashar et al. 2015; Goga et al. 2015; Zhang et al. 2013; Zhu et al. 2013). Previous literature has introduced specific terms to describe bots involved in certain malicious acts. *Commercial bots*, according to Subrahmanian et al. (2016), include e.g. spam bots, and pay bots "which copy content from respected sources and paste it into micro URLs that pay the bot creator for direct traffic to that site" (p. 38). Bots which operate on social media with a fake identity or have the aim of impersonation are often described as *sybils* – 18% of the examined papers used that term (e.g. Goga et al. 2015; Paradise et al. 2014). According to Goga et al. (2015), the three main types of identity attacks are celebrity impersonation attacks (duplication of a celebrity account by a sybil account), social engineering attacks (which aim at motivating friends/followers to disclose private data) or doppelgänger bot attacks (which are copies of user accounts, in order to use their identity to slip through the networks detection systems). However, the term *sybil* is not clearly distinguishable from the term *bot*, as bots operate with fake identities, too. Therefore, the terms *sybil* and *bot* are often used interchangeably.

Additional terms which are linked to the malicious behaviour of bots are *astroturfing*, *misdirection* and *smoke screening*. Ratkiewicz et al. (2011) describe political astroturf as "political campaigns disguised as spontaneous 'grassroots' behaviour that are in reality carried out by a single person or organisation. This is related to spam but with a more specific domain context, and potentially larger consequences" (p. 297). Astroturf is often linked to the intention to influence the opinion in a political debate and to create the impression that a vast majority is in favour of a certain position. In contrast to that, *smoke screening* entails the use of context-related hashtags on Twitter, to distract the readers from the main point of the debate (e.g. to use the hashtag *#syria* but talk about something unrelated to the war) (Abokhodair et al. 2015). *Misdirection* is similar to smoke screening, but goes a step further by using context-related hashtags without referring to the topic at all (e.g. use *#syria* but talk about something which is not related to Syria) (Abokhodair et al. 2015). All three forms of bot attacks can lead to a misconception of events and can influence e.g. the popularity of certain hashtags and topics in the related network. Subrahmanian et al. (2016) describe bots which seek to influence public discourse as "influence bots". Although influence bots have mainly been investigated in the context of political debates, it is easy to see how they could be used in a commercial context. The spreading of negative opinions about a certain product, brand or service can lead to a distorted perception of that product in the public opinion and a lasting damage for the enterprise. Whether in a political or a commercial context, as soon as bots aim at distorting the public perception, ethical concerns arise.

Finally, there are also bots that merely *are* – without being outright malicious or benign. The humoristic Twitter bots, as described by Veale et al. (2015), exemplify this type. Their only aim is to create funny and linguistically correct posts. Veale et al. identified two generations in the development of linguistic bots. While those of the first generation only made use of superficial language resources and did not manipulate text on a semantic level, second-generation bots apply semantic techniques and theories. Therefore, second-generation bots are linguistically more difficult to distinguish from human users. For these bots, we propose a third, *neutral* level for the intent, between malicious and benign.

4.2 Imitation of human behaviour

Besides the distinction between malicious and benign bots, another commonly made distinction is between bots that mimic human behaviour and those that do not (Boshmaf et al. 2013). Bots which pretend to be human users are often referred to as *social bots* (Boshmaf et al. 2011; He et al. 2017; Hegelich and Janetzko 2016; Igawa et al. 2016; Stieglitz et al. 2017). Abokhodair et al. (2015) describe social bots as automated social actors, which differ in their social skills and their intentions. They consider those automated social actors as dangerous which aim at being recognised as humans or companies, and which highlight one point of view to establish the impression of uniformity of opinions. Some of the previously described benign bots imitate human behaviour to some degree, too. For example, the humoristic bots described by Veale et al. (2015) produce messages which are very similar to content produced by human users. However, those bots do not try to hide that they are based on an algorithm. Often the profile description even states clearly that they are bots. As the boundaries between human and non-human behaviour are fading in some cases, some authors introduced the term *cyborg* to describe accounts which cannot be clearly categorised. These accounts can be humans who make use of automation techniques or bots which are managed by human beings (Chu et al. 2010). As social bots become more and more sophisticated, constantly advancing their cover-up techniques, their detection becomes increasingly challenging (Chavoshi et al. 2017; Everett et al. 2016). Researchers who programmed social bots report that social network operators were slow to identify and remove their bots (Boshmaf et al. 2011; Freitas et al. 2015).

As mentioned before, not all bots use sophisticated strategies in order to appear human. This is also true for accounts referred to as malicious. *Spam bots* often publish a large number of nearly identical messages in a short time. To the human observer, it is immediately clear that they are bots. Additionally, some bots that attempt to engage and converse with humans are based on simple rules. For example, on Twitter they might respond to all tweets that mention a keyword out of a predefined set by tweeting a generic response (Salto Martínez and Jacques García 2012). Moon (2017) also describes several accounts that do not attempt to hide that they are bots. The above considerations directly lead to the following categorisation of bot accounts on social media.

As we have shown, there are many different types of bots. For researchers, it might be quite helpful to distinguish between different classes of bots on social media because e.g. they have a different impact on communication on social media or because they require different approaches to be identified.

Also, as we have mentioned before, the terms used to describe different types of bots in the social media context are often imprecise – especially in the context of social bots. We argue that not every bot on the social media is a *social* bot, and that the term *social media* does not automatically imply that every automated (bot) account on such a platform is by definition a *social* bot. Instead, the term *social* refers to the imitation of human behaviour and the act of pretending to be a human with whom a social interaction is possible, "to act in ways that are similar to how a person might act in the social space" (Abokhodair et al. 2015, p. 840). Thus, we propose *imitation of human behaviour* to be the second dimension to discriminate between different kinds of bots on social media. Social bots are those bots that attempt to imitate humans to a large degree, while in contrast, a mere spam bot which only uses social media to disseminate a lot of messages exhibits a low degree of imitation.

These proposed dimensions are combined in the scheme shown in Table 1, which provides a way to organise different types of bots on social media. This categorisation covers accounts on social media sites that are controlled by bots, but differ regarding their intent and the degree to which they imitate human behaviour. Examples of social bots can be seen in the first row. They are social in the sense that they imitate human users to a high degree by writing original messages, sending friend requests, and sharing or retweeting information by other users. This definition is not limited to harmful accounts because helpful accounts can be as social, or even more social, than harmful ones. As we pointed out earlier, most definitions of social bots reflect these circumstances.

		Intent (Ferrara et al. 2016)		
		Malicious	Neutral	Benign
Imitation of human behaviour (Boshmaf et al. 2013)	High: *Social* bots	• Astroturfing bot (Ratkiewicz et al. 2011) • Social botnets in political conflicts (Abokhodair et al. 2015) • Infiltration of an organisation (Elyashar et al. 2015) • Influence bots (Subrahmanian et al. 2016) • Sybils (Alarifi et al. 2016; Goga et al. 2015) • Doppelgänger bots (Goga et al. 2015)	• Humoristic bots (Veale et al. 2015)	• Chat bots (Salto Martínez & Jacques García 2012)
	Low to none	• Spam bots (Wang 2010) • Fake accounts used for botnet command & control (Sebastian et al. 2014) • Pay bots (Subrahmanian et al. 2016)	• Nonsense bots (Wilkie et al. 2015)	• News bots (Lokot & Diakopoulos 2016) • Recruitment bots (Flores-Saviaga et al. 2016) • Public Dissemination Account (Yin et al. 2014) • Earthquake warning bots (Haustein et al. 2016) • Editing Bots, Anti-Vandalism Bots on Wikipedia (Tsvetkova et al. 2017)

Table 1. Categorisation scheme of social media bot accounts

5 Conclusion

This paper provides a comprehensive literature analysis on a very new topic that is becoming increasingly important for research and practice. Based on the insights from the review, a categorisation scheme was developed that includes and differentiates bots in social media on two dimensions: the intent (Ferrara et al. 2016) and the imitation of human behaviour (Boshmaf et al. 2013). This paper is the first to combine these two dimensions into a six-category system. For a coherent categorisation, the dimension of 'imitation of human behaviour' is divided into *high* and *low to none* while the dimension 'intent' includes *benign*, *neutral,* and *malicious*. By this means all social media bots, which have been analysed and discussed in the literature review could be classified. We follow Boshmaf et al. (2013) in defining *social* bots as bots which imitate human behaviour to a high degree, and give examples of such behaviour found in the literature. The literature review shows that the majority of papers on bots on social media address malicious bots. Social bots with a neutral or benign intent are an exception and are researched rarely. We do not assert that the proposed categorisation reflects the absolute truth, or is the only way to bring structure into the diversity of bots on social media. However, in our opinion it is a first helpful step for researchers and practitioners to categorise bots on social media.

Therefore, the first contribution of our article, which results from the systematic literature review, is to make this unstructured and heterogeneous research field more accessible. This article offers researchers an overview which will be especially helpful for those academics and practitioners who start investigating the phenomenon of social bots. Furthermore, researchers who already are engaged with the topic benefit from the categorisation, as it facilitates the localisation of the scope of their

research in that field. Second, our categorisation system contributes to the research field by allowing bots to be assessed and analysed faster. Finally, we pointed out what separates *social* bots from other types of bots, leading to a more unified understanding of the phenomenon which can serve as the starting point for further analyses.

In further investigations, we plan to test the presented categorisation system empirically, to further prove its practical applicability. Further research may also examine which types of bots exist in different domains of social media communication. While, for example, harmful human-like bots that seek to influence human behaviour might be more frequent in politics than in sports and art, humorous bots might be prevalent in entertainment but not involved in a discussion about a current crisis situation. In addition to that, further research in this area could identify and improve effective methods for identifying bots. In this sense, the categorisation scheme raises the question whether different techniques are more helpful for certain bot categories. It can also be tested if the comparatively low number of papers researching potentially helpful and neutral bots which mimic human behaviour is due to this type of bot actually being rare, or if there is simply less research interest in bots which do not potentially harm social media users.

Overall, in this paper we pointed out a way to structure research on bots in social media and contributed to a broader understanding of this topic, thus providing a foundation for a more focused approach for future research.

6 References

Abokhodair, N., Yoo, D., and McDonald, D. W. 2015. "Dissecting a Social Botnet," in *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing - CSCW '15*, pp. 839–851 (doi: 10.1145/2675133.2675208).

Alarifi, A., Alsaleh, M., and Al-Salman, A. 2016. "Twitter turing test: Identifying social machines," *Information Sciences*, (372), pp. 332–346 (doi: 10.1016/j.ins.2016.08.036).

Bhat, S. Y., and Abulaish, M. 2013. "Community-based features for identifying spammers in Online Social Networks," in *Proceedings of the 2013 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining, ASONAM 2013*, pp. 100–107 (doi: 10.1145/2492517.2492567).

Bokobza, Y., Paradise, A., Rapaport, G., Puzis, R., Shapira, B., and Shabtai, A. 2015. "Leak sinks: The threat of targeted social eavesdropping," in *Proceedings of the 2015 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining, ASONAM 2015*, pp. 375–382 (doi: 10.1145/2808797.2808891).

Boshmaf, Y., Logothetis, D., Siganos, G., Lería, J., Lorenzo, J., Ripeanu, M., Beznosov, K., and Halawa, H. 2016. "Íntegro: Leveraging victim prediction for robust fake account detection in large scale OSNs," *Computers and Security*, (61), pp. 142–168 (doi: 10.1016/j.cose.2016.05.005).

Boshmaf, Y., Muslukhov, I., Beznosov, K., and Ripeanu, M. 2011. "The Socialbot Network: When bots socialize for fame and money," in *ACM International Conference Proceeding Series*, pp. 93–102 (doi: 10.1145/2076732.2076746).

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

Brachten, F., Stieglitz, S., Hofeditz, L., Kloppenborg, K., and Reimann, A. 2017. "Strategies and Influence of Social Bots in a 2017 German state election," in *Proceedings of the Australasian Conference on Information Systems*.

vom Brocke, J., Simons, A., Niehaves, B., Riemer, K., Plattfaut, R., and Cleven, A. 2009. "Reconstructing the giant: On the importance of rigour in documenting the literature search process," in *17th European Conference on Information Systems (ECIS 2009)*, Verona, pp. 2206–2217 (available at http://aisel.aisnet.org/ecis2009%0Ahttp://aisel.aisnet.org/ecis2009/161).

Chavoshi, N., Hamooni, H., and Mueen, A. 2017. "DeBot: Twitter bot detection via warped correlation," in *IEEE International Conference on Data Mining, ICDM*, pp. 817–822 (doi: 10.1109/ICDM.2016.86).

Chu, Z., Gianvecchio, S., Wang, H., and Jajodia, S. 2010. "Who is tweeting on twitter: Human, bot, or cyborg?," in *Proceedings - Annual Computer Security Applications Conference, ACSAC*, pp. 21–30 (doi: 10.1145/1920261.1920265).

Chu, Z., Gianvecchio, S., Wang, H., and Jajodia, S. 2012. "Detecting automation of Twitter accounts: Are you a human, bot, or cyborg?," in *IEEE Transactions on Dependable and Secure Computing*, (Vol. 9), pp. 811–824 (doi: 10.1109/TDSC.2012.75).

Cooper, H. M. 1988. "Organizing knowledge syntheses: A taxonomy of literature reviews," *Knowledge in Society*, (1:1), p. 104 (doi: 10.1007/BF03177550).

Dewey, C. 2016. "One in four debate tweets comes from a bot. Here's how to spot them.," *Washington Post* (available at https://www.washingtonpost.com/news/the-intersect/wp/2016/10/19/one-in-four-debate-tweets-comes-from-a-bot-heres-how-to-spot-them/; retrieved November 17, 2016).

Edwards, C., Edwards, A., Spence, P. R., and Shelton, A. 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," in *Computers in Human Behavior*, (Vol. 33), pp. 372–376 (doi: 10.1016/j.chb.2013.08.013).

Elyashar, A., Fire, M., Kagan, D., and Elovici, Y. 2015. "Guided socialbots: Infiltrating the social networks of specific organizations' employees," *AI Communications*, (29:1), pp. 87–106 (doi: 10.3233/AIC-140650).

Everett, R. M., Nurse, J. R. C., and Erola, A. 2016. "The anatomy of online deception: What makes automated text convincing?," in *Proceedings of the ACM Symposium on Applied Computing*, (Vol. 04–08–Apri), pp. 1115–1120 (doi: 10.1145/2851613.2851813).

Ferrara, E., Varol, O., Davis, C., Menczer, F., and Flammini, A. 2016. "The rise of social bots," *Communications of the ACM*, (59:7), pp. 96–104 (doi: 10.1145/2818717).

Flores-Saviaga, C., Savage, S., and Taraborelli, D. 2016. "LeadWise: Using online bots to recruit and guide expert volunteers," in *Proceedings of the ACM Conference on Computer Supported Cooperative Work, CSCW*, (Vol. 26–Februar), pp. 257–260 (doi: 10.1145/2818052.2869106).

Forelle, M. C., Howard, P. N., Monroy-Hernandez, A., and Savage, S. 2015. "Political Bots and the Manipulation of Public Opinion in Venezuela," *SSRN Electronic Journal*, pp. 1–8 (doi: 10.2139/ssrn.2635800).

Freitas, C. A., Benevenuto, F., Ghosh, S., and Veloso, A. 2015. "Reverse engineering socialbot infiltration strategies in twitter," in *Proceedings of the 2015 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining*, pp. 25–32 (doi: 10.1145/2808797.2809292).

Fuchs, M. 2016. "Warum Social Bots unsere Demokratie gefährden," *Neue Züchricher Zeitung* (available at http://www.nzz.ch/digital/automatisierte-trolle-warum-social-bots-unsere-demokratie-gefaehrden-ld.116166; retrieved November 17, 2016).

Goga, O., Venkatadri, G., and Gummadi, K. P. 2015. "The doppelgänger bot attack: Exploring identity impersonation in online social networks," in *Proceedings of the ACM SIGCOMM Internet Measurement Conference, IMC*, (Vol. 2015–Octob), pp. 141–153 (doi: 10.1145/2815675.2815699).

Guilbeault, D., and Woolley, S. 2016. "How Twitter Bots Are Shaping the Election," *The Atlantic* (available at http://www.theatlantic.com/technology/archive/2016/11/election-bots/506072/; retrieved November 17, 2016).

Haustein, S., Bowman, T. D., Holmberg, K., Tsou, A., Sugimoto, C. R., and 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), pp. 232–238 (doi: 10.1002/asi.23456).

He, Y., Li, Q., Cao, J., Ji, Y., and Guo, D. 2017. "Understanding socialbot behavior on end hosts," *International Journal of Distributed Sensor Networks*, (13:2), p. 155014771769417 (doi: 10.1177/1550147717694170).

Hegelich, S., and Janetzko, D. 2016. "Are social bots on Twitter political actors? Empirical evidence from a Ukrainian social botnet," in *Proceedings of the 10th International Conference on Web and Social Media, ICWSM 2016*, pp. 579–582.

Igawa, R. A., Barbon Jr, S., Paulo, K. C. S., Kido, G. S., Guido, R. C., Proença Júnior, M. L., and da Silva, I. N. 2016. "Account classification in online social networks with {LBCA} and wavelets," *Information Sciences*, (332), pp. 72–83 (doi: http://dx.doi.org/10.1016/j.ins.2015.10.039).

Larsson, A. O., and Moe, H. 2015. "Bots or journalists? News sharing on Twitter," *Communications*, (40:3), pp. 361–370 (doi: 10.1515/commun-2015-0014).

Lobe, A. 2016. "Erzwitschere dir deine Wahl!," *Frankfurter Allgemeine Zeitung* (available at http://www.faz.net/aktuell/feuilleton/debatten/twitter-bots-von-donald-trump-greifen-in-us-wahlkampf-ein-14455894.html; retrieved November 17, 2016).

Lokot, T., and Diakopoulos, N. 2016. "News Bots," *Digital Journalism*, (4:6), pp. 682–699 (doi: 10.1080/21670811.2015.1081822).

Main, W., and Shekokhar, N. 2015. "Twitterati identification system," in *Procedia Computer Science*, (Vol. 45), pp. 32–41 (doi: 10.1016/j.procs.2015.03.076).

Mayzlin, D., Dover, Y., and Chevalier, J. 2014. "Promotional reviews: An empirical investigation of online review manipulation," *American Economic Review*, (104:8), pp. 2421–2455 (doi: 10.1257/aer.104.8.2421).

Moon, B. 2017. "Identifying Bots in the Australian Twittersphere," in *Proceedings of the 8th International Conference on Social Media & Society - #SMSociety17*, New York, NY, USA: ACM (doi: 10.1145/3097286.3097335).

Nagaraja, S., Houmansadr, A., Piyawongwisal, P., Singh, V., Agarwal, P., and Borisov, N. 2011. "Stegobot: A covert social network botnet," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, (Vol. 6958 LNCS), pp. 299–313 (doi: 10.1007/978-3-642-24178-9_21).

Pantic, N., and Husain, M. I. 2015. "Covert botnet command and control using twitter," in *ACM International Conference Proceeding Series*, (Vol. 7–11–Decem), pp. 171–180 (doi: 10.1145/2818000.2818047).

Paradise, A., Puzis, R., and Shabtai, A. 2014. "Anti-reconnaissance tools: Detecting targeted socialbots," *IEEE Internet Computing*, (18:5), pp. 11–19 (doi: 10.1109/MIC.2014.81).

Ratkiewicz, J., Conover, M. D., Meiss, M., Gonçalves, B., Flammini, A., and Menczer, F. 2011. "Detecting and Tracking Political Abuse in Social Media," in *Proceedings of the Fifth International AAAI Conference on Weblogs and Social Media*, Barcelona, Spain, pp. 297–304.

Salto Martínez, R., and Jacques García, F. A. 2012. "Development and implementation of a chat bot in a social network," in *Proceedings of the 9th International Conference on Information Technology, ITNG 2012*, pp. 751–755 (doi: 10.1109/ITNG.2012.147).

Savage, S., Monroy-Hernández, A., and Höllerer, T. 2016. "Botivist: Calling volunteers to action using online bots," in *Proceedings of the ACM Conference on Computer Supported Cooperative Work, CSCW*, (Vol. 27), pp. 813–822 (doi: 10.1145/2818048.2819985).

Sebastian, S., Ayyappan, S., and Vinod, P. 2014. "Framework for design of Graybot in social network," in *Proceedings of the 2014 International Conference on Advances in Computing, Communications and Informatics, ICACCI 2014*, pp. 2331–2336 (doi: 10.1109/ICACCI.2014.6968575).

Stieglitz, S., Brachten, F., Berthelé, D., Schlaus, M., Venetopoulou, C., and Veutgen, D. 2017. "Do Social Bots (Still) Act Different to Humans? - Comparing Metrics of Social Bots with Those of Humans," in *Social Computing and Social Media. Human Behavior: 9th International Conference, SCSM 2017, Held as Part of HCI International 2017, Vancouver, BC, Canada, July 9-14, 2017, Proceedings, Part I*, G. Meiselwitz (ed.), Cham: Springer International Publishing, pp. 379–395 (doi: 10.1007/978-3-319-58559-8_30).

Subrahmanian, V. S., Azaria, A., Durst, S., Kagan, V., Galstyan, A., Lerman, K., Zhu, L., Ferrara, E., Flammini, A., and Menczer, F. 2016. "The DARPA Twitter Bot Challenge," *Computer*, (49:6), pp. 38–46 (doi: 10.1109/MC.2016.183).

Thieltges, A., Schmidt, F., and Hegelich, S. 2016. "The devil's triangle: Ethical considerations on developing bot detection methods.," in *AAAI Spring Symposium - Technical Report*, (Vol. SS-16-01), pp. 253–257.

Tsvetkova, M., García-Gavilanes, R., Floridi, L., and Yasseri, T. 2017. "Even good bots fight: The case of Wikipedia," *Plos One*, (12:2), p. e0171774 (doi: 10.1371/journal.pone.0171774).

Veale, T., Valitutti, A., and Li, G. 2015. *Twitter: The best of bot worlds for automated wit Lecture*

Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), (Vol. 9189) (doi: 10.1007/978-3-319-20804-6_63).

Wang, A. H. 2010. "Detecting spam bots in online social networking sites: A machine learning approach," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, (Vol. 6166 LNCS), pp. 335–342 (doi: 10.1007/978-3-642-13739-6_25).

Wattal, S., Schuff, D., Mandviwalla, M., and Williams, C. B. 2010. "Web 2.0 and Politics: The 2008 US Presidential Election and An E-Politice Research Agenda," *MIS Quarterly*, (34:4), pp. 669–688.

Wilkie, A., Michael, M., and Plummer-Fernandez, M. 2015. "Speculative method and Twitter: Bots, energy and three conceptual characters," *Sociological Review*, (63:1), pp. 79–101 (doi: 10.1111/1467-954X.12168).

Xie, K., and Lee, Y.-J. 2015. "Social media and brand purchase: Quantifying the effects of exposures to earned and owned social media activities in a two-stage decision making model," *Journal of Management Information Systems*, (32:2), pp. 204–238 (doi: 10.1080/07421222.2015.1063297).

Yin, P., Ram, N., Lee, W.-C., Tucker, C., Khandelwal, S., and Salathé, M. 2014. *Two sides of a coin: Separating personal communication and public dissemination accounts in Twitter Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, (Vol. 8443 LNAI) (doi: 10.1007/978-3-319-06608-0_14).

Zhang, J., Zhang, R., Zhang, Y., and Yan, G. 2013. "On the impact of social botnets for spam distribution and digital-influence manipulation," in *2013 IEEE Conference on Communications and Network Security, CNS 2013*, pp. 46–54 (doi: 10.1109/CNS.2013.6682691).

Zhu, Q., Clark, A., Poovendran, R., and Başar, T. 2013. "Deployment and exploitation of deceptive honeybots in social networks," in *Proceedings of the IEEE Conference on Decision and Control*, pp. 212–219 (doi: 10.1109/CDC.2013.6759884).

Acknowledgements: This work was in part supported by the Deutsche Forschungsgemeinschaft (DFG) under grant No. GRK 2167, Research Training Group "User-Centred Social Media".

Paper 2: Defining Bots in an Enterprise Context

Fact sheet

Type	Conference paper (research in progress)
Authors	Stieglitz, Stefan; Brachten, Florian; Kissmer, Tobias
Year	2018
Venue	International Conference on Information Systems (ICIS)
Full citation	Stieglitz, S., Brachten, F. & Kissmer, T. 2018. "Defining Bots in an Enterprise Context," in *Proceedings of the International Conference on Information Systems (ICIS)*.

Defining Bots in an Enterprise Context

Short Paper

Stefan Stieglitz
University of Duisburg-Essen
Duisburg, Germany
stefan.stieglitz@uni-due.de

Florian Brachten
University of Duisburg-Essen
Duisburg, Germany
florian.brachten@uni-due.de

Tobias Kissmer
University of Duisburg-Essen
Duisburg, Germany
tobias.kissmer@uni-due.de

Abstract

Bots, Virtual Assistants and Virtual Agents are well known in a personal environment. Technologies like "Apple Siri" or "Amazon Alexa" serve as digital assistants to enhance both accessibility and productivity. Yet, these technologies have to play with different rules in the enterprise context. Workflows within a digital workplace are different from what consumers are used to in a private surrounding. This concludes into the need for a definition for Enterprise Bots. In this short paper, we aim to provide insights in the current state of literature in this area and derive a definition for Enterprise Bots that is reusable by scholars and practitioners. Furthermore, we propose a research model based on the decomposed Theory of Planned Behavior, that both validates this definition and allows further insights into the topic of Enterprise Bots. Finally, we conclude by testing this model in a survey at a multi-national company and find that 10 of 12 used constructs are highly reliable.

Keywords: enterprise bots, virtual assistants, virtual agents, social bots, digital workplace

Introduction

Social media platforms have become an inherent part of everyday life for most people in the western world. As it is now possible to reach a wide audience beyond one's own acquaintance and as there is money to be made (say as an influencer on Twitter or Instagram), there are now incentives in generating a huge following and extending one's reach. One outlet of these tendencies is the use of automation on social media accounts so that a message can be spread on more channels simultaneously with the aim of pushing the possibility of being heard. So called social bots, automated accounts on social media platforms which post messages that at first glance cannot be distinguished from human messages, are a possible way to increase the reach on social media platforms. Amongst others, mainly due to the possible interference in the outcome of elections (presumably by influencing the voters' opinions) the research on social bots has seen a rise in the recent years (Stieglitz et al. 2017). Research has shown that under certain circumstances those accounts do have the potential to influence the behavior of people (Munger 2017). However, while many of this research concentrated on negative effects of bots, there are also those fields of application in which automated communications can be potential helpful to humans (Stieglitz et al. 2017). There are utilizations such as the *Woebot* which offers a simple form of talk therapy via a web interface. Furthermore, there are whole ecosystems built around automated communication such as *Amazon's Alexa* or *Apple's Siri*. However, the

latter can rather be seen as virtual assistants and in contrast to social bots, are mostly used for 1:1 communication. As the term *assistant* indicates these oftentimes voice controlled applications can be used to assist its users with everyday tasks be it setting a timer or dictating messages which can then be send.

While these examples stem from the private sector, it is easy to see how the overall acceptance of and familiarity with such technologies holds potential for enterprises and slowly migrates into it. Following years of organizational transformation research with the support of information systems (Orlikowski 1996), the Digital Transformation at the workplace bears a high complexity, companies drive large initiatives to move in this direction (Matt et al. 2015). Employees need to get more productive, more efficient and they eventually need to reduce the use of paper within their working environment (Mock 2017). Most companies follow this trend (Dery et al. 2017) which eventually leads to an omnipresent automation in both services and processes as well as in the personal professional environment (Egeli 2016). In order to make this workplace automation accessible for a broad mass of users, natural language interfaces and conversational access are being added to recent software solutions. This leads to the unavoidable need of process knowledge for a user about manifold points of entry in order to satisfy various needs that arise during everyday work life.

Given these developments in both private and enterprise environments, we can see a trend towards the automation in both domains. There have been several approaches of joining both in the past years –most of them being unsuccessful. Probably the most prominent example is *Clippy*, a digital assistant that served as the single-point-of-entry for several processes that could be executed with Microsoft Office Products. While the intention was obvious, it was often described as the "*most hated character*" (Geller 2014) of the Microsoft Office suite.

As bots bring a benefit for private life that eases the little things, this effect might as well be applicable to the enterprise domain. Current work mostly focuses on possible harmful influence either of growing automation in the enterprise context or on social bots in the private domain. Nevertheless, we believe that the assumptions can be validated within an enterprise context and Enterprise Bots (in the following referred to as EB) can support companies to automate processes and provide the users a service that makes them more efficient around the digital workplace, though the literature is rare to non-existing in this regard.

Thus, we structure the short paper as follows. At first, we will present the related work with a focus on the public bot domain. We further seek for literature in relation to bots in an enterprise context and join the outcome of both into the derivation of an initial definition of EB (refer to RQ1 below) that we base our subsequent study on. We believe to be the first to define EB, which is why we initially test our definition in a pre-study with students. In conjunction with the results from both private and corporate domains, we will derive hypothesis that let us assume the comparability of both domains. We further aim to propose a research model, that will let us answer these overall research questions:

RQ1 How can Enterprise Bots be defined?

RQ2 What influences the intention-to-use of Enterprise Bots for users?

RQ3 How strong is the influence of the employees' trust towards Enterprise Bots in the usage intention?

Our model is based on the decomposed Theory of Planned Behavior by Taylor and Todd 1995. We further aim to design a survey that we conduct at a large European manufacturing company with almost 100,000 employees, in order to validate our model and contribute to the findings in the area of EB. We conclude by reporting first reliability scores for our proposed model constructs and the outlook on further steps of our research.

Theoretical Background and Related Work

A World of Difference: Public and Enterprise Bots

While the foundation for human computer interaction in a natural way is artificial intelligence (AI), a term which stems from Alan Turing in the 1950s (Turing 1950), the topic garnered more widespread attention amongst others, due to supercomputers like IBM Watson and in the end user area by Siri, the virtual assistant (VA) which Apple introduced in 2011 with the iPhone 4s (Guzman 2017). Through this and similar

assistants of competitors which have been preinstalled on smartphones since then, such programs penetrated into everyday life and made people familiar with their use. Another point of contact with (seemingly) intelligent machines which has received more attention especially in the recent past are social bots (Stieglitz et al. 2017). This term describes artificial social media accounts that try to hide their artificial nature by behaving as human-like as possible in order to be taken fully by other social media users. Social Bots are postulated to try to bring certain issues into the public perception and to simulate broader support for these issues than is actually the case (Stieglitz et al. 2017).

Although similarities exist, social bots should not be confused with virtual assistants. As Guzman points out, "*both [vocal social agent]s and socialbots are forms of AI and share a technological lineage. While agents can function across platforms, socialbots are autonomous programs operating in social media networks*" (2017, p.70). This example shows that for technologies that communicate with humans in a human-like-manner the terms used to describe those phenomena are used interchangeably even though different aspects are described. Accordingly, Strohmann et al. point out, that e.g. "there are various synonyms for VAs [...] [and s]imilar to the variety of names there is also no single definition, but several definition approaches [...] tasks and characteristics." (2018, p.3). For example, while Quarteroni sees both, chatbots and virtual agents, as subclass of *conversational agents* (2018), Io and Lee equate chatbots and conversational agents (2018). Klopfenstein et al. call Siri and the likes *public voice-enabled Virtual Private Assistant* in the field of conversational interfaces (2017), Io and Lee call Siri *chatbot* (2018) and Siri itself states that it is a *personal assistant* (Guzman 2017). To obviate confusion in the course of this paper, we will use the term virtual assistant (VA) when referring to software programs, that can be addressed via voice or text and that can respond to the users input (i.e. *assist*) with sought after information or carry out certain actions (like playing a certain song or generating a reminder, cf. Strohmann et al., 2018, Quateroni, 2018). When talking about social bots, we refer to autonomous bot accounts operating on *social* media platforms (hence the term 'social', cf. Stieglitz, Brachten, et al. 2017, Guzman 2017). Nevertheless, the debate over social bots and the associated research interest in the phenomenon has brought insights that could apply to VAs. In particular, the potential for influencing people is of interest to researchers, as it is an integral part of most of the underlying effects of social bots. Accordingly, some publications use the *Computer are Social Actors Framework*, which - based a series of experiments from the 1990s - postulates that humans perceive computers as social entities, even when it is known that the other person is a computer (Nass and Moon 2000). People in the experiments behaved to computers the same way as other people would behave in social situations. Based on these results, recent studies have shown that people perceive social bots in a similar way to other people (Edwards et al. 2014). Furthermore, it could be shown that under certain circumstances social bots were able to change people's behavior (Munger 2017). While VAs in general do not aim to change the behavior of their users (and thus exercise power / authority over them) the results can't be carried over to virtual assistants in their entirety – instead, with VAs the aspect of "serving" and subordination plays a greater role (Guzman 2017). Still, the findings described indicate important aspects of the relationship between humans and machines, showing that interactions are not per se considered 'artificial' and suggest that there is some acceptance in dealing with VA.

These findings refer to private, non-work related aspects of VAs. Nevertheless, the topic is also of interest to enterprises. Besides AI and machine learning which are their own fields of research and offer great potential in the analysis and optimization of business processes (or open completely new business fields), the use of VAs is of interest to companies as well. On the one hand, there is the aspect of external communication with customers, where e.g. the workload of call centers can be reduced by using VAs to intercept certain customer requests or users can be specifically addressed without having to employ additional staff (McTear et al. 2016). In addition, VAs also have potential for use within a company to also relieve service centers or enable employees to work more efficiently, as natural language requests can be quickly made and processed without the need for additional overhead, which has limited presence and capacity for processing such requests. Still, while there are potentials and uses for such VAs, research in enterprise contexts has rarely been concerned with them. A thorough literature search with various terms (*"enterprise bot", "enterprise social bot", "virtual enterprise assistant"* or *"enterprise chatbot")* hardly yielded results – if any. The papers that were at least partially fitting mostly researched said applications in a healthcare context or as a means to get in contact with potential customers but not to primarily reduce the workload within the company. As the latter is an essential part of discussions around organizational transformation (Besson and Rowe 2012), agility and flexibility have to be fostered. These dogmas prevail in modern organizations and lead to initiatives towards digital workplaces (Dery et al. 2017; Köffer 2015). As

productivity increase and user adoption are hot topics within the digital transformation, concepts like digital nudging become attractive for corporations (Kissmer et al. 2018). At the same time, companies need to manage a heterogenous workforce that has different needs based on their demographic profile (Meske et al. 2016). But given the impact of social bots or VA in private life (e.g. *Apple Siri*), this familiarity and prior findings could be carried over to the enterprise area (Quarteroni 2018). It becomes obvious, that in order to support digital transformation initiatives within companies, concepts like EB have to be looked into – which to our knowledge hasn't happened so far. As we did not find existing literature on the subject of EB and because of the confusion on the terminology surrounding VAs, we do see the need to delimit the virtual assistants used in the private sector (like Amazons Alexa or Apples Siri) from similar automated systems in an enterprise context. In the following we thus establish the definition for Enterprise Bots as a simpler, clearer term with the simplicity and novelty possibly being useful when conducting surveys (as in our case) and to present a label for the technology to the test subjects.

Defining Enterprise Bots

Based on the considerations mentioned above, we summarize the benefits and purpose of social bots, organizational transformation and the tendencies towards digital workplaces in a definition. It is essential to be able to fall back on a uniform terminology in the future as well as to have a common understanding for the study. Hence, we define Virtual Assistants which are applied in an internal enterprise context, labeled Enterprise Bots, as follows:

An Enterprise Bot is an automated user service that provides casual and conversational interactions with complex enterprise systems and processes. Enterprise Bots can e.g. answer questions or perform smaller tasks. A user can interact with it by just typing or speaking a request in natural language. It can only act in a passive way, meaning it must actively be triggered by the user.

This definition has been positively pre-tested with a group of 13 students. With this, we intend to ensure the quality and a common understanding before launching a larger study within the case company.

Hypothesis, Research Model and Study Approach

As we found in the related work, bots are commonly used in private life – either actively (e.g. *Amazon Alexa* or *Apple Siri*) or passively (e.g. social bots on twitter). We further found, that current trends (digital transformation, digital workplace) lead towards new concepts, such as digital nudging. Amongst others these concepts drive a company to an increased productivity. Thus, EB are potentially helpful to fulfill these requirements, as they are currently placed to increase productivity in a personal environment. Since this is a major goal of the digitization as mentioned, EB are assumed to be fostering these goals. In order to answer the research questions, we aim to propose a model, that helps to understand both the applicability of bots in an enterprise context and the applicability of the definition provided by us in the previous chapter. Based on this model, we derive hypotheses and indicate first meta-results of our survey towards data integrity and descriptive means.

As previously shown, there seems to be a degree of acceptance for the interaction with machines (Nass and Moon 2000). However, as the findings stem mostly from the private sector in which people are normally not paid by a third party to fulfill their tasks in an efficient manner, other requirements apply for the enterprise use of virtual entities. Therefore, before companies make investments to implement any type of systems, it is important to evaluate the acceptance and willingness of employees to use these systems. A common model for studying the acceptance of technology has been the unified theory of acceptance and use of technology (UTAUT), which is based on the theory of planned behavior (TPB). As per Taylor and Todd (1995), several extensions of the TPB exist, such as the UTAUT. Furthermore, one of the most reliable extension of the TPB is the Decomposed Theory of Planned Behavior (dTPB), which is particularly suitable in the case of innovative subjects. It is based on constructs from the innovation literature and provides a more complete understanding of the actual usage intention. It includes constructs towards the mutual superior and peer influence (Taylor and Todd 1995). As we aim to contribute to the understanding of Enterprise Bots and thus focus on the professional perception by users, we are especially interested in the users' perceived ability and the individual's beliefs. By focusing on specific beliefs, the model enables us to derive assumptions towards adoption and usage, adding managerial relevance. Hence, the decomposition of the well-established TPB constructs directly benefits our study. As the dTPB has been proven to be highly

reliable and tested several times (Sadaf et al. 2012; Shih and Fang 2004) we found it particularly suitable for our study.

Our research model is illustrated below in figure 1. All scale relations except *Trust* are based on the dTPB (Taylor and Todd 1995). To fit our research needs in the enterprise context and to address corresponding matters, we removed the variable *Compatibility* and added *Trust* based on Reid and Levy (2008) as an already well-tested construct relation. As *compatibility* referred to the actual system compatibility and since integration and compatibility is a natural requirement towards bots, we are not expecting a major benefit and a rather low contribution from the inclusion of this variable. On the other hand, *Trust* was of particular interest, since EB are fulfilling smaller tasks autonomously for the user which requires the user to trust the bots work. Since, as mentioned, *Trust* was non-existing in the original dTPB, we took the construct from the study of Reid and Levy (2008) who extended the Theory of Planned Behaviour towards the construct of Trust and positively tested the overall reliability and item integrity. This way, we are able to keep the model integrity towards the influence of *Trust* on both *Perceived usefulness* and *Perceived ease-of-use*.

We further added moderators for a more extensive insight. For one we chose to add a set of demographic questions, since demographic attributes of employees can result in different behaviors (Meske et al. 2016). Furthermore we included the personal innovativeness based on Agarwal and Prasad (1998) and the technology acceptance scale by Neyer et al. (2012), indicated as "System usage behavior" in figure 1. As EB are an innovative technology that can change or at least support the daily workflow, we want to see the impact of a respondent's personal attitude towards technology and whether this impacts the usage intention for EB or not. We further removed the latent variable for peer and superior influence in order to make the direct influence onto the usage intention of each more visible. Since we can't measure the actual usage, we keep usage intention as the terminating latent variable in our model and do not include the original "Usage Behavior" variable. An overview of all scales and adjustments is shown in table 1.

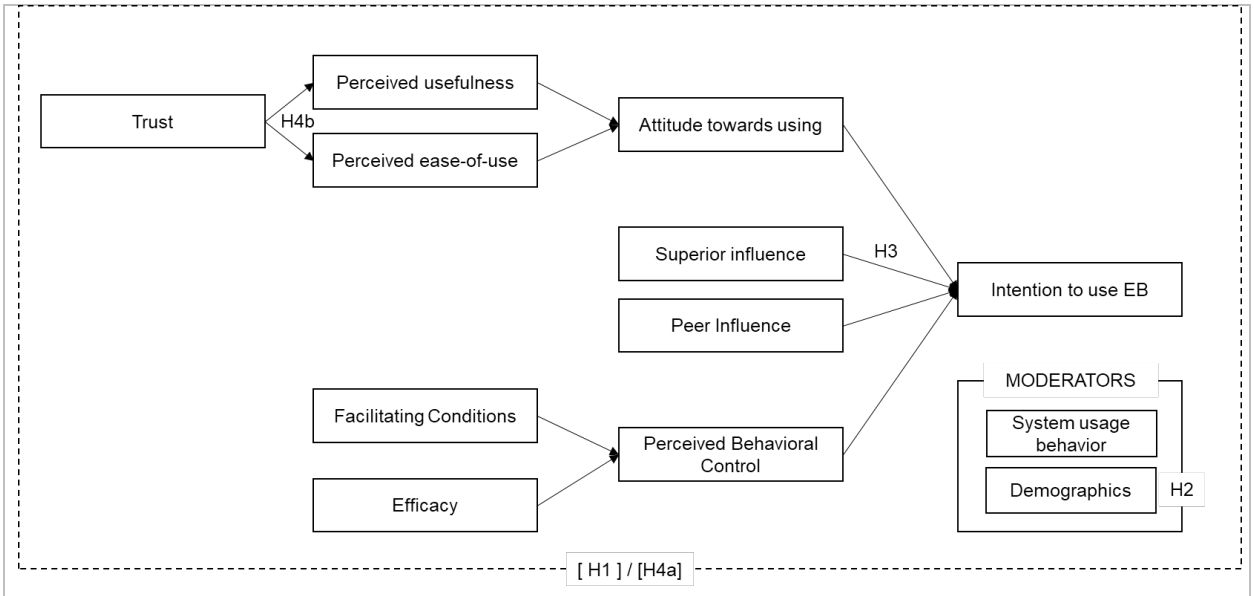

Figure 1. Research Model

Given this research model, we propose hypotheses based on the related work that was outlined earlier and with the aim to answer our research questions. We are planning to test these against the survey results once they are available. Based on prior findings we expect all constructs to positively correlate with the corresponding latent variable. We base this assumptions on prior validations of the dTPB model (e.g. Sadaf et al. 2012) which indicated for example a higher usage attitude if the perceived usefulness is high as well. A lower score in trust will on the other hand result in a lower attitude towards using and an increased perceived usefulness will lead to a higher attitude to use EB. This will help us to validate both the model and the definition. Thus, we hypothesize:

H1: *All model correlations are positively flagged.*

Disapproving this hypothesis would indicate that either the individual's intention is not reasonable or our definition was not understood properly.

Based on several studies , we expect the demographic influence to be significant, as we measure a rather digital technology (cf. Meske et al. 2016, 2018). We thus hypothesize:

H2: *Latent variables are significantly influenced by demographic attributes of employees.*

In contrast to e.g. social bots that exist in private life, Enterprise Bots are used in a professional context, which differs significantly from the private context (e.g. due to different security measures, stricter hierarchies, etc.). Given a natural hierarchic influence in companies (Stieglitz et al. 2014), we expect the superior influence to be higher than the peer influence, allowing us to derive adoption strategies. Due to organizational hierarchies that prevail in almost all large companies, we expect both, *Superior* and *Peer Influence* to be positively correlated, reflecting the hierarchical and peer influence (ref. to H1 as well).

H3: *The latent variable "Superior Influence" has a greater influence on usage intention than "Peer Influence"*

As EBs in our definition are referred to as simplifying instruments that are made to facilitate work, we hypothesize a positive attitude of the respondents towards the latent variables. The positive attitude implicates an above median mean for these constructs. Nevertheless as skepticism is rather high in novel, digital working environments (Köffer 2015), we postulate that the *Trust* in EB initially is low, resulting in a lower mean.

H4a: *The mean for all connected constructs (except Trust) is higher than the median.*

H4b: *The mean for Trust is lower than the median.*

To test our definition and model, we conducted an online-survey at a large German headquartered multi-national manufacturing company with over 160 subsidiaries in 50 countries globally. The company currently does not run any kind of application that would fit our definition of EB. Thus, we needed to establish a common understanding of the concept which we did by providing respondents with our definition of Enterprise Bots. The survey was structured according to our research model and the scales (ref. table 1) were taken from literature related to the dTPB. The preliminary findings should provide insights in the users' intention towards using Enterprise Bots. The survey was sent out on March 5th 2018 to 183 employees at a Romanian location of the company, generating 64 responses in total over the course of two weeks. In the next section we present preliminary findings based on the responses with the focus on data and construct reliability.

Construct Reliability and Implications

Among the 64 responses, 1 had to be excluded for missing data generating our final dataset of n=63. All scales used were captured on a 7-point Likert scale with 1 indicating a rejection of the statement and 7 an agreement. Table 1 outlines the constructs together with the scale(s) that were used to measure the constructs as well as Cronbach's Alpha indicating the reliability of answers within our dataset. Further, the mean values are reported for all constructs.

Construct	**Scale**	**Cronbach's Alpha α**	**Mean**
Attitude towards using	Shih and Fang 2004; Taylor and Todd 1995	.960	4.633
Perceived usefulness	Taylor and Todd 1995	.414	4.675
Perceived ease-of-use	Taylor and Todd 1995	.532	3.228
Trust	Reid and Levy 2008	.948	4.409
Perceived behavioral control	Taylor and Todd 1995	.918	4.850

Facilitating Conditions	Lu et al. 2005; Taylor and Todd 1995	.823	4.903
Efficacy	Taylor and Todd 1995	.836	4.300
Superior influence	Taylor and Todd 1995	.928	4.733
Peer influence	Taylor and Todd 1995	.969	4.533
System usage behavior[1]:	(-)	(-)	(-)
Personal Innovativeness	Agarwal and Prasad 1998	.834	4.295
Technology Acceptance	Neyer et al. 2012	.965	4.835
Intention to use EB	Nicolaou and Mcknight 2006; Taylor and Todd 1995	.963	4.580

Table 1. Construct reliability with Cronbach's Alpha

Summarizing the *Personal Innovativeness* and the *Technology Acceptance*, we calculated an α of .930 and a mean of 4.565. Almost all constructs yielded an excellent reliability with values of α between 0.8 and 0.9, indicating that most of our research model is highly reliably and hence consistent. *Perceived usefulness* and the *Perceived ease-of-use* were not reliable and did not show specifically flawed items that would increase the reliability when left out. As mentioned previously, EB were not present in the company at the time when the survey was run. We assume that this impacts the reliability of the constructs. Both scales are based on Taylor and Todd 1995 and were not altered. This is rather unexpected and unusual compared to findings from prior studies (Reid and Levy 2008; Sadaf et al. 2012). As an anonymous reviewer suggested, this might be due to an insufficient understanding of our definition by the test subjects as, like we mentioned before, the company we tested in does not run any application similar to what we examined as enterprise bots. We intent to include a deeper analysis in our complete research by conducting interviews to analyze the root cause and propose actions accordingly. We observed a mean above the median (4) for almost all values, indicating a slight positive mindset towards EB by the users' which is also reflected by the mean of 4.580 for the intention to use EB.

The latter findings indicate that hypotheses **H4a** and **H4b** have to be rejected, as for **H4a** the *Perceived ease-of-use* mean is below the median and for **H4b** the mean of 4.409 for *Trust* is above the median. The latter finding indicates that trust towards EB is (pending further analysis of the data) not perceived as negative. However, these results need further statistical testing (e.g. variance and value deployment).

Conclusion, Contribution and Further Research

In this short paper we introduced the term of Enterprise Bots (EB). By analyzing the current state of literature, we showed the lack of research on bots within the enterprise domain and furthermore pointed out key motivators for the existence of bots in the public environment. We introduced a first definition for EB and outlined the research model that is the base for a quantitative survey we conducted. Preliminary findings based on answers by 63 subjects indicate that the chosen scales are highly reliable in most cases. Just two constructs were showing a reliability below an acceptable threshold of 0.7, which will be followed up on in the full paper of this study.

However, limitations apply. The survey (and thus the existing scales) had to be customized to fit the future tense and hence match the non-existing system availability (EBs don't exist at the moment in our case company). This is causing a slight deviation of words for few items but does not impact the overall reliability of the model, as all questions were altered (this approach was used by Shih and Fang (2004) as well and has proven to be valid). As mentioned earlier one important aspect that needs deeper investigation are the low reliability scores for *Perceived usefulness* and *Perceived ease-of-use*, two of the most frequently used constructs in usage intention research. As one anonymous reviewer proposed that the lack of understanding

[1] As *System usage behavior* is not a latent variable it was not included in our model and is and table 1 doesn't show corresponding values.

of definition of EBs by the test subjects may be a reason (possibly caused by the non-existence of EBs in the company). We plan to follow up this suggestion by conducting interviews with test subjects in the full paper and possibly discuss adjustments to the definition, for example to put more emphasis on the business context. As a last limitation we see that rather young respondents returned the survey (remark: due to space limitations in this short paper, we did not include a table with a demographic overview) and hence there is a possibility to have a demographic based bias which has not been captured. We will examine this in the following study on this subject and evaluate the significance of the influence of age and other moderating factors towards the answers.

All in all our study contributes in a unique way to the current state of research around bots and virtual assistants. Neither did we find a fitting definition nor an enterprise-oriented usage intention survey. Companies currently examining these technologies and service providers offer solutions accordingly (e.g. Microsoft Cortana) that are still in a rather immature state. Although this field is proceeding fast, science has yet to catch up. We aim to partially contribute to closing this research gap with our study by providing a novel definition for EB and outlining the difference between public bots and EB. The previously stated additional research questions and hypotheses will be answered in a complete research paper, building upon this short paper.

We conclude by encouraging scholars to provide feedback on literature, definition and the used research model. With this research we aim to provide the basis for further, more extensive analysis in the area of EB, such as the need of a user adoption for EB or the actual usage compared to the usage intention we derived.

References

Agarwal, R., and Prasad, J. 1998. "A Conceptual and Operational Definition of Personal Innovativeness in the Domain of Information Technology," *Information Systems Research* (9:2), pp. 204–215. (https://doi.org/10.1287/isre.9.2.204).

Besson, P., and Rowe, F. 2012. "Strategizing Information Systems-Enabled Organizational Transformation: A Transdisciplinary Review and New Directions," *Journal of Strategic Information Systems* (21), pp. 103–124. (https://doi.org/10.1016/j.jsis.2012.05.001).

Dery, K., Sebastian, I. M., and van der Meulen, N. 2017. "The Digital Workplace Is Key to Digital Innovation," *MIS Quarterly Executive* (16:2), pp. 135–152. (http://www.misqe.org/ojs2/index.php/misqe/article/viewFile/767/460).

Edwards, C., Edwards, A., Spence, P. R., and Shelton, A. 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), Pergamon, pp. 372–376. (https://doi.org/10.1016/J.CHB.2013.08.013).

Egeli, M. 2016. "Existierende Maturity Models Für Digital- Und Mobile Business," in *Erfolgsfaktoren von Mobile Business*, Wiesbaden: Springer Fachmedien Wiesbaden, pp. 27–41. (https://doi.org/10.1007/978-3-658-12771-8_4).

Geller, T. 2014. "How Do You Feel?," *Communications of the ACM* (57:1), pp. 24–26. (https://doi.org/10.1145/2555809).

Guzman, A. L. 2017. "Making AI Safe for Humans: A Conversation With Siri," in *Socialbots and Their Friends - Digital Media and the Automation of Sociality*, R. W. Gehl and B. Maria (eds.), New York: Routledge, pp. 69–85.

Io, H. N., and Lee, C. B. 2018. "Chatbots and Conversational Agents: A Bibliometric Analysis," *IEEE International Conference on Industrial Engineering and Engineering Management* (2017–Decem), pp. 215–219. (https://doi.org/10.1109/IEEM.2017.8289883).

Kissmer, T., Potthoff, T., and Stieglitz, S. 2018. "Enterprise Digital Nudging: Between Adoption Gain and Unintended Rejection," in *AMCIS 2018 Proceedings*, New Orleans, LA, August 16. (https://aisel.aisnet.org/amcis2018/AdoptionDiff/Presentations/12).

Klopfenstein, L. C., Delpriori, S., Malatini, S., and Bogliolo, A. 2017. "The Rise of Bots: A Survey of Conversational Interfaces, Patterns, and Paradigms.," *Proceedings of the 2017 Conference on Designing Interactive Systems*, pp. 555–565. (https://doi.org/10.1145/3064663.3064672).

Köffer, S. 2015. "Designing the Digital Workplace of the Future – What Scholars Recommend to Practitioners," *Icis*, pp. 1–21.

Matt, C., Hess, T., and Benlian, A. 2015. "Digital Transformation Strategies," *Business and Information*

Systems Engineering, pp. 339–343. (https://doi.org/10.1007/s12599-015-0401-5).

McTear, M., Callejas Carrion, Z., and Griol Barres, D. 2016. *The Conversational Interface : Talking to Smart Devices*.

Meske, C., Kissmer, T., and Stieglitz, S. 2016. "Influence of Demography and Management Responsibility on the Adaptation of Technologies for the Support of Organizational Ad Hoc Communication," in *Multikonferenz Wirtschaftsinformatik, MKWI 2016* (Vol. 1).

Meske, C., Kissmer, T., and Stieglitz, S. 2018. "Global Adoption of Unified Communication Technologies as Part of Digital Transformation in Organizations," in *Multikonferenz Wirtschaftsinformatik, MKWI 2018*.

Munger, K. 2017. "Tweetment Effects on the Tweeted: Experimentally Reducing Racist Harassment," *Political Behavior* (39:3), Springer US, pp. 629–649. (https://doi.org/10.1007/s11109-016-9373-5).

Nass, C., and Moon, Y. 2000. "Machines and Mindlessness: Social Responses to Computers," *Journal of Social Issues* (56:1), Wiley/Blackwell (10.1111), pp. 81–103. (https://doi.org/10.1111/0022-4537.00153).

Neyer, F. J., Felber, J., and Gebhardt, C. 2012. "Entwicklung Und Validierung Einer Kurzskala Zur Erfassung von Technikbereitschaft," *Diagnostica* (58:2), pp. 87–99. (https://doi.org/10.1026/0012-1924/a000067).

Orlikowski, W. J. 1996. "Improvising Organizational Transformation over Time: A Situated Change Perspective," *Information Systems Research*, pp. 63–92. (https://doi.org/10.1287/isre.7.1.63).

Quarteroni, S. 2018. "Natural Language Processing for Industry: ELCA's Experience," *Informatik-Spektrum* (41:2), pp. 105–112. (https://doi.org/10.1007/s00287-018-1094-1).

Reid, M., and Levy, Y. 2008. "Integrating Trust and Computer Self-Efficacy with TAM: An Empirical Assessment of Customers' Acceptance of Banking Information Systems (BIS) in Jamaica," *Journal of Internet Banking and Commerce* (12:3), pp. 1–18. (https://doi.org/10.1007/978-3-531-92534-9_12).

Sadaf, A., Newby, T. J., and Ertmer, P. A. 2012. "Exploring Factors That Predict Preservice Teachers' Intentions to Use Web 2.0 Technologies Using Decomposed Theory of Planned Behavior," *Journal of Research on Technology in Education* (45:2), pp. 171–196. (https://doi.org/10.1080/15391523.2012.10782602).

Shih, Y.-Y., and Fang, K. 2004. "The Use of a Decomposed Theory of Planned Behavior to Study Internet Banking in Taiwan," *Internet Research* (14:3), pp. 213–223. (https://doi.org/10.1108/10662240410542643).

Stieglitz, S., Brachten, F., Ross, B., and Jung, A.-K. 2017. "Do Social Bots Dream of Electric Sheep? A Categorisation of Social Media Bot Accounts," *Proceedings of the 28th Australasian Conference on Information Systems - Paper 206*, pp. 1–11. (http://arxiv.org/abs/1710.04044).

Stieglitz, S., Riemer, K., and Meske, C. 2014. "Hierarchy or Activity? The Role of Formal and Informal Influence in Eliciting Responses From," *Proceedings of the European Conference on Information Systems (ECIS) 2014, Tel Aviv, Israel*, pp. 1–14.

Strohmann, T., Fischer, S., Siemon, D., Brachten, F., Lattemann, C., Robra-Bissantz, S., and Stieglitz, S. 2018. "Virtual Moderation Assistance: Creating Design Guidelines for Virtual Assistants Supporting Creative Workshops," in *Proceedings of the 22nd Pacific Asia Conference on Information Systems - Paper*, pp. 3580–3594.

Taylor, S., and Todd, P. A. 1995. "Understanding Information Technology Usage: A Test of Competing Models," *Information Systems Research* (6:2), pp. 144–176. (https://doi.org/10.1287/isre.6.2.144).

Turing, A. M. (1950). Computing Machinery and Intelligence. In D. C. Ince (Ed.), Mechanical Intelligence (pp. 133–160). New York, NY: North-Holland.

Paper 4: Strategies and Influence of Social Bots in a 2017 German state election

Fact sheet

Type	Conference paper
Authors	Brachten, Florian; Stieglitz, Stefan; Hofeditz, Lennart; Kloppenborg, Katharina; Reimann, Annette
Year	2017
Venue	Australasian Conference on Information Systems (ACIS)
Full citation	Brachten, F., Stieglitz, S., Hofeditz, L., Kloppenborg, K. & Reimann, A. (2017). Strategies and Influence of Social Bots in a 2017 German state election. In: *Proceedings of the 28th Australasian Conference on Information Systems (ACIS)*, Paper 207.

Strategies and Influence of Social Bots in a 2017 German state election – A case study on Twitter

Florian Brachten
University of Duisburg-Essen
Duisburg, Germany
Email: florian.brachten@uni-due.de

Stefan Stieglitz
University of Duisburg-Essen
Duisburg, Germany
Email: stefan.stieglitz@uni-due.de

Lennart Hofeditz
University of Duisburg-Essen
Duisburg, Germany
Email: lennart.hofeditz@stud.uni-due.de

Katharina Kloppenborg
University of Duisburg-Essen
Duisburg, Germany
Email: katharina.kloppenborg@stud.uni-due.de

Annette Reimann
University of Duisburg-Essen
Duisburg, Germany
Email: annette.reimann@stud.uni-due.de

Abstract

As social media has permeated large parts of the population it simultaneously has become a way to reach many people e.g. with political messages. One way to efficiently reach those people is the application of automated computer programs that aim to simulate human behaviour - so called social bots. These bots are thought to be able to potentially influence users' opinion about a topic. To gain insight in the use of these bots in the run-up to the German Bundestag elections, we collected a dataset from Twitter consisting of tweets regarding a German state election in May 2017. The strategies and influence of social bots were analysed based on relevant features and network visualization. 61 social bots were identified. Possibly due to the concentration on German language as well as the elections regionality, identified bots showed no signs of collective political strategies and low to none influence. Implications are discussed.

Keywords: political bots; social bots; social media; Twitter; state election

1 Introduction

Social media have gained importance in political communication over the last years. People as well as political actors use social media such as Twitter to debate political topics or to conduct political online campaigns (Yang et al. 2016). Twitters retweet system combined with its public nature strongly adds to the diffusion of information (Stieglitz and Dang-Xuan 2012). However, social media like Twitter also attract people who aim to abuse their functionalities and apply their potential as an efficient way to spread messages to a large audience with little effort (Rinke 2016). One potential danger in this context is that users could attempt to manipulate public opinion or to disrupt political communication. Within social networks such as Twitter, an effective tool for accomplishing this feat is the use of so called social bots (Woolley 2016).

Social bots are automated social media accounts designed to mimic human behavior (Abokhodair et al. 2015; Ferrara et al. 2016; Freitas et al. 2015). Through the simulation of human behaviour they are at first glance not easily recognizable as artificial accounts (Ferrara et al. 2016). This in turn could lead to human users misjudging the importance of the messages spread by such accounts eventually leading to being influenced in favour of the messages at display. The accounts differ on their level of sophistication with low-level-accounts, merely aggregating information from websites and using it to produce simple messages, e.g. on Twitter. A more sophisticated social bot on the other hand can be conversational and aim at passing as a human (Abokhodair et al. 2015).

The application of such accounts has been observed in several political contexts such as the Brexit debate in 2016 or the US presidential election in 2016 where social bots were responsible for roughly one-fifth of the conversation on Twitter (Howard and Kollanyi 2016). They potentially influenced users' opinion about the election as one candidate seemed to have more support than the other (Bessi and Ferrara 2016; Kollanyi et al. 2016). Kindled by observations of the use of these accounts in important votes, a debate considering the use of such accounts in the state election in 2017 in the most populous German state of North-Rhine Westphalia (NRW) has been a topic in the media and politics of the country (Rinke 2016). Driven by the ongoing debate about potential dangers of social bots and by the statement of the right wing populist party *Alternative für Deutschland* (*Alternative for Germany - AfD*) to potentially use social bots, all other major parties officially refrained from using social bots during their campaigns ("Das sagen die NRW-Parteien zu Social Bots" 2016). Accordingly, social bots in support of the right-winged *AfD* have been identified on Facebook by the popular media (Bender and Oppong 2017).

Besides the detection of social bots themselves, another important part in research is the detection and identification of different strategies social bots utilize while the attempt to take part in a discussion (Abokhodair et al. 2015). This is especially true for political context where social bots are thought to manipulate users' perceptions of political actors or parties (Bessi and Ferrara 2016; Woolley 2016). Furthermore, as social bots are a rather new phenomenon, the literature on social bot usage in German politics or votes is sparse. Therefore, one goal of this paper is to deepen the understanding of social bot use in political contexts and more specifically in Germany on the example of the NRW state election in May 2017. To do so, social bots on Twitter were detected and analyzed concerning their strategies and the influence they exerted on human users. Twitter was chosen for comparability with previrous findings, as the majority of research on the topic of social bots takes place on Twitter (Stieglitz et al. 2017).

Following this introduction section 2 will present state of the art social bot detection techniques and findings on strategies and influence of social bots which will conclude in the research questions. Section 3 presents the dataset and methods used. Findings are presented in section 4 and discussed in section 5.

2 Theoretical Background

Among the research on social bots one of the largest streams are papers with the focus on the use of social bots in politics. The application of these bots could be shown in several different political scenarios such as Venezuelan politics in general (Forelle et al. 2015), the Syrian civil war (Abokhodair et al. 2015), US-American mid-term elections (Mustafaraj and Metaxas 2010) or the conflict between Ukraine and Russia in 2015 (Hegelich and Janetzko 2016). Furthermore, in two of the most far reaching votes in recent years, namely the 2016 UK vote to leave the European Union as well as the 2016 US presidential election, the application of bots was also detected (Bessi and Ferrara 2016; Howard and Kollanyi 2016). Especially as the outcome of the latter two events was contrary to what the polls had indicated in the forerun, the debate in the aftermath in parts also considered social media

and within those social bots to have played a role in the outcome. If follows, that with the election of a new parliament in Germany in the fall of 2017 some concerns are raised as to what extent social bots might be used on this occasion as well. However, research about the use of social bots in German politics is rare. In a data memo, Neudert et al. (2017) analyzed the use of social bots during the German Federal Presidency Election in February 2017. The overall impact of social bots was rather small as the authors found only 22 automated accounts with a total of about 6,000 tweets out of more than 36,500 accounts with about 120,000 tweets. However, as the German Federal President is not elected directly by the people this election may have been of less interest for social bot use. As North-Rhine Westphalia is Germanys most populous federal state (with 17.9 million citizens) elections here are considered a pre-election test for overall Germany. Regarding the election of a new German parliament in fall of 2017, this paper aims at expanding the findings on social bot usage in German politics on the example of the 2017 German federal election in NRW. In order to compare the bots' behavior the first step is to assess the strategies applied by bots in earlier studies.

2.1 Social Bot Strategies in Politics

Woolley (2016) conducted a qualitative content analysis of English news articles on social bots. The author found that governments and other actors deployed social bots during elections, country-specific political conversations or crises and for preemptive online security purposes. The strategies that social bots used varied depending on country and political entity. One strategy of social bots was to disseminate pro-government or pro-candidate tweets to sway public opinion in their favour. An article of the New York Times described this strategy in a state prosecutors' allegations in South Korea, where over one million tweets were sent from social bots. Politicians also bought social bots to boost their follower lists. In the US presidential election in 2012 Mitt Romney acquired nearly 117,000 followers in about 24 hours through presumptive social bot activities (Woolley 2016). Another finding was that communication surrounding elections was targeted with so called *Twitter bombs*, i.e. the flooding of a hashtag used by opposing parties through a social botnet and thus barraging that hashtag. Those practices were labeled *smokescreening* or *misdirecting* by Abokhodair et al. (2015) who observed similar practices in the context of the Syrian civil war. Spam tactics were also used for attacking various other countries and commercial organizations. Another term often used in the context is the so called *astroturfing* - tweeting or retweeting a (political) opinion to suggest broad consensus around this opinion. In other words, astroturfing is a term for the simulation of an artificial grassroot movement.

Abokhodair et al. (2015) used a content analysis to identify the strategies used by social bots in the Syrian botnet. They searched for content that did not match the hashtags used to detect smoke screening or misdirecting. Chu et al. (2012) considered features like the URL frequency as an efficient detection feature for social bots by analyzing 500,000 accounts. They found that social bots tend to include URLs in tweets to redirect users to external content and argued that this is due to Twitter's character limit of 140 characters.

According to Zhang et al. (2013a), a good indicator for astroturfing is if there are many automated tweets for a topic or a hashtag. Bessi and Ferrara (2016) observed how social bots affect political discussion in the context of the 2016 US presidential election. Through analyzing the sentiments of tweets that used hashtags related to the presidential election the authors found social bots to produce more positive content in support of a candidate than humans do. This strategy resembles astroturfing, as it tries to generate the impression that a certain opinion is more strongly represented than another.

A study by Ratkiewicz et al. (2011) researched how Twitter can be exploited through astroturfing campaigns. To answer this, they analyzed Twitter data that they obtained from the astroturfing detection tool *Truthy*. They found that astroturfing campaigns had to be detected early, because after passing Twitter's spam detector they can do great harm in a political campaign.

To sum it up, social bots can be implemented with different objectives in mind. In political contexts, they are applied to boost the number of followers, influence public opinion or disrupt communication (Woolley 2016). These objectives are strongly linked to different strategies. To create a forged grass root support for a specific agenda social bots tweet or retweet certain opinions (astroturfing) (Abokhodair et al. 2015; Zhang, Zhang, et al. 2013). They disrupt communication by flooding hashtags with unrelated but similar content (smoke screening) or by guiding users' attention away from a topic (misdirection) (Abokhodair et al. 2015). However, these social bot strategies are still quite new and therefore there is further need for research on the identification of strategies in political context. As mentioned before, research lacks findings considering social bots in the political context in Germany and as such also considering the strategies applied. To close this gap, our first research question is:

RQ1: *Were there bot activities in the context of the state election 2017 of NRW on Twitter and which strategies were used by social bots?*

Besides the question of strategies social bots apply another important aspect to answer is to what extend these accounts can influences public opinion. As mentioned before, through so called astroturfing bot accounts are thought to artificially generate the impression that a certain opinion has stronger support than it really does. While this and other similar mechanisms may seem retraceable in in theory it is important to look at the actual findings in influence of social bots.

2.2 Influence of Social Bots

Social media can be seen as an important tool for advancing democratic discussions, as it is a very powerful communication medium (Bessi and Ferrara 2016) and facilitates information diffusion (Stieglitz and Dang-Xuan 2012). Thus, political messages on Twitter can potentially be spread quickly and reach a high number of users. Nevertheless, determining the influence of social media interaction on political decision making and behavior is hard to assess and current research offers mixed results. On one side, according to a study by Kushin and Yamamoto (2010), social media only played a minor role in affecting situational political involvement and political self-efficacy. On the other side, different studies show an influence of social media use on offline political participation and political expression (Gil de Zúñiga et al. 2014). But what happens when social bot strategies, like astroturfing and smoke screening, are used in order to manipulate the content in social media?

Considering social bots, the first step to determine their impact is to detect the reach of these accounts. Bessi and Ferrara (2016) explored the level of network embeddedness of social bots in the US presidential election in 2016 to ascertain the reach of the manipulations. They found that humans retweeted other humans and social bots equally as often indicating that social bots can be just as effective in spreading information as human's accounts are. The authors concluded that the presence of social bots has the potential to negatively affect democratic political discussion and thus to manipulate public opinion and endanger the integrity of elections. Human users retweeting social bots indicate difficulty for normal users to differentiate between social bot and human user which is in accordance with findings by Everett et al. (2016). The authors confronted Internet users with genuine comments generated by humans as well as automatically generated comments. In 50% of the cases the users could not differentiate whether a comment came from another user or was generated automatically. Similarly, Freitas et al. (2015) found an akin number of obtained followers between social bots that solely retweeted other users tweets compared with bots who retweeted and additionally produced automatic tweets. The authors concluded that this is due to users not being able to identify automatically generated tweets, possibly as a result of messages on Twitter being short and often informal or grammatically incoherent which makes the identification of automated content difficult for human users.

The effectiveness of social botnets for digital influence manipulation was investigated by Zhang et al. (2013b). The authors built their own social botnet for spam distribution on Twitter and tried to boost the influence of the network. To measure the digital influence, they used scores by three commercial digital-influence services (Klout, Kred, and Retweet Rank) that were based on how frequently accounts are retweeted and mentioned in a certain amount of days. In their experiments Zhang et al. (2013b) concluded that the three presented measurement methods of digital influence could easily be manipulated by social bots. The manipulation of the measurement by social bots is possible, because reach plays an important role and is often very high for social bot accounts.

These examples show that social bots in general are able to trick other users into acting as if the automated accounts were human. This might be the reason for human users retweeting social bots equally often as human users, thus spreading their content and in turn leading to people potentially being influenced by social bots (Bessi and Ferrara 2016).

In this paper, the digital influence is measured mainly through the reach of social bots. According to Bessi and Ferrara (2016) the digital influence in political contexts played an important role in the US presidential election 2016. In European countries like Germany research in this field is still sparse. Thus, the second research question is:

RQ2: *What influence did social bots have in the NRW state election 2017 campaigns on Twitter?*

The answers to these two questions will show 1. if social bots were used and 2. if so what influence they had in a German state election. Findings will add to the body of current research on the topic of social bots and will expand it with knowledge on the use in German politics and 2. broaden knowledge about

the actual importance of social bots. The following section describes the research design used to answer these questions.

3 Research Design

To answer the research questions, we captured and analysed a dataset of Tweets concerning the NRW state election in May 2017.

3.1 Dataset

The dataset was collected from the 23rd of April to the 21st of May 2017 via the Twitter SEARCH API and covers a month of tweets about the NRW state election 2017 which took place on the 14th of May. The following general keywords marking tweets about the state election were used for the selection of tweets: *Landtagswahl* (the German expression for state election), *Landtagswahl NRW*, *#landtagswahl*, *#landtagswahlnrw*, *#ltwnrw* and *#ltwnrw17*. Furthermore, mentions of and tweets by the official accounts of parties in the state parliament of NRW at that time were tracked. Additionally, the account of the right wing populist party AfD was tracked as well, as there were indicators of social bots connected to the AfD (e.g. Bender and Oppong 2017). Thus *@gruenenrw*, *@CDUNRW_de*, *@fdp_nrw*, *@nrwspd*, *@PiratenNRW* and *@AlternativeNRW* were included in the dataset.

As the aforementioned keywords included the general term *#landtagswahl* and as a state election in another German federal state (namely Schleswig-Holstein) also took place in May 2017, the initial dataset included some tweets concerning the latter election which had to be removed. Hence, tweets including the keyword *ltwsh* or mentions of the official Schleswig-Holstein accounts of the local parties (*@Gruene_SH*, *@SH_CDU*, *@FDP_SH*, *@spdsh*, *@piratenparteish*, *@AfD_LV_SH*) were dismissed when they did not contain the additional keywords *NRW* or *Nordrhein*. This led to 6,166 tweets being excluded with a total of 182,995 tweets remaining. Out of the remaining tweets, 125,468 were retweets and 57,527 were original tweets. A total of 33,481 unique accounts were analyzed. Further descriptive statistics are presented in table 1.

Keywords	No. of tweets	Keywords	No. of tweets
Landtagswahl	18,595	@gruenenrw	13,370
Landtagswahl NRW	1,274	@CDUNRW_de	7,863
#landtagswahl	3,748	@fdp_nrw	10,045
#landtagswahlnrw	139	@nrwspd	6,541
#ltwnrw	76,673	@PiratenNRW	17,512
#ltwnrw17	48,538	@AlternativeNRW	17,447

Table 1. Descriptive statistics of the dataset. Number of tweets includes original tweets and retweets.

3.2 Methods

For the detection of social bots, BotOrNot and Python version 3.6.1 in combination with R version 3.4.0 were used. BotOrNot is one of the first social bot detection interface for Twitter and employs supervised machine learning classification algorithms trained with 15,000 examples for social bot accounts and 16,000 examples for human accounts and more than 5 million tweets in total to assess whether an account is likely to be a bot - or not (Davis et al. 2016). In the past papers used this system to classify accounts in their gathered sample (Ferrara et al. 2016). The systems offers an open source API with a possible Python interface. As BotOrNot cannot detect accounts exhibiting a mixture of human and social bot characteristic the system should be combined with other approaches (Ferrara et al. 2016). Thus, BotOrBot was only the first part in the detection process described in the following section.

3.2.1 Detection of Social Bots in the Sample

The schematic description of the identification process is shown in figure 1. As a first step BotOrNot was utilized to detect social bots (I) as it has been tested on a large amount of data and generates reliable results for simple to medium sophisticated social bots (Davis et al. 2016). BotOrNot offers a public API endpoint which was accessed using Python. The detection program generates a score ranging from 0-1 with a higher score indicating stronger social bot-like behavior. For an account to be

declared a social bot a score of at least 0.7 was decided on which is consistent with prior studies (e.g. Bessi & Ferrara 2016). A second check (II) was then conducted by analyzing six features that proved to be useful in detecting social bots (verified accounts, Tweet frequency, follower-to-followee-ratio, account language, account creation time, and profile picture – cf. Alarifi et al. 2016; Chu et al. 2012; Ferrara et al. 2016; Neudert et al. 2017). If an account displayed social bot-like behavior for a feature, the account scored 1 for this feature otherwise it was assigned 0. A summation of the feature scores results in a score ranging from 0 to 6 for each account. A higher score indicates a higher probability of the account being a social bot. Accounts with a score of at least 5 were classified as social bots.

To make sure that the classification was successful two manual checks were conducted by two coders each. The first manual check (III) concerned the accounts that were classified as social bots in step I and II. The goal here was to judge if the accounts were correctly classified as social bots. Additionally, 50 random accounts from the dataset were examined manually (IV) to judge if the automated detection may have missed social bot accounts. Accounts that were already examined in the first manual check were not included in the second manual check. Accounts that were classified as humans by both coders were given a 0, accounts that were classified as social bots by both coders were given a 1 and otherwise the account was given a 2. This results in a score of 0-2. Only those accounts that were consistently classified as social bots by the coders (score = 1) were declared as social bots – irrespective of their BotOrNot and feature analysis scores. Criteria for the manual checks were: Was the content of tweets original, intelligent and *human-like*, i.e. did the user report what he/she was doing or feeling (Chu et al. 2012)? Did tweets contain irony, sarcasm or jokes (Abokhodair et al. 2015; Ferrara et al. 2016)? Were there references to friends, family members, etc.? Was a tweet part of a conversation between friends (Abokhodair et al. 2015)? Was the profile information individualized (Freitas et al. 2015)? Did the account follow suspicious accounts? Did the account have suspicious followers (Chu et al. 2012; Ferrara et al. 2016)?

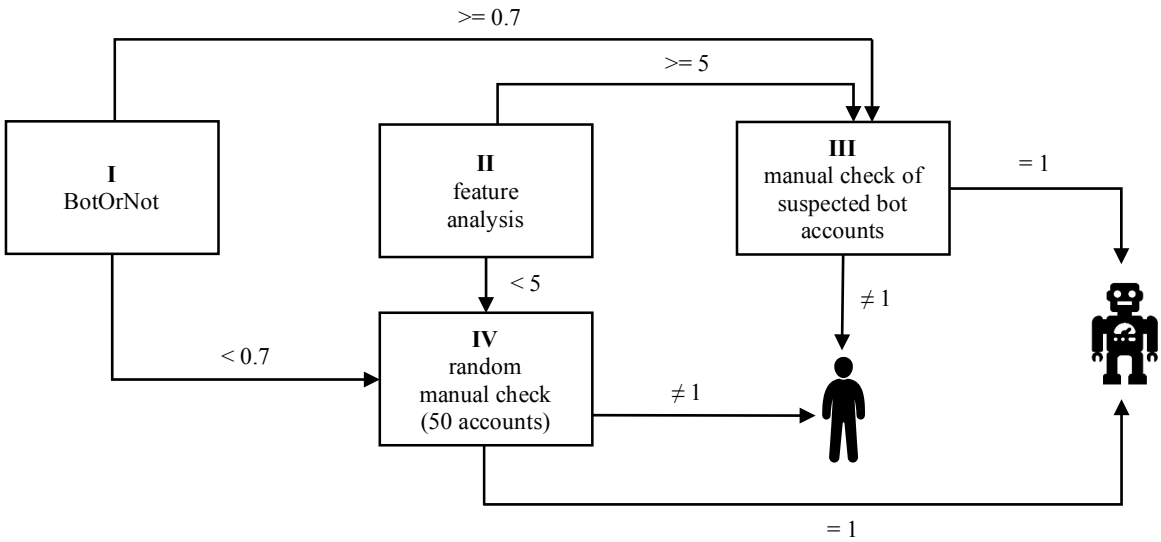

Figure 1. Social bot detection procedure. Arabic numbers reflect the social bot detection scores described in section 3.2.1. The scores determined which path was taken. Roman numerals show the order in which the steps were conducted.

Cohen's kappa adjusted by Byrt et al. (1993) was calculated to assess the consistency between the two coders. This adjusted version diminishes effects caused by prevalence, in this case unequal appearances of social bots and humans in the subsets. Both manual checks resulted in a satisfying inter-rater reliability (k = 0.93 for the manual check of social bots and k = 0.91 for the random check).

Through this procedure, a total of 61 social bot accounts could be identified which made up around 0.2% of all 33,481 accounts in the dataset.

3.2.2 Methods for Strategy Analysis

After detecting social bots their strategies were analyzed. First, a check for smoke screening and misdirecting was conducted. The two strategies were combined as they are not completely

distinguishable from each other. For a tweet to be classified as smoke screening / misdirection the tweet content had to be unrelated to the hashtag used but still related to the topic of the state election or politics in general (Abokhodair et al. 2015). If the content was completely unrelated to the hashtag and had no connection to politic topics the tweet was classified as spam. The content of the tweets was checked manually by one coder. Additionally, the URL frequency of social bot accounts was calculated, i.e. the total number of URLs used by social bots was divided by the total number of tweets by social bots (Chu et al. 2012).

Furthermore, the tweets were checked for astroturfing, i.e. conveying support of a certain opinion or party. One characteristic of astroturfing is a high number of automated tweets concerning a certain topic (Zhang et al. 2013a) Therefore, the most used hashtags and keywords by social bots as well as by human users were determined to identify relevant topics in these groups. To analyze if topics were artificially pushed by social bots a ratio was calculated by dividing the number of times a popular hashtag was used by the number of tweets. This was done for human users and social bots individually to gain a metric that allowed comparison of hashtag frequency between humans and social bots. The same procedure was applied to popular keywords.

Automated tweets that could not be identified as smoke screening / misdirecting, spam or astroturfing were analyzed for similarities in order to uncover other strategies.

3.2.3 Methods for Influence Analysis

Finally, the influence of social bots was evaluated by assessing their reach. A retweet rank was calculated to uncover the most influential accounts. The rank was calculated sorting the accounts by the number of times their original tweets were retweeted by other users. To determine if human users retweeted social bots, it was distinguished whether those users were other social bots or human users.

4 Results

4.1 Behaviour of Detected Social Bots

The 61 identified social bot accounts tweeted 336 times in total with 199 messages being original tweets and 137 being retweets. In comparison, human users in the dataset tweeted 182,659 times during the data collection period. Of these, 121,664 were original tweets and 60,995 retweets. On average, social bots posted M = 5.60 (SD = 14.00, range = 1 - 96) tweets during the collection period, while human users *posted* M = 5.47 (SD = 19.41, range = 1 – 680). The values of the standard deviation and range indicate that the tweet frequency for social bots as well as for human users is skewed in that few users produced the most content. More precisely, the three most active social bot accounts tweeted 96, 40, and 35 times each in the dataset with *Rambonelli*, an AfD supporting social bot, being the most active. The most active human accounts tweeted 680, 609, and 553 times each and were accounts associated with the pirate party. The account *AlternativeNRW* ranks on fifth place while accounts of other parties were being less active.

4.2 Detected Strategies

No indication for the occurrence of smoke screening or misdirecting could be found as hashtags or URLs were either related to the content or completely unrelated. Those tweets containing URLs and hashtags unrelated to the topic were classified as spam. This strategy was discovered 24 times in the social bot dataset. However, the tweet frequency was too low to speak of an extensive strategy. URLs that were related to the content often led to news articles of various official and unofficial media sites. This behavior, together with the aforementioned spamming, is also reflected in a high URL frequency of 0.95 URLs per bot tweet.

The low overall occurrence of social bots in the dataset also led to the absence of astroturfing as this strategy constitutes of large numbers of tweets in favour of a certain topic. Said behavior could not be observed. The top five hashtags used by social bots were *#ltwnrw*, *#afd*, *#ltwnrw17*, *#nrw* and *#nachrichten*. The first four of these were also in the top five hashtags used by human accounts, but in a different order. On average humans shared the most used hashtag in 40% of their tweets. The same cannot be said for social bots. Here, the top hashtags were only used in 10% of the tweets each (tables 2, 3 and 4). The top keywords used by social bots were *Landtagswahl*, *NRW*, *Nordrhein-Westfalen*, *SPD* and *CDU* (the both biggest parties, competing for the lead in the vote). *Landtagswahl* was used in 42% of the tweets on average but concerns the state election in general rather than being connected to a certain party. As for *SPD* and *CDU* those keywords were used rather seldom. Additionally, humans used very similar top keywords (tables 5 and 6).

	Human accounts	Social Bots
Number of Tweets	182,659	336
Number of Hashtags	396,796	597

Table 2: Tweet statistics for human and social bot accounts

Top 5 Hashtags H	Count	Ratio
#ltwnrw	71,428	0.39
#ltwnrw17	46,858	0.26
#nrw	21,378	0.12
#afd	21,104	0.12
#piraten	7,815	0.04

Table 3: Top hashtags used in human tweets.

Top 5 Hashtags SB	Count	Ratio
#ltwnrw	59	0.11
#afd	53	0.10
#ltwnrw17	50	0.09
#nrw	38	0.07
#nachrichten	18	0.03

Table 4: Top hashtags used in social bot tweets.

Nonetheless, some social bot accounts posted party-specific content. One example for this behavior was exhibited by the account *Rambonelli* which had a follower count of 1,081 a status count of 34,982 statuses (96 in this dataset) and a friend count of 1,024. The social bot tweeted or retweeted exclusively content that was related to topics of the AfD like "*RT @MdLFernandes: +++ Die neue Fahne der Arbeiter ist blau ! +++ Bei der Landtagswahl haben die meisten Arbeiter die #AfD gewÃ¤hlt.... httpsâ€¦*"[sic!] (meaning '*The new colour of workers is blue! Most blue collar workers voted for the #AfD in the state election*').

Top 5 Keywords	Count	Ratio
NRW	13,977	0.08
Landtagswahl	12,399	0.07
Wählen	8,660	0.05
CDU	5,631	0.03
SPD	5,514	0.03

Table 5: Top keywords used in human tweets.

Top 5 Keywords	Count	Ratio
Landtagswahl	142	0.42
NRW	94	0.28
Nordrhein-Westfalen	45	0.13
SPD	32	0.10
CDU	32	0.10

Table 6: Top keywords used in social bot tweets.

4.3 Detected Influence

For the detection of the influence of social bots in our dataset a retweet rank was calculated. From all detected social bot accounts only three produced original tweets were retweeted: Those were *afdwtal* on rank 904 with one original tweet and six retweets, all of them by human accounts, *pilgerpetrus* on rank 1,550 with one original tweet and three retweets, all of them by human accounts and *Immoprofil_JM* on rank 1,753, also with one original tweet, two retweets, all of them by social bot accounts. The most retweeted accounts in the whole dataset were *PiratenNRW*, *WDR* and *tagesschau* (the two latter both being news outlets) with 1,240, 1,172 and 1,070 retweets respectively.

5 Discussion

As the results show we only found very little social bot activity in our study (RQ1). Out of 33,481 unique accounts only 61 accounts were classified as social bots. Furthermore, social bot accounts on average posted only 5.6 tweets during the whole collection period. These results differ from research concerning the US presidential election where "political bot activity reached an all-time high" (Kollanyi et al. 2016, p4.) with 15% of the studied accounts being classified as social bots and the discovered social bots made up a significant part of the conversation (Bessi and Ferrara 2016). Concerning the first research question which strategies were used by social bots in the NRW state election 2017 campaigns on Twitter (RQ1) it can be stated that no collective strategies could be detected. While there were some social bot accounts that shared URLs to news articles the referenced websites were diverse and mostly concerned the state election in general instead of pushing news related to a specific party. Therefore, the underlying objective of these social bots does not seem concerned with manipulating communication about the NRW state election by shedding a positive or negative light on specific political parties or actors. In addition to the sharing of news articles indicators for spam were found for individual social bots. These social bots used hashtags concerning the state election and added content

and / or URLs that were unrelated to the topic. As the content and URLs were completely unrelated to the state election and politics in general these tweets were not classified as smoke screening / misdirecting. Thus, social bots did not try to influence communication by hiding content associated with a certain hashtag behind unrelated political content. They also did not try to direct users' attention to other political topics as was the case for the Syrian botnet analyzed by Abokhodair et al. (2015). Both the sharing of news and spamming are reflected in the URL frequency which indicates that every social bot generated tweet on average contained an URL.

Another strategy, that was expected to be used by social bots during the NRW state election, was astroturfing. As indicated by the low hashtag and keyword ratios, social bots did not try to push a topic collectively. Additionally, more than half of the top 5 hashtags and keywords were related to the state election in general which does not indicate the creation of artificial support for a certain party. There were, however, individual accounts sending sophisticated propaganda which Woolley (2016) described as a characteristic of social bots in political contexts. In the NRW state election dataset just a few social bots used this strategy namely *Rambonelli* or *afdwtal*. Both accounts are in support of the AfD that was accused of using social bots and artificially increasing their follower number and reach (Bender and Oppong 2017; Reuter 2017). However, we did not find a network of social bots that used this strategy. One possible explanation for the weak activities is that the use of social bots is in a test phase and should not arouse much attention. Furthermore, the language as well as the regional nature of the state election could be a reason for the apparent low application of social bots. Most other studies that could show a higher degree of bot activity concentrated on large events in English-speaking countries (e.g. the Brexit vote or the US election). However, literature also presents examples for application in non-English speaking contexts (Abokhodair et al. 2015) as well as smaller scale political events (US mid-term elections - Mustafaraj and Metaxas 2010). Nevertheless, both of the latter examples either used English language or were in the context of larger political events (for example the Syrian civil war). This could mean that if neither of these criteria is at hand (as in the current example), the application of social bots may not be seen as an efficient tool as the costs for a potential influence could be too high. One possible way to test this assumption is to analyze the social media communication surrounding the election of a new parliament for whole Germany and then compare these findings to our current study. With the latter election being a larger political event, comparable to the vote for a new government in other countries, the usage of social bots could be seen as being more worthwhile than on the regional level.

In accordance with the findings on strategies used by social bots in the NRW state election, regarding the influence of social bots in the NRW state election 2017 campaigns (RQ2) it can be derived that social bots exerted almost no influence. Only three of the social bot accounts were retweeted at all and their highest retweet rank was 904 with six retweets for *afdwtal*. Nevertheless, these retweets as well as those from *pilgerpetrus* were all carried out by human accounts which means that there was some human-social bot interaction, but to a very small extent. *Immoprofil_JM* was not retweeted by human accounts but twice by other social bots. However, this small number can rather be evaluated as coincidence than evidence for a social botnet. This is supported by the social network analysis that was conducted based on Bessi and Ferrara (2016) who explored the level of network embeddedness to measure the reach of accounts. It turned out that in the current sample social bots were largely outside the main graph which means that they were hardly networked in this dataset. Most of the connections seemed to be outside the graph. In this respect, the results differ greatly from those of Bessi and Ferrara (2016), who found many social bots within their social graph of the US presidential elections 2016.

As with all research some limitations to our findings apply. First, one important part of every study on social bots is the identification of bots in a dataset. While we attempted to apply a method which was at the same time efficient to use as well as effective it is possible that among the accounts classified as social bots there are some human users and the other way around. Also, while we relied on an external tool which promises to use over 1,000 different variables in the process of identification, it is still possible that BotOrNot is not able to detect German bots as it may be with other languages. Here, further research is needed, if maybe bots are sophisticated in a way that they differ depending of the country they are targeted at. The other important aspect to consider is that we explicitly only concentrated on German accounts, not regarding English tweets that may have also occurred within our dataset and may have enlarged the body of social bots and thus their influence. However, a closer look into the data did not indicate that this was the case.

In summary, only few social bots were present in the Twitter communication about the NRW state election 2017. These social bots did not pursue collective strategies and were hardly networked with one another or with humans.

6 Conclusion

The present study investigated the use of social bots, their strategies and influence in political contexts on the example of the NRW state election 2017 on Twitter. Contrary to findings from different countries and elections this study could reveal only a small number of accounts showing enough signs of automation that they could be classified as social bots. The only strategies that could be identified were traditional spamming without a political concept and the sharing of diverse news articles concerning the state election. In general, social bots were only loosely connected to the network and to each other and no social botnets were identified. This is already indicative for a low reach of social bots in the dataset and is further supported by social bots being seldom retweeted. These results deviating from other research in political contexts could be due to the language used (German) as well as the regional nature of the state election. Future studies thus could a) validate these findings on different elections on a similar level and b) compare these findings to other political events in Germany of the same size as well as larger votes such as the elections to the Bundestag.

Consequently, according to this study there are almost no social bots, let alone political bots used in the NRW state election campaigns 2017 on Twitter. One reason for these findings the German language that makes it harder for international social bot developers to create social bot campaigns for political events in Germany.

Even though the results indicated a very small use of social bots with no collective strategies and low reach more research on social bot use in Germany is desirable. As argued before the NRW state election might simply not have been of enough importance for social bots to be deployed. One interesting use case is the upcoming German federal election 2017 as it is the biggest political election in Germany. At this event, the practice using social bots on Twitter or other social media like Facebook should be analyzed. Especially Facebook should be considered due to the popularity of it in Germany.

7 References

Abokhodair, N., Yoo, D., and McDonald, D. W. 2015. "Dissecting a Social Botnet," in *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing - CSCW '15*, pp. 839–851 (doi: 10.1145/2675133.2675208).

Alarifi, A., Alsaleh, M., and Al-Salman, A. 2016. "Twitter turing test: Identifying social machines," *Information Sciences*, (372), pp. 332–346 (doi: 10.1016/j.ins.2016.08.036).

Bender, J., and Oppong, M. 2017. "Frauke Petry und die Bots," *Frankfurter Allgemeine Zeitung Online* (available at http://www.faz.net/aktuell/politik/digitaler-wahlkampf-frauke-petry-und-die-bots-14863763.html; retrieved July 31, 2017).

Bessi, A., and Ferrara, E. 2016. "Social bots distort the 2016 U.S. Presedential election online discussion," *First Monday*, (21:11), pp. 1–14.

Byrt, T., Bishop, J., and Carlin, J. B. 1993. "Bias, prevalence and kappa," *Journal of Clinical Epidemiology*, (46:5), pp. 423–429 (doi: 10.1016/0895-4356(93)90018-V).

Chu, Z., Gianvecchio, S., Wang, H., and Jajodia, S. 2012. "Detecting automation of Twitter accounts: Are you a human, bot, or cyborg?," in *IEEE Transactions on Dependable and Secure Computing*, (Vol. 9), pp. 811–824 (doi: 10.1109/TDSC.2012.75).

"Das sagen die NRW-Parteien zu Social Bots.," 2016. *Westdeutscher Rundfunk* (available at http://www1.wdr.de/nachrichten/landespolitik/parteien-nrw-social-bots-100.html; retrieved July 31, 2017).

Davis, C. A., Varol, O., Ferrara, E., Flammini, A., and Menczer, F. 2016. "BotOrNot: A System to Evaluate Social Bots," *Proceedings of the 25th International Conference Companion on World Wide Web*, (: 1602.009), pp. 273–274 (doi: 10.1145/2818717).

Everett, R. M., Nurse, J. R. C., and Erola, A. 2016. "The anatomy of online deception: What makes automated text convincing?," in *Proceedings of the ACM Symposium on Applied Computing*, (Vol. 04–08–Apri), pp. 1115–1120 (doi: 10.1145/2851613.2851813).

Ferrara, E., Varol, O., Davis, C., Menczer, F., and Flammini, A. 2016. "The rise of social bots," *Communications of the ACM*, (59:7), pp. 96–104 (doi: 10.1145/2818717).

Forelle, M. C., Howard, P. N., Monroy-Hernandez, A., and Savage, S. 2015. "Political Bots and the Manipulation of Public Opinion in Venezuela," *SSRN Electronic Journal*, pp. 1–8 (doi:

10.2139/ssrn.2635800).

Freitas, C. A., Benevenuto, F., Ghosh, S., and Veloso, A. 2015. "Reverse engineering socialbot infiltration strategies in twitter," in *Proceedings of the 2015 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining*, pp. 25–32 (doi: 10.1145/2808797.2809292).

Gil de Zúñiga, H., Molyneux, L., and Zheng, P. 2014. "Social Media, Political Expression, and Political Participation: Panel Analysis of Lagged and Concurrent Relationships," *Journal of Communication*, (64:4), Wiley Subscription Services, Inc., pp. 612–634 (doi: 10.1111/jcom.12103).

Hegelich, S., and Janetzko, D. 2016. "Are social bots on Twitter political actors? Empirical evidence from a Ukrainian social botnet," in *Proceedings of the Tenth International Conference on Weblogs and Social Media (ICWSM-2016)*, Palo Alto, CA: The AAAI Press, pp. 579–582.

Howard, P. N., and Kollanyi, B. 2016. "Bots, #StrongerIn, and #Brexit: Computational Propaganda during the UK-EU Referendum," (available at http://arxiv.org/abs/1606.06356).

Kollanyi, B., Howard, P. N., and Woolley, S. C. 2016. "Bots and automation over twitter during the second U.S. presidential debate.," *Comprop data memo*, (2).

Kushin, M. J., and Yamamoto, M. 2010. "Did Social Media Really Matter? College Students' Use of Online Media and Political Decision Making in the 2008 Election," *Mass Communication and Society*, (13:5), pp. 608–630 (doi: 10.1080/15205436.2010.516863).

Mustafaraj, E., and Metaxas, P. 2010. "From obscurity to prominence in minutes: Political speech and real-time search," *WebSci10: Extending the Frontiers of Society On-Line*, p. 317 (available at http://repository.wellesley.edu/computersciencefaculty/9/).

Neudert, L.-M., Kollanyi, B., and Howard, P. N. 2017. "Junk News and Bots during the German Federal Presidency Election: What Were German Voters Sharing Over Twitter?," Oxford (available at http://comprop.oii.ox.ac.uk/wp-content/uploads/sites/89/2017/03/What-Were-German-Voters-Sharing-Over-Twitter-v6-1.pdf).

Ratkiewicz, J., Conover, M. D., Meiss, M., Gonc, B., Flammini, A., and Menczer, F. 2011. "Detecting and Tracking Political Abuse in Social Media," *Icwsm*, pp. 297–304 (doi: 10.1145/1963192.1963301).

Reuter, M. 2017. "Größter AfD-Twitter-Account ist ein Scheinriese," *Der Tagesspiegel* (available at http://www.tagesspiegel.de/medien/datenjournalismus/twitter-datenanalyse-groesster-afd-twitter-account-ist-ein-scheinriese/19691492.html; retrieved July 31, 2017).

Rinke, A. 2016. "Hintergrund - Sorge um gekaufte digitale Hilfe im Wahlkampf 2017," (available at http://de.reuters.com/article/deutschland-parteien-sozialemedien-idDEKCN12L0EJ; retrieved July 31, 2017).

Stieglitz, S., Brachten, F., Ross, B., and Jung, A.-K. 2017. "Do Social Bots Dream of Electric Sheep? A Categorisation of Social Media Bot Accounts," *Proceedings of the Australasian Conference on Information Systems*.

Stieglitz, S., and Dang-Xuan, L. 2012. "Political communication and influence through microblogging - An empirical analysis of sentiment in Twitter messages and retweet behavior," *Proceedings of the Annual Hawaii International Conference on System Sciences*, pp. 3500–3509 (doi: 10.1109/HICSS.2012.476).

Woolley, S. C. 2016. "Automating power: Social bot interference in global politics," *First Monday*, (21:4) (doi: 10.5210/fm.v21i4.6161).

Yang, X., Chen, B.-C., Maity, M., and Ferrara, E. 2016. "Social Politics: Agenda Setting and Political Communication on Social Media," in *Social Informatics: 8th International Conference, SocInfo 2016, Bellevue, WA, USA, November 11-14, 2016, Proceedings, Part I*, E. Spiro and Y.-Y. Ahn (eds.), Cham: Springer International Publishing, pp. 330–344 (doi: 10.1007/978-3-319-47880-7_20).

Zhang, J., Carpenter, D., and Ko, M. 2013. "Online astroturfing: A theoretical perspective," *19th Americas Conference on Information Systems, AMCIS 2013*, (4:1), pp. 2559–2565 (available at http://www.scopus.com/inward/record.url?eid=2-s2.0-84893308397&partnerID=40&md5=f6d9be2d64336eb6f42f53f56d4e253b).

Zhang, J., Zhang, R., Zhang, Y., and Yan, G. 2013. "On the impact of social botnets for spam distribution and digital-influence manipulation," in *2013 IEEE Conference on Communications and Network Security, CNS 2013*, pp. 46–54 (doi: 10.1109/CNS.2013.6682691).

Paper 5: Social Bots in a Commercial Context - A Case Study on SoundCloud

Fact sheet

Type	Conference paper (research in progress)
Authors	Ross, Björn; Brachten, Florian; Stieglitz, Stefan; Wikström, Patrik; Moon, Brenda; Münch, Felix Victor; Bruns, Axel
Year	2018
Venue	European Conference on Information Systems (ECIS)
Full citation	Ross, B., Brachten, F., Stieglitz, S., Wikström, P., Moon, B., Münch, F., and Bruns, A. 2018. "Social Bots in a Commercial Context – A Case Study on Soundcloud," in *Proceedings of the European Conference on Information Systems (ECIS)*.

SOCIAL BOTS IN A COMMERCIAL CONTEXT – A CASE STUDY ON SOUNDCLOUD

Research in Progress

Ross, Björn, University of Duisburg-Essen, Duisburg, Germany, bjoern.ross@uni-due.de

Brachten, Florian, University of Duisburg-Essen, Duisburg, Germany, florian.brachten@uni-due.de

Stieglitz, Stefan, University of Duisburg-Essen, Duisburg, Germany, stefan.stieglitz@uni-due.de

Wikström, Patrik, Queensland University of Technology, Brisbane, Australia, patrik.wikstrom@qut.edu.au

Moon, Brenda, Queensland University of Technology, Brisbane, Australia, brenda.moon@qut.edu.au

Münch, Felix Victor, Queensland University of Technology, Brisbane, Australia, felixvictor.muench@hdr.qut.edu.au

Bruns, Axel, Queensland University of Technology, Brisbane, Australia, a.bruns@qut.edu.au

Abstract

Recently, automated communication on social media has seen increased attention. Social bots, social media accounts controlled by algorithms that mimic human behaviour, have been found to attempt to influence users in several political contexts. However, their use in a commercial context, e.g. to boost sales of a product by aggressively promoting it with thousands of messages, has so far been neglected. To address this shortcoming, this paper examines the case of the social media music platform SoundCloud. We gathered a dataset of six months of activity, comprising 15,850,069 tracks and 12,125,095 comments. We then calculated a comment uniqueness score for highly active accounts to assess the variability of their comments. First analyses show that some accounts post suspiciously repetitive comments. These accounts also frequently repost existing content, but contribute little original content. An analysis of the commenting network further underlines that these accounts differ clearly from regular users. We conclude that the comment uniqueness metric can be used as an indicator to distinguish bots from humans, and that a considerable proportion of SoundCloud comments are likely to emanate from bots or semi-automated accounts. The implications of these findings and future plans are discussed.

Keywords: social bots, SoundCloud, social media, automated communication, social network analysis.

1 Introduction

Social media platforms have become an important source for customers to inform themselves about products and services they are interested in (Gu, Park, and Konana, 2012). Much research has shown that product information which is shared by other customers has a larger effect on buying decisions

than promotions and advertisements initiated by the selling companies (Zhou and Duan, 2016). As a consequence, there is a growing interest in influencing customer reviews and the visibility of products and services on social media platforms. While those are crucial to the success of companies (especially if they generate the majority of revenue on the internet), findings from research on automated communication on social media indicate that a part of the seemingly human-generated content is actually produced by software algorithms, called social bots (Varol et al., 2017). Findings from the field of politics show that bots apparently try to influence the opinion climate on certain topics (Hegelich and Janetzko, 2016). Simultaneously, research on whether social bots try to influence customers of companies is sparse. Given their apparent pervasiveness in political contexts, it seems plausible that they are also being used for commercial purposes, but the extent of their usage and their specific aims are unclear. To address this gap, this paper examines the application of bots in a commercial context using the example of the music industry and more specifically the social network SoundCloud (see also Bruns et al., 2018).
This short paper presents preliminary results on the following research questions:

1. How can bots be differentiated from humans on the music sharing platform SoundCloud?
2. How do these bots behave, especially how do they interact with tracks and human users?

To address these questions, it is first necessary to develop a bot identification method for this platform. We present a simple metric that separates the accounts into two distinct groups and we then apply it on a data set spanning six months of activity from SoundCloud. The analysis of this data reveals that the two groups of accounts also differ in several other behavioural attributes. One of the groups closely resembles what is known about bots, underlining the validity of the proposed identification method. In a network analysis, we also find that the bot-like accounts focus on tracks that were uploaded by specific accounts, suggesting that they might be used to boost the visibility of these accounts. This research is part of a larger study with the overall goal of examining the relevancy of these bots to the music industry, that is, to determine whether or not they are successful in influencing consumers' purchasing and streaming behaviour.
The paper is structured as follows. Section 2 presents the necessary background information on social bots and the music industry. Section 3 describes the data set and introduces the comment uniqueness metric which we use to separate user types. Section 4 describes our preliminary results applying this metric to the data set, and presents a network analysis. We discuss the implications of this finding and plans for further research in Section 5. Section 6 contains concluding remarks.

2 Background

2.1 Social bots

Social bots have seen a rise in attention in recent years, especially on Twitter (Stieglitz et al., 2017). Social bots are accounts that are based on 'computer algorithms that automatically produce content and interact with humans on social media, trying to emulate and possibly alter their behavior' (Ferrara et al., 2016, p. 96). The assumption that these accounts try to influence the human users' behaviour is one of the main drivers of research on this topic, as authors try to assess the actual importance of bots. One of the main domains of research has been the use of these bots in a political context (e. g. Abokhodair, Yoo, and McDonald, 2015; Brachten et al., 2017; Hegelich and Janetzko, 2016). While a possible influence in other scenarios may have limited impact, bot influence on politics, especially elections, may in theory alter their outcomes and thus affect the lives of many citizens. Elections with a global impact, such as the 2017 French presidential election (Ferrara, 2017) or the 2016 US presidential election (Bessi and Ferrara, 2016), are among the events researched in the political context, as are elections with a smaller scale or regionally limited importance such as the 2010 US midterm elections (Ratkiewicz et al., 2011) or 2017 German state elections (Brachten et al., 2017). Besides these scenarios, bots have also been researched in political conflicts such as the civil war in Syria (Abokhodair, Yoo, and McDonald, 2015) or the Ukrainian crisis (Hegelich and Janetzko, 2016). All of the studies found at least some use of social bots – and often,

their strategies could be observed. The studies showed that bots simulate broad agreement on certain topics, even though this consensus might not be prevalent in the general population (a strategy often called *astroturfing*), or try to hijack a hashtag by posting unrelated messages with the hashtag, thus misdirecting human users that search using these hashtags (Abokhodair, Yoo, and McDonald, 2015).

While this behaviour could be observed, the actual ability of bots to alter human behaviour has not been shown on a larger scale. There are, however, findings which suggest that human users treat social bots as they would other human users (Edwards et al., 2016), which suggests that these bots have the same potential to influence human behaviour as other human users. Furthermore, it could also be shown that, under certain circumstances, bots that directly address human users regarding a specific topic were able to alter the future behaviour of these users on social media (Munger, 2017). Some research also indicates that the number of sources of a message (i.e. the number of social bots that spread a message) rather than the number of messages from a single source determines the probability for this message to spread throughout a network (Mønsted et al., 2017). The authors used positive, encouraging messages to test their model, and state that the findings might have implications not only in a political context but also for improving marketing and advertising strategies. While the latter domain seems to hold great potential for social bot usage, little to no research in this context has been conducted. As the possibility of altering the behaviour of human users and of spreading messages in favour of a specific topic seems to exist, findings also suggest that user-generated content on social media can be used to point consumers to a particular product (i.e. make the target group aware that the product exists) and, in the form of user recommendations, affect users' purchasing intentions (Aggarwal and Singh, 2013; Bai, Yao, and Dou, 2015). The potential for misuse of automated communication in this context as well as the need for research in this area has become apparent.

A domain which seems suitable for this purpose is the music industry. In contrast to mere online retailers, there are social media music platforms on which a lot of user interactions take place. These platforms are often as unregulated as Twitter, on which most findings on social bots rely (Stieglitz et al., 2017). In addition, previous research has shown that social media activity has significant effects on music sales (Chen, De, and Hu, 2015).

2.2 Social media and the music industry

The digital transformation of the music economy is generally considered to have started in 1999 when peer-to-peer filesharing services emerged as a mainstream (albeit illegal) mechanism for online music distribution (Alderman, 2008; Knopper, 2009). During the years that followed, the digital transformation primarily affected music distribution and dramatically improved the consumers' access to music (Wikström, 2012). While at first analogue broadcast radio remained the most influential tastemaker, social media platforms soon emerged, such as MySpace in 2003 and Facebook, Twitter and YouTube in 2006, and they gradually took over that role (Mjos, 2013; Wikström, 2013). This logic has been further reinforced as music streaming platforms that serve as social media platforms in themselves (i.e. YouTube and SoundCloud) have become the de facto standard for music distribution (ibid.).

As these social media platforms increasingly gain influence as tastemakers in the music economy, the economic incentives also increase to try to influence the communication dynamics on these platforms in one direction or another.

3 Method

3.1 Data

SoundCloud is a music streaming platform headquartered in Berlin, Germany, that was launched in 2007 (Mahroum, 2016). The service allows its users to upload, listen and comment on audio files of all kinds (not only songs but also speeches and podcasts). While the platform has struggled to establish a sustainable

	Descriptive statistics[1]			Correlations			
	% Zero[2]	Mean	Max	Downld.	Fav.	Reposts	Comments
Playback count	4.7	1,119.15	116,423,945	.260[3]	.842[1]	.727[1]	.323[4]
Download count	87.2	1.86	92,535		.343[3]	.268[3]	.201[5]
Favouritings count	34.7	20.39	1,447,597			.849[1]	.328[4]
Reposts count	79.5	3.22	127,565				.514[4]
Comment count	85.8	0.96	15,132				

[1]All tracks (n = 15,850,069) [2]Percentage of tracks with value 0
[3]Downloadable tracks only (n = 2,409,814) [4]Commentable tracks only (n = 15,781,151)
[5]Downloadable and commentable tracks only (n = 2,366,729)

Table 1. Descriptive statistics and correlation table for track interactions on SoundCloud

business model (Cook, 2017), it has become one of the most widely used social music platforms with 40 million registered users (July 2013) and 175 million unique monthly listeners (December 2014) (Walker, 2015). While there are other social music platforms, none of these platforms have been able to challenge SoundCloud in terms of user numbers. This pivotal position of SoundCloud in the music economy and the social media features offered by the platform are the main reasons that the platform has been selected as the empirical context for this paper. In combination, they create an exceptional opportunity for studying the prevalence of bots on a large online platform and the bots' capability to influence the behaviour of the human users on the platform.

Social media analytics (Stieglitz et al., 2018) was used to analyse the behaviour of SoundCloud users. We collected the metadata of all publicly accessible tracks uploaded to SoundCloud between 1 January and 30 June 2017, including the users who uploaded them, all comments on these tracks and the authors of the comments. The data was collected through SoundCloud's restricted HTTP API[1], which we gained access to after an application. This API allows the crawling of tracks uploaded at an arbitrary point in the past. Data collection took place from July to September 2017. Each track was crawled at least a month after it was uploaded. To ensure that the metadata of tracks can be compared regardless of when they were uploaded, we only consider the first month of comments in the analysis.

The final dataset comprises a total of:

- **15,850,069 tracks**, including the number of times a track has been played, downloaded, favourited, reposted or commented on. Favouriting refers to publicly marking a track as a favourite, and reposting refers to adding a track to one's profile page to share it with one's followers. Each account can only favourite or repost a track once. Not all tracks permit downloading or commenting. While 15,781,152 tracks (99.6%) in the dataset are commentable, only 2,409,814 tracks (15.2%) can be downloaded. The metrics for these fields were calculated based only on the tracks that allow commenting or downloading, respectively. Descriptive statistics for the track interaction metrics are given in Table 1.
- **12,125,095 comments** on the tracks, including the date and time of each comment, the account making the comment and the text of the comment.
- **1,983,846 users** who wrote the comments, including the number of followers, the number of accounts being followed (at most 2,000), the number of playlists created, the number of tracks reposted and the number of tracks uploaded.

To understand how users' interactions with tracks are related, consider the correlations of these metrics (Table 1). These correlations mean that the data are consistent with our assumption that comments on tracks are associated with other forms of interaction. While the correlations are of course not enough to establish a sequence of events or causation, such as more comments on some tracks by bots leading to more plays by humans, this first result is enough to warrant further investigation.

[1] https://developers.soundcloud.com/docs/api/

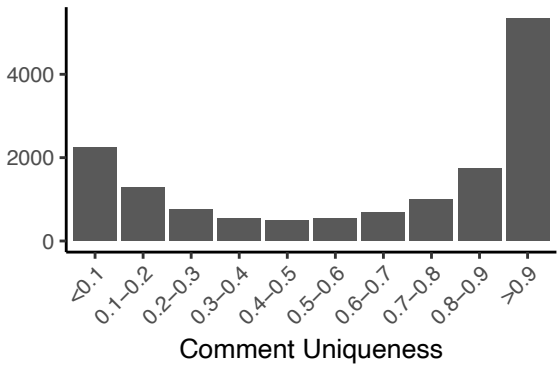

Figure 1. Number of users in each range of comment uniqueness. Only the 14,600 users who have written at least 100 comments are considered.

3.2 Comment uniqueness as a measure of behavioural complexity

It has been suggested that there are active bots on SoundCloud (Jordanous, Allington, and Dueck, 2015). These bots might interact with the tracks in several ways, such as download them, repost them or comment on them. For most forms of interaction, only aggregate counts are available. More detailed information can be examined for the comments, including the date, time, content and author of a comment. A preliminary look at the comments suggested that some of them seemed to be very repetitive. We therefore define a metric that captures how unique the comments posted by a user are. This method is similar to the identification of bots by examining the entropy of posting intervals. A high entropy reflects the behavioural complexity and unpredictability hypothesised to be typical of humans (Chu et al., 2012; Ghosh, Surachawala, and Lerman, 2011). Likewise, we propose an identification method based on the predictability of the texts of users' comments.

To measure the diversity of the comments by a user, we divide the number of distinct comments a user has made by the total number of comments k by the user. We call the result comment uniqueness (CU). Two comments are considered different if they differ by at least one character. Formally, let $C = (c_1, c_2, \ldots c_k)$ denote the tuple of comments posted by a user. Then,

$$CU = \frac{|\{c \in C\}|}{k}.$$

Comment uniqueness takes values between $1/k$ and 1. It will equal 1 if every comment by a user is different from the others. In contrast, it will equal $1/k$ if all of a user's comments are identical, and approach zero the more identical comments are made.

As described above, we assume that automated accounts post repetitive comments whereas contributions by humans are more diverse. This assumption does not hold for bots that generate arbitrary sequences of characters and post them as comments, or generate comments from complex rules. However, we assume that many bots follow simplistic rules such as choosing one out of a number of possible predefined comments (e.g. 'nice', 'great song', 'cool'). This type of activity would result in a very low comment uniqueness.

The complexity of a user's commenting behaviour can only be estimated reliably if the user has posted sufficient comments. For example, three identical comments are arguably not enough to reliably determine whether or not a user will always post repetitive comments. The minimum threshold used in the following is 100. It was deliberately set high, to examine whether there are accounts which can be confidently classified as suspicious. The 14,600 users with at least 100 comments only make up a tiny fraction (0.74 %) of all SoundCloud users, but are together responsible for almost a third (3,781,195 or 31.18 %) of comments.

Figure 1 shows that, contrary to what one might expect, these highly active users are divided into two groups by the comment uniqueness metric while only few occupy the space in between those groups. This

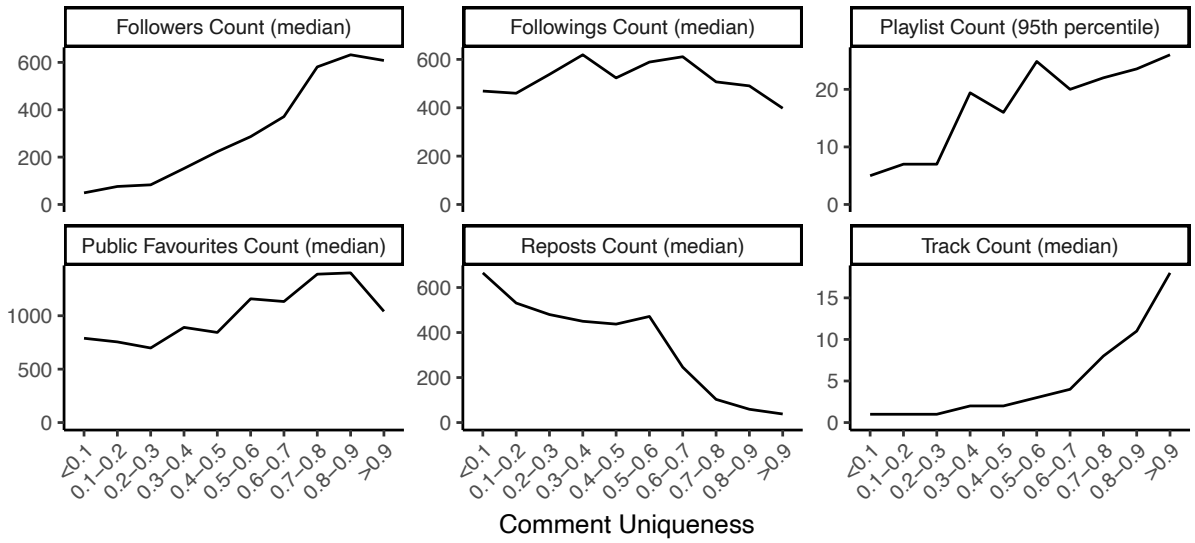

Figure 2. Statistics for users by comment uniqueness. Only users who have written at least 100 comments are considered. The median is used because it is less sensitive to outliers than the mean. With playlist count, the 95th percentile is used because the median is 0 in most groups.

suggests that comment uniqueness indeed offers a useful way of distinguishing between groups of users. For further analyses, such as visualising the user types in a network, a threshold to separate the groups is needed. A simple option is the global comment uniqueness, i.e., the number of distinct comments in the data set divided by the total number of comments ($5{,}706{,}663/12{,}125{,}095 = 0.471$).

4 Preliminary Results

4.1 Behaviour of the two groups

The previous section showed how the highly active users of SoundCloud can be divided into two separate groups based on their commenting behaviour. Not every active SoundCloud user can be clearly assigned one of these categories; instead, the users form a continuum between two extremes. However, the data in Figure 1 showed that most users are much closer to one of these two archetypes than the other. Next, we examine these extremes more closely to see how they differ in other behaviour.

Figure 2 compares comment uniqueness with other metrics of user behaviour. Users with a lower comment uniqueness (that is, those we consider potential bots) tend to have far fewer followers (people who are interested in updates about what they are doing). Between the two extremes, this difference is very pronounced, as the accounts with $CU > 0.9$ have 12.4 times the median number of followers as accounts with $CU > 0.1$ (608/49). They also upload far fewer tracks (with medians of 18 and 1, respectively) and create fewer playlists (although many active users do not create playlists at all). In contrast, they repost 17.5 times as many tracks (665/38). They also follow about the same number of people and publicly favourite about the same number of tracks.

We conclude that the comment uniqueness metric results in a meaningful distinction between two types of highly active users (i.e. those with 100 or more comments in the examined six-month timespan):

- The first type comprises those who actively participate in the community by writing varied comments, uploading tracks, and sometimes creating playlists, and who perhaps as a result have a high number of followers.
- The second type consists of those who are active, but write highly repetitive comments and mainly repost tracks uploaded by others, and who therefore do not contribute to the community as much.

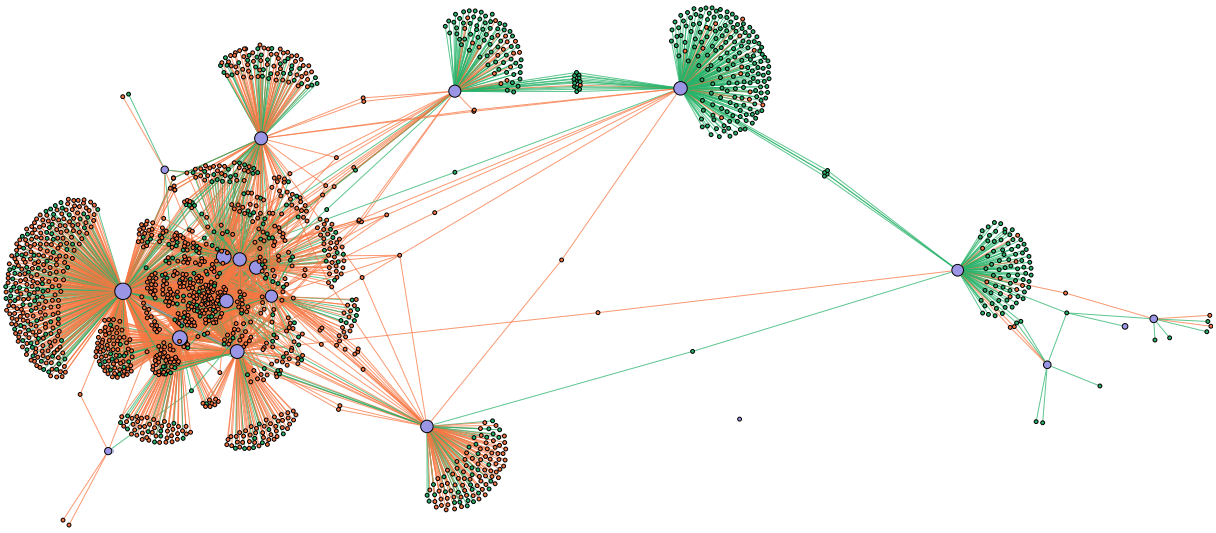

Figure 3. Network visualisation of the 25 most commented tracks on SoundCloud. Nodes represent users, edges comments. The 19 purple nodes are the track authors (some authors uploaded more than one of the 25 most commented tracks). Other users are only included if they made at least 100 comments (across all tracks). Users who post highly repetitive comments ($CU < 0.471$) are coloured red, users who post diverse comments ($CU > 0.471$) green.

4.2 Network visualisation

Given that the users can be divided into two groups which differ clearly in several behavioural attributes, the next question is: Do members of these groups interact with the same tracks or do the groups represent separate communities? To address this question, we visualised the network of highly active users and their comments (Figure 3) using the open source network visualisation package Gephi (Bastian, Heymann, Jacomy, et al., 2009). Several observations can be made.

At the centre of the network is a group of eight track authors (artists or publishers, shown as large nodes) whose tracks are frequently commented on by the same users. In this central group, the share of commenting 'bots' (red nodes) is much higher than for the other track authors. In addition, many of the 'bots' commented on tracks by several of the central authors, while many of the humans only commented on one. The other track authors have few or no commenting users in common with each other, or with this central group. In addition, the remaining track authors are mostly commented on by humans.

5 Discussion

5.1 Implications

We identified two archetypes of accounts on SoundCloud. The highly active users on the platform form two groups, which more or less correspond to the archetypes. One of these groups represents about a third of the highly active users, and they aggressively post repetitive comments. To validate the comment uniqueness metric, we examined the behaviour of these accounts in terms of other types of interactions (see Figure 2). Their low number of followers and lack of original content conform with prior findings on bot activity on social media (Ferrara et al., 2016; Ratkiewicz et al., 2011).

Our early results relying on comment uniqueness thus indicate that there is a substantial number of bots active on SoundCloud. These bots do not contribute to the platform in the way humans do, and in the way the platform was intended: by uploading tracks or curating content through the creation of playlists, both of which attract followers. The comment network for the top 25 most commented tracks (Figure 3) shows that the accounts exhibiting bot-like behaviour (shown in red) are highly focused on some of the tracks. It is therefore plausible that they might be attempting to influence the discussion on these tracks to prop

them up artificially or to increase the visibility of the commenting accounts themselves. This might help artists gain an unfair advantage over other artists who do not engage in such activities. However, there are very few 'human' accounts active for these tracks, so it is not clear whether the bots are successful in influencing activity by non-bot accounts.

Our findings have important implications for SoundCloud and the wider music industry, since a considerable amount of activity seems to be generated by bots or semi-automated accounts. This raises the question, for example, to what extent success on SoundCloud – in terms of the sheer number of interactions – depends on technical expertise rather than musical creativity. This finding also has implications for other domains, since bots could equally be used in other commercial contexts to promote goods, including other social networks and other types of products. SoundCloud, through its open API and large userbase, simply made it possible to study this phenomenon in a relevant context.

5.2 Limitations and outlook

The presented preliminary findings have several limitations which will be addressed in further research. Although the introduced comment uniqueness measure yields promising results, it currently only identifies one kind of bot behaviour, namely repetitive comments. Furthermore, human-operated accounts with a simple commenting behaviour could lead to false positives, while the comment uniqueness measure is not able to detect bots with more sophisticated, randomised routines either. To do so, we will examine other attributes and factors. Triangulating and complementing our results obtained with the comment uniqueness measure, for example with entropy measures regarding content and timing of comments, is the logical next step.

Furthermore, our network visualisation of comment activities is promising: a more detailed analysis of the comment network, as well as collecting the follower network of a sample of active users will provide us with more possibilities to unveil bot-like behaviour. Our results already point at differences in following behaviour, as the following count seems to be uniform between both groups of accounts while the follower count increases with comment uniqueness. Therefore, we are confident that community detection in combination with density, clustering, and centrality measures will offer another possibility to further support and extend our results.

In addition, it is not clear whether these suspected bots are successful at convincing human accounts that a track is more popular than it is, and thereby causing them to become more active on the tracks themselves. To examine this causal relationship, we plan to carry out a time series analysis that can establish the order of events, that is, whether comments by bots lead to subsequent comments by non-bot accounts. Additionally, we will use time-series data to determine if the selected period of one month from upload is optimal for a sample of track activity.

Finally, the collected data does not contain deleted accounts and their activities, because the SoundCloud API does not provide an archive. Therefore, bots already deleted by SoundCloud are excluded from our analysis. The true extent of bot activity is probably even higher than what we are able to observe.

6 Conclusion

Summarising, we have found a simple method to differentiate between two groups of SoundCloud accounts which exhibit characteristic behaviours. Further analysis, for example regarding the active contribution to the platform in the form of track uploads or curation of playlists, supported our interpretation of these groups exhibiting bot-like or human-like behaviour respectively. This strong evidence for the widespread existence and visibility of automated accounts on SoundCloud is by itself relevant and strongly suggests that further efforts in detecting and investigating these automated accounts are needed. The manipulation of activity records arguably impacts platform dynamics, thereby affecting the success of tracks. This not only has implications regarding the user experience of SoundCloud, but also cultural, monetary, and economical consequences for the platform, its users, and artists.

References

Abokhodair, N., D. Yoo, and D. W. McDonald (2015). "Dissecting a Social Botnet." In: *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing – CSCW '15*, pp. 839–851. ISBN: 9781450329224. DOI: 10.1145/2675133.2675208.

Aggarwal, R. and H. Singh (2013). "Differential influence of blogs across different stages of decision making: The case of venture capitalists." *MIS Quarterly* 37 (4), 1093–1112.

Alderman, J. (2008). *Sonic boom: Napster, MP3, and the new pioneers of music*. London: Fourth Estate.

Bai, Y., Z. Yao, and Y.-F. Dou (2015). "Effect of social commerce factors on user purchase behavior: An empirical investigation from renren.com." *International Journal of Information Management* 35 (5), 538–550. DOI: 10.1016/j.ijinfomgt.2015.04.011.

Bastian, M., S. Heymann, M. Jacomy, et al. (2009). "Gephi: an open source software for exploring and manipulating networks." In: *Proceedings of the Third International Conference on Weblogs and Social Media (ICWSM-2009)*. Vol. 8, pp. 361–362.

Bessi, A. and E. Ferrara (2016). "Social bots distort the 2016 U.S. Presidential election online discussion." *First Monday* 21 (11). ISSN: 13960466. DOI: 10.5210/fm.v21i11.7090.

Brachten, F., S. Stieglitz, L. Hofeditz, K. Kloppenborg, and A. Reimann (2017). "Strategies and Influence of Social Bots in a 2017 German state election." In: *Proceeedings of the Australasian Conference on Information Systems*, Paper 207.

Bruns, A., B. Moon, F. Münch, P. Wikström, S. Stieglitz, F. Brachten, and B. Ross (2018). "Detecting Twitter Bots That Share SoundCloud Tracks." In: *Proceedings of the International Conference on Social Media + Society*.

Chen, H., P. De, and Y. Hu (2015). "IT-enabled broadcasting in social media: An empirical study of artists' activities and music sales." *Information Systems Research* 26 (3), 513–531. DOI: 10.1287/isre.2015.0582.

Chu, Z., S. Gianvecchio, H. Wang, and S. Jajodia (2012). "Detecting automation of twitter accounts: Are you a human, bot, or cyborg?" *IEEE Transactions on Dependable and Secure Computing* 9 (6), 811–824. DOI: 10.1109/TDSC.2012.75.

Cook, J. (2017). *SoundCloud raised $70 million in debt funding*. URL: http://uk.businessinsider.com/soundcloud-raises-70-million-debt-funding-companies-house-documents-2017-3 (visited on 10/13/2017).

Edwards, C., A. J. Beattie, A. Edwards, and P. R. Spence (2016). "Differences in perceptions of communication quality between a Twitterbot and human agent for information seeking and learning." *Computers in Human Behavior* 65 (Supplement C), 666–671. ISSN: 0747-5632. DOI: https://doi.org/10.1016/j.chb.2016.07.003.

Ferrara, E., O. Varol, C. Davis, F. Menczer, and A. Flammini (2016). "The rise of social bots." *Communications of the ACM* 59 (7), 96–104. DOI: 10.1145/2818717.

Ferrara, E. (2017). "Disinformation and social bot operations in the run up to the 2017 French presidential election." *First Monday* 22 (8). ISSN: 13960466. DOI: 10.5210/fm.v22i8.8005.

Ghosh, R., T. Surachawala, and K. Lerman (2011). "Entropy-based Classification of 'Retweeting' Activity on Twitter." In: *Proceedings of the KDD Workshop on Social Network Analysis (SNA-KDD)*.

Gu, B., J. Park, and P. Konana (2012). "The impact of external word-of-mouth sources on retailer sales of high-involvement products." *Information Systems Research* 23 (1), 182–196. DOI: 10.1287/isre.1100.0343.

Hegelich, S. and D. Janetzko (2016). "Are social bots on Twitter political actors? Empirical evidence from a Ukrainian social botnet." In: *Proceedings of the 10th International Conference on Web and Social Media, ICWSM 2016*, pp. 579–582. ISBN: 9781577357582.

Jordanous, A., D. Allington, and B. Dueck (2015). "Measuring cultural value using social network analysis: A case study on valuing electronic musicians." In: *Proceedings of the Sixth International Conference on Computational Creativity*, pp. 110–117.

Knopper, S. (2009). *Appetite for self-destruction: The spectacular crash of the record industry in the digital age*. New York: Free Press.

Mahroum, S. (2016). "SoundCloud in Berlin, Sharing the Vibe." In: *Black Swan Start-ups: Understanding the Rise of Successful Technology Business in Unlikely Places*. London: Palgrave Macmillan UK, pp. 59–72. ISBN: 978-1-137-57727-6. DOI: 10.1057/978-1-137-57727-6_4.

Mjos, O. J. (2013). *Music, social media and global mobility: MySpace, Facebook, YouTube*. New York: Routledge Chapman & Hall.

Mønsted, B., P. Sapieżyński, E. Ferrara, and S. Lehmann (2017). "Evidence of complex contagion of information in social media: An experiment using Twitter bots." *PLOS ONE* 12 (9), 1–12. DOI: 10.1371/journal.pone.0184148.

Munger, K. (2017). "Tweetment Effects on the Tweeted: Experimentally Reducing Racist Harassment." *Political Behavior* 39 (3), 629–649. ISSN: 1573-6687. DOI: 10.1007/s11109-016-9373-5.

Ratkiewicz, J., M. D. Conover, M. Meiss, B. Gonçalves, A. Flammini, and F. Menczer (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. eprint: 1011.3768.

Stieglitz, S., M. Mirbabaie, B. Ross, and C. Neuberger (2018). "Social media analytics – Challenges in topic discovery, data collection, and data preparation." *International Journal of Information Management* 39, 156–168. DOI: 10.1016/j.ijinfomgt.2017.12.002.

Stieglitz, S., F. Brachten, B. Ross, and A.-K. Jung (2017). "Do Social Bots Dream of Electric Sheep? A Categorisation of Social Media Bot Accounts." In: *Proceedings of the Australasian Conference on Information Systems*, Paper 206.

Varol, O., E. Ferrara, C. A. Davis, F. Menczer, and A. Flammini (2017). "Online Human-Bot Interactions: Detection, Estimation, and Characterization." In: *Proceedings of the Eleventh International Conference on Web and Social Media*, pp. 280–289.

Walker, R. (2015). *Can SoundCloud Be the Facebook of Music?* URL: https://www.bloomberg.com/news/features/2015-07-10/can-soundcloud-be-the-facebook-of-music- (visited on 10/13/2017).

Wikström, P. (2012). "A typology of music distribution models." *International Journal of Music Business Research* 1 (1), 7–20.

— (2013). *The music industry: Music in the cloud*. Cambridge: Polity Press.

Zhou, W. and W. Duan (2016). "The sales impact of word-of-mouth distribution across retail and third-party websites." In: *Proceedings of the Thirty-seventh International Conference on Information Systems (ICIS)*.

Acknowledgements

This work was supported in part by the UA-DAAD project "Emergent Music Engagement Practices via SoundCloud" and in part by the Deutsche Forschungsgemeinschaft (DFG) under grant No. GRK 2167, Research Training Group "User-Centred Social Media".

Paper 8: Virtual Moderation Assistance- Creating Design Guidelines for Virtual Assistants Supporting Creative Workshops

Fact sheet

Type	Conference paper
Authors	Strohmann, Timo; Fischer, Simon; Siemon, Dominik; Brachten, Florian; Robra-Bissantz, Susanne; Stieglitz, Stefan; Lattemann, Christoph.
Year	2018
Venue	Pacific Asia Conference on Information Systems (PACIS)
Full citation	Strohmann, T., Fischer, S., Siemon, D., Brachten, F., Robra-Bissantz, S., Stieglitz, S. & Lattemann, C. (2018). Virtual Moderation Assistance-Creating Design Guidelines for Virtual Assistants Supporting Creative Workshops. In: *Proceedings of the 22nd Pacific Asia Conference on Information Systems* (PACIS).

Virtual Moderation Assistance: Creating Design Guidelines for Virtual Assistants Supporting Creative Workshops

Completed Research Paper

Timo Strohmann
Technische Universität Braunschweig
Braunschweig, Germany
t.strohmann@tu-braunschweig.de

Simon Fischer
Jacobs University Bremen gGmbH
Bremen, Germany
si.fischer@jacobs-university.de

Dominik Siemon
Technische Universität Braunschweig
Braunschweig, Germany
d.siemon@tu-braunschweig.de

Florian Brachten
University of Duisburg-Essen
Duisburg, Germany
florian.brachten@uni-due.de

Christoph Lattemann
Jacobs University Bremen gGmbH
Bremen, Germany
c.lattemann@jacobs-university.de

Susanne Robra-Bissantz
Technische Universität Braunschweig
Braunschweig, Germany
s.robra-bissantz@tu-braunschweig.de

Stefan Stieglitz
University of Duisburg-Essen
Duisburg, Germany
stefan.stieglitz@uni-due.de

Abstract

To remain competitive, businesses need to develop innovative and profitable products, processes and services. The development of innovation relies on novel ideas, which can be generated during creative workshops. In this context the Design Thinking approach, a problem-solving methodology based on collaboration, user-centricity and creativity, may be used. However, guidance and moderation of this process require a vast amount of skills and knowledge. As technologies like artificial intelligence have the potential of making machines our collaboration partner in the future, creating virtual assistants adapting human behaviors is promising. To reduce cognitive dissonance and stress on both the moderators and participants, we investigate the potential of a virtual assistant to support moderation in a Design Thinking process to improve innovative output as well as perceived satisfaction. We therefore developed design guidelines for virtual assistants supporting creative workshops based on qualitative expert interviews and related literature following the Design Science Research Methodology.

Keywords: Virtual Assistant, Creative Workshops, Moderation, Design Thinking, Creativity, Artificial Intelligence

Introduction

In a world with fierce competition and where new technologies emerge constantly, businesses and professionals need to adapt to challenges and chances. To survive it is inevitable to develop innovative and profitable products, processes and services to remain competitive (Weerawardena and Mavondo 2011; Kung and Schmid 2015). The development of innovations relies on the ability to conceive, appraise and refine novel ideas. (Helpman et al. 2010; Hennessey and Amabile 2010; Somech and Drach-Zahavy 2013; Forés and Camisón 2016). As creativity is hindered in competitive environments, sustaining innovation processes relies on creative individuals and groups (Siemon et al. 2015). To foster creative thinking in corporate environments dedicated creativity workshops became increasingly popular since cost and time efficiency are major concerns for firms (Gabriel et al. 2015).

As the business environment becomes more complex, problems become wicked. To solve wicked problems, which neither have a definitive formulation or solution, teamwork becomes necessary, as research shows that team efforts outrun the cumulated individual outcome (Buchanan 1992; Connoly et al. 1993; Paulus 2000; Santanen et al. 2000, 2004; Nijstad and Stroebe 2006; Bell and Kozlowski 2008; Paulus and Nijstad 2009). In this context, Design Thinking (DT) gained attention in corporate and scientific communities due to its holistic and collaborative approach for the innovation of products, services, product-service systems or processes (Chasanidou et al. 2015). DT is a problem solving methodology that is used to address *wicked problems* by utilizing creative and user-centered approaches in heterogeneous groups (Buchanan 1992; Brown 2008; Stickdorn and Schneider 2010; Carlgren et al. 2014). In this iterative and emerging process a variety of methods, tools and techniques engage participants both cognitively and empathically (Brem and Spoedt 2017). Regardless of what approach is used, these processes rely on a large amount of information and knowledge, which calls for the implementation of knowledge management or group support systems. However, addressing both moderators and participants with one system calls for a differentiated evaluation of processes (Gabriel et al. 2015). Recent developments in the field of artificial intelligence may offer the tools to address this issue.

Following this line of thought, we examine how the moderation of DT workshops may be supported through the utilization of a Virtual Assistant (VA). We explore the components of such a VA and propose guidelines for its design using the Design Science Research Methodology (DSRM). Our generated design guidelines can be used by researchers, as well as designers implementing specific VAs for the support of creative processes to improve innovative and creative processes like the DT approach in a workshop setting. The development of our guidelines is based on a qualitative research approach with five expert interviews as well as the review of existing research results. In this paper we address the question: **What are the requirements that need to be considered to design VAs supporting DT workshops?**

The remainder of this paper is structured as follows. First, we present the theoretical background on DT and creativity workshops as well as VAs. In the third section, we present our research methodology followed by the guidelines we propose for VAs in the fourth section. Lastly, we discuss the implications, limitations and potentials of these results and conclude our findings.

Theoretical Background

Design Thinking and Workshop Settings

DT is a holistic, collaborative approach for the innovation of products, services, product-service systems or processes that is inherently based on creativity, multidisciplinarity, and user-centricity (Brown 2008; Stickdorn and Schneider 2010; Carlgren et al. 2014). The origins of DT can be traced back to the end of the 1960s when Simon (1969) introduced the development of artifacts based on "designerly thinking", which conveyed a design science procedure as a useful interdisciplinary approach in the area of business and economics. The terms "design" and "thinking" connotes a strong connection to the manner designers or the science of design operates (Brenner et al. 2016). DT is a stand-alone strategic innovation development approach that consists of a process, methods, and a mindset (Brenner et al. 2016).

DT employs divergent thinking to create choices through insights into consumer behavior, alternative solutions and interactive experiences to improve creative activities such as ideation (Brown 2008). This not only avoids efficient brainstorming that fosters conservative and inflexible ideas, it also heavily increases complexity, and thus convergent methods are used to empower participants to eliminate options and make choices (Martin 2009). As DT is a user centered approach to problem solving, empathy builds the foundation to reveal the users core needs and problems (Bellet and Maloney 1991; Brown 2008). Therefore the DT process in its entirety is focused towards understanding the context of needs and problems, observing user interaction with humans and objects, refining the issue before coming up with ideas and merging them to a product, service or process prototype, that is presented to and tested with the user (Buchanan 1992). As the users' issues can be manifold, there is a variety of methods for each process phase (Understand, Observer, Point-of-View, Ideate, Prototype and Test) that allow for a tailored creation process. These phases heavily alter between their intended modes of thinking. DT relies on inductive (generalization), deductive (prediction of consequences) and also on abductive reasoning (formation of explanatory hypothesis) (Brown 2008). The latter is considerably less prominent in business and engineering education but vitally important for generating novel ideas (Dunne and Martin 2006). The underlying DT mindset involves aspects such as user-centricity and abductive reasoning that are not easily taught and internalized by participants with traditional business mindsets (Heiman and Burnett 2007). Additionally, while performing DT workshops, the design of the space and surroundings, such as moveable furniture, tools and materials, visualization of new ideas are used to foster creativity and interaction (Grots and Pratschke 2009). This interactive communication is considered as an important feature of DT (Brereton and McGarry 2000) as the "DT team" needs to organize, share and develop their ideas in the early phases of the DT process and must often interact practically during the development of prototypes (Brown 2009).

Guidance and moderation of this process is usually left to one or two trained design thinkers per group as, although rising in popularity, DT education is scarce (Heiman and Burnett 2007). Guiding this process often encounters barriers as participants are accustomed to traditional innovation or decision processes using deterministic and analytical thinking that some DT phases and methods deliberately avoid in order to foster new ideas (Heiman and Burnett 2007; Chasanidou et al. 2015). This induces that individuals or groups find themselves stuck in arguments or abandoned by creativity (Morehen et al. 2013). Utilizing dedicated methods within the DT process while encouraging creativity, interaction and fun as well as providing appropriate surroundings, tools, materials and visualizations to ensure a smooth process therefore is a challenge for moderators. Hence, we investigate the potentials of a VA to support DT moderation in a DT process to reduce cognitive dissonance and stress on moderators and participants to improve process success and perceived satisfaction.

Virtual Assistants

There are various synonyms for VAs such as virtual personal assistants (McTear et al. 2016), vocal social agents or digital assistants (Guzman 2017), voice assistants (McTear et al. 2016) and intelligent agents (Balakrishnan and Honavar 2001). Similar to the variety of names there is also no single definition, but several definition approaches as well as tasks and characteristics of VA, which are considered below. VAs are computer software programs performing functions in an intelligent manner, like helping users on their own initiative (Skalski and Tamborini 2007). According to Balakrishnan and Honavar (2001) an intelligent manner in terms of VAs can be defined as *"behaviors that would be characterized as intelligent if performed by normal human beings under similar circumstances"* (Balakrishnan and Honavar 2001). VAs carry out tasks for the users and can answer various questions (Zhao 2006). The interaction takes place in a human-like way. Thus, VAs need to interact with humans in a social way, follow social norms of interpersonal communication and need to be able to use and understand natural human language (Zhao 2006; Skalski and Tamborini 2007; Guzman 2017). Humans can interact with VAs via voice or text (Chatbots) and the VA might respond either with voice, text or visual information (Pearl 2016).

According to Balakrishnan and Honavar (2001) there are several requirements for VAs to perceive input from its environment and act upon it. To perform an action the VA has to interpret the given input. The actions are then chosen autonomously according to the intended goal to be achieved. As VAs are usually

used in dynamic environments, mechanisms to learn and adopt from instruction to instruction are needed for sustainable, reliable operation. Lastly, a VA has to be able to collaborate and communicate in groups.

While VAs are used in more and more areas of everyday life (McTear et al. 2016), like Siri, Google Now, Microsoft Cortana, Amazon Echo or Google Home (Pearl 2016), there are only few examples where VAs support creative processes (Siemon et al. 2015; Strohmann et al. 2017), concrete design instructions are missing. Seeber et al. (2018) states that artificial intelligence technology like VAs have the potential to become our smart collaboration partner in the future which is why the opportunities enabled by smart technology for collaborative processes have to be explored. Our Guidelines for VAs supporting the moderation of creative processes are a step into this direction.

Methodology

The objective of our research is to generate design guidelines that can be used by researchers, as well as by designers to develop and implement specific VAs for the support of creative processes. We follow the DSRM, which aims to create innovative and novel artifacts in order to solve organizational and human problems (Hevner et al. 2004). Several process models and frameworks for Design Science exist that cover the extent and specification of the artifact that is designed within the methodology. According to Gregor and Hevner (2013) the levels of artifact abstraction in DSR is vast and ranges from models, processes, instantiations, methods or software. Artifacts can be described by the two attributes "abstract" and "specific" and by their general maturity of knowledge. Design guidelines and constructs, methods, models, or technological rules are artifacts that are in between a solely abstract or specific artifact and contribute nascent design theories or operational principles. In our research, the design guidelines that serve as a foundation for further refinement, adjustment or implementation, thus, represent the design artifact.

The development of our guidelines is based on a qualitative research approach with five expert interviews as well as existing research results. Therefore, we have reviewed literature concerning the topics Moderation, DT and VAs. The findings of reviewing the literature form the basis of the interview guidelines, as well as for validation and refinement of our proposed design guidelines.

Expert Interviews

Due to the lack of previous research we conducted five expert interviews. Expert interviews have grown in popularity as a reliant method to obtain knowledge that is otherwise hard to discover (Bogner et al. 2009; Meuser and Nagel 2009). As the term *expert* is used inflationary (Meuser and Nagel 2009), in this study the interviewees' expertise is distinguished from "everyday" and "common-sense knowledge" as they have been either participants or moderators in various creativity or DT workshops for at least two years. Additionally, three participants have an IS research background. Grounded in the novelty of this endeavor the expert interviews were employed as an exploratory tool, and hence conducted as openly as possible as suggested by Bogner et al. (2009).

Participants

Potential participants were directly approached by the researchers. Prior to the selection of the experts, we defined specific requirements. For the purpose of our research, an expert should have at least two years of experience in planning and conducting creativity workshops (e.g. DT) or experience in being a workshop participant. In order to gain divergent insights, we sought experts from different professions. A background in IS research, ideally artificial intelligence, knowledge management or group support systems, was optional but ideal. We acquired two IS researchers (E1 and E4) with experience as participants in various creativity workshops including DT. One researcher and DT coach (E2) and one freelancer, specializing in innovation workshops and change management with a system engineering and design background (E5). One expert is an employee of a financial services company (E3) and works as a manager of innovations. This expert has more than 10 years of workshop experience (participants and moderator). Participants were 31 years old on average, with one female and four male experts. The interviews took place at the workplaces of the experts.

Semi-structured Interviews

An interview guideline was developed based on literature from the area of moderation and DT as well as on prior experience, research on design science and existing design principles for VAs. As the interviews were to be carried out on DT and creativity workshop experts from different backgrounds, semi-structured interviews were chosen to give the interviewees enough space to elaborate on issues and the possibility to point out aspects that had not come up in the preparation of the guideline. Furthermore, the method should also ensure that all relevant aspects were captured and in regard to the coding process, answers were at least in part comparable. The interview guidelines were roughly divided into six parts according to previous studies such as Cohen (2004) and Pearl (2016), who specified requirements regarding representational (part 4) and interaction design (part 5) of VAs. The whole interview was structured as follows:

- Part 1 introduced the interviewer and again summed up the purpose of the research as the participants had already received relevant information when they were recruited.
- Part 2 asked the interviewee to elaborate on his background as well as his current occupation and relation to the DT process.
- Part 3 consisted of questions considering the general approach to the DT process and the tasks of a moderator therein.
- Part 4 asked the interviewee for his evaluation of the potential for support by VAs within the DT process and the importance of different appearances of such VAs.
- Part 5 included questions as to how interactions with a VA could ideally take place within the situation of a DT workshop.
- Part 6 concluded the interview with the possibility for the interviewee to give an outlook on how she thinks the future of VAs might look like as well as to state aspects that had not been addressed in the interview.

The interviews were conducted between 20th December 2017 and 25th January 2018 and took between 55 and 65 minutes. After conducting the interviews, they were transcribed and coded in MAXQDA version 18. The following section describes the coding in detail.

Coding of the Interviews

The expert interviews are analyzed using codes as an efficient data-labeling and data-retrieval device (Miles and Huberman 1994). Following Miles and Hubermanns (1994) advice prior to the first expert interview an initial list of general codes was created. As Saldaña (2015) stated several coding cycles are needed for analyzing qualitative data. The coding effort has therefore been divided into three cycles, going from a general view on the data to a more specific and was done following Saldañas (2015) manual for qualitative researchers. The three coding cycles are visualized in Figure 1 with an adapted version of Saldañas (2015) code-to-theory model for qualitative inquiry.

The coding was done collaboratively by four researchers to distribute the effort of conducting and coding the interviews as well as to get different views on the qualitative data. According to Miles and Huberman (1994) the first cycle of coding started directly after each interview. With each newly coded interview the codings done so far were refined and aggregated with the new codings. Hence every coding cycle had several iterations. Additionally to the initial list of codes, more codes emerged inductively during the analyzing progress and were added to the list (Miles et al. 2014).

The coding effort in the first cycle was distributed to the four researches who coded their respective interviews. One of the researchers was appointed the codebook editor. Following Guest and MacQueen (2008) the codebook editor is responsible to update, revise and maintain the list of codes for the group. So, in our case the editor combined and aggregated all codings after each new interview, cycle and iteration. In the first coding cycle the general codes defined prior are used to get the data structured. General codes like "Tasks of the Moderator" or "Design Thinking Process" were assigned to the data. The first coding cycle structured the data and prepared it for the second coding cycle.

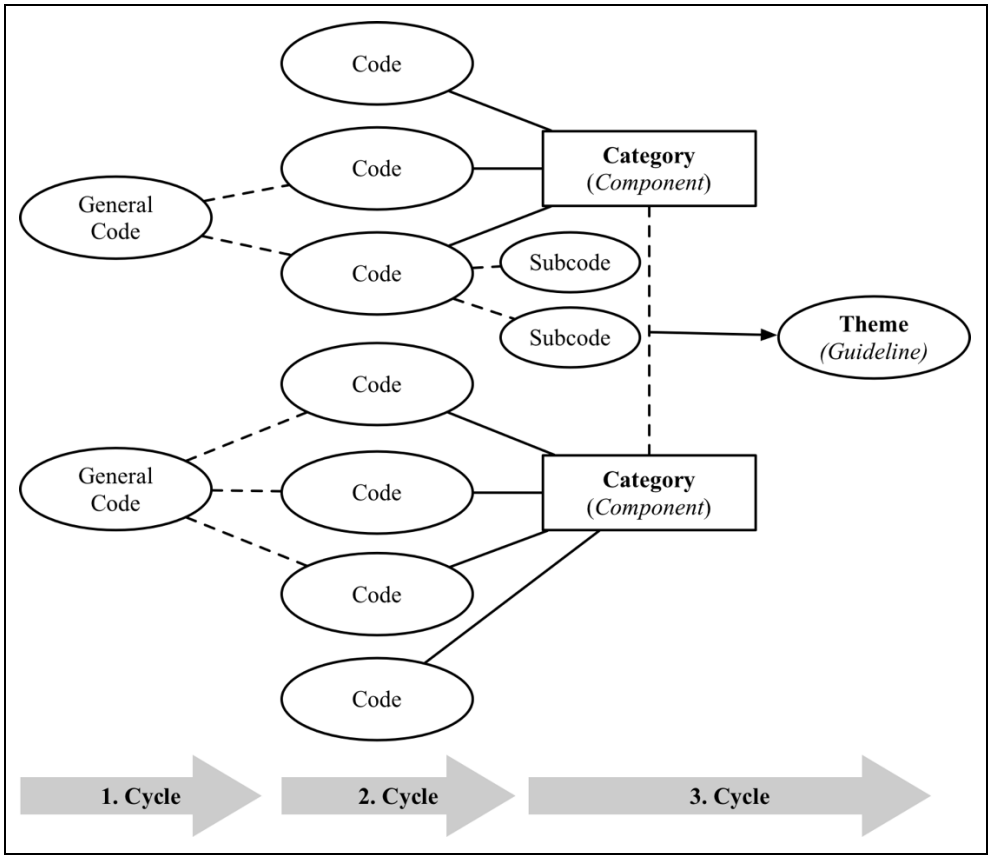

Figure 1. Code-to-Guideline Model adapted from (Saldaña 2015)

In the second coding cycle the researchers dived deeper into the data and created specific codes and subcodes. For example, a general code like "tasks of the moderator" got subdivide into subcodes like "acquire participants" or "intervene in case of stagnation". In this phase the actual meaning of the data got uncovered. In summary, 219 codes and 537 codings were created.

In the third and last cycle of coding the specific codes and subcodes of the second cycle were combined to categories (components of the guidelines) which were then structured into themes (guidelines).

Creating the Design Guidelines for Virtual Assistants

The design guidelines were derived from the results of the three coding cycles. As the coding started with a general view, it was specified from iteration to iteration and cycle to cycle. In the last coding cycle the focus was set specifically on possible design guidelines. They were validated and refined using corresponding literature. The final guidelines are structured into three main categories: First general conditions for a VA supporting creative processes, second characteristics of the VA and third tasks of the VA (divided into preparation and execution).

The components of each guideline as well as the guidelines themselves are shown in the following visualized in three tables. The components are directly related to the codes from the interview coding. Components with a high relevance (mentioned by 4 or more experts) are highlighted in bold. If there

are contrary statements of the experts leading to opposing components, both opinions are listed in the column *Components* with a "*vs.*" between them. Some components are divided into subcomponents which are listed as bullet points under the associated component. The literature used for validation and refinement is listed in the column *Supported By*. In addition, several components are corroborated with quotes from the expert interviews. As the interviews were conducted in German language, the quotes are translated into English.

General Conditions for a Virtual Assistant supporting Creative Processes

The first category is about general conditions that should exist for a VA supporting creative processes. Guideline 1.1 reflects the fear of the participants that machines get uncontrollable. As Expert 3 asked: *"How far do I humanize machines, so that I can allow them either manipulate me or not manipulate me?"*. Thus, while designing a VA the designer has to mind ethics, be conscientious and create an environment in which the user feels comfortable.

As all interviewees are moderators themselves it is clear that they do not want a VA to replace them. Looking at this from a technological perspective, it presumably is too early to completely replace a human moderator. Instead a VA may first support and collaborate with a human moderator and deepen its functionalities step by step. This is why we propose guideline 1.2 containing tasks and use cases the interviewed experts would like to have support for and also think a VA is able to support.

Table 1. General Conditions

Guideline	Component	Supported By
1.1 Create a trustful and transparent environment	Mind ethics	Bostrom and Yudkowsky 2014
	Be careful with growth of machine intelligence	
	Minimize feeling of observation	
1.2 A VA supporting (not replacing) the human moderator	Method and tool support	Strohmann et al. 2017
	Giving tips and nudges	
	Provide multimedia-based instructions	
	Process support	
	Explain rules and remember to follow them in case of violation	
	Documentation	
	Time-keeping	
	Consequent reminder	
1.3 Conditions for proactive actions	Human moderator has to accept proactive actions / VA has to cooperate with human moderator	Friedman 1989; Pearl 2016
	proactive action only with high quality	
	Acting proactively based on events (external or workshop-related)	
	Context (emotions, hierarchies, no personal criticism, concentration, interruption) and sensitiveness important for proactive features	

Guideline 1.3 summarizes conditions for proactive actions. Although planning and preparing the workshop is essential to its quality, the moderator needs to adjust to the participants needs and the emerging topics. The VA respectively is expected to proactively support ad hoc changes based on events. For example, if the participants are stagnating, the VA needs to recognize this and act on it. In such situations the participants would normally not be able to act by themselves, which is why proactive actions are needed. The context of a situation during a workshop is one of the most important

components for performing proactive actions mentioned by all experts. When thinking of widely used VA like Siri or Google Now, most of the time they only act when prompted. For a VA moderating a creative process it is vital to refrain from being limited to reactive interaction, but also, for example, act based on a certain event or context. A requirement for that is high quality of the action, meaning it is seen as appropriate by the participants and moderator and it has to fit to the situation. Expert 3 reinforces this by saying: *"Equating machines with humans is heavily depended on the quality. (…) If the assistant expresses unqualified garbage at the wrong time, then it is a tool, I want to turn off."*. Thinking of guideline 1.2 the human moderator has to tolerate and accept a proactive behavior of the VA, as it is prone to interfere with the moderator's course of action. Thus, it may also be conceivable that the VA cooperates with the human moderator by just giving hints for proactive actions.

Characteristics of the Virtual Assistant

The second category contains guidelines concerning characteristics of the VA to ensure a high quality of interaction and representational consistency. Guideline 2.1 aims for the appearance of the VA. Combining the expert's statements it turns out that the VA should appear more like a human than as a robot. While the technology should not be visible (Expert 3 and 5) it has to be clear, that the VA is not a real human, which corresponds to guideline 1.1. Beside to a human appearance, like face and body (Expert 4), the VA may also use gestures or expressions (Expert 2). If the VA interacts using voice, the gender can either be male or female. Expert 4 thinks that a male voice would be more authoritarian, while Expert 1 says it should be possible to choose the gender of the VA.

Table 2. Characteristics

Guideline	Component	Supported By
2.1 More human than robot	Human appearance (face, body, visual specialties) vs. no human appearance	Cohen et al. 2004; Pearl 2016
	Facial expressions and gestures	
	Technology should not be present / visible	
	Voice (Male vs. Female, Possibility to choose)	
2.2 Seamless way of interaction supported by intelligent features	Lighten the mood (jokes, anecdotes)	Zhao 2006; Skalski and Tamborini 2007; Hassenzahl et al. 2015; Pearl 2016; Guzman 2017
	Visualize, guide and support methods	
	Knowledge about methodology and tools	
	Collect / Use knowledge from previous workshops	
	Recognize emotions	
	Easy to use	
	Different levels of support	
	Reaction by intent, gesture, keyword or button	
	Audiovisual interaction	
	Recognize / Differentiate voices of participants - Measure participation - Identify personal traits / characteristics	
	Natural way of interaction	
	Follow social norms of interpersonal communication	
	Understand natural human language	

2.3 Personality helps for acceptance	Authority	Nielsen 1995; Cohen et al. 2004; Pearl 2016
	Humanly answer for non-existent functionalities	
	Name	
	Mannerism	
	Show emotions	

Guideline 2.2 is all about the way of interaction between humans and the VA as well as the use of intelligent features. As the components of guideline 2.2 originate from requirements of DT experts, they naturally apply to DT Workshops. However, the examined literature on creative workshops suggests a high coherence for creative workshop support in general, as a similar atmosphere and work modes are intended. The VA should interact with the participants in a natural way, meaning that the participants can use natural language as they would with a human moderator. Several features might be used to make the interaction as natural and sophisticated as possible. Collected knowledge from previous workshops, for example, enables machine learning. With the recognition of emotions and different voices of participants, the VA can respond individually, fit the situation and learn about the participants. Expert 2, for example, said *"The virtual assistant could identify personal characteristics of the participants, in order to be able to act according to them when a critical situation occurs."*. Besides the proactive actions mentioned in guideline 1.3, the VA has to be addressable by the participants using an intent, gesture, keyword or button.

Guideline 2.3 is related to guideline 2.1 and focuses on the personality helping for acceptance. Therefore human-related behaviors like mannerisms or emotions as well as a name might help the participants accepting the VA.

Tasks of the Virtual Assistant as Moderation Support

The third and last category considers the tasks of the VA as moderation support. In this respect, we propose three guidelines containing several components and list supporting literature from the field of moderation of creativity and DT workshops. Guideline 3.1 deals with the preparation of the workshop, i. e. all activities that take place prior to the workshop. According to the interviewees, one fundamental component is the definition of the design challenge. Furthermore, several organizational tasks have to be handled in advance (e. g. the preparation of the workshop concept, the acquisition of participants or the organization of dates).

Table 3. Tasks of the Virtual Assistant as a Moderator

Guideline	Component	Supported by
3.1 Preparation of the Workshop	**(Support to) Define the design challenge**	Brem and Spoedt 2017
	Prepare methods, tools, concept and time table - Catalog of methods - First (introducing) questions - Dive into topic	Adam and Trapp 2015
	Acquire participants	
	Organize dates and detail the agenda	
	(Support) team composition	Geschka 1986

3.2 Active Support of the Execution	**Motivate, activate and stimulate**	Adam and Trapp 2015
	Encourage mindset and create atmosphere	Steinert 1992; Adam and Trapp 2015
	Intervene in case of problems (fast recognition and reaction as a requirement, motivation, explain methods, remind to follow rules, make adjustments)	Geschka 1986; Lempiala 2010; Adam and Trapp 2015
	Post processing and documentation	Brem and Spoedt 2017
	Keep results of current and prior methods or phases present	
	Time and process management - Initiate breaks - Initiate transition to new phase - Iteration	
	Visualize, explain, support and suggest methods and tools	Nielsen 2012; Adam and Trapp 2015; Gabriel et al. 2015
3.3 Passive Support of the Execution	Observe and record activities	Lempiala 2010; Nielsen 2012; Adam and Trapp 2015; Brem and Spoedt 2017
	Recognize / detect if goal is reached	
	Recognize / detect problems (Stagnation, disruption, boredom, demotivation, frustration, disorientation, loss of focus, diverging from the topic, misuse or non-acceptance of methods, criticism, violation of rules, too similar ideas, challenge is not the real problem)	
	Recognize / detect results	
	Recognize / detect the process	

The other two guidelines are related to the execution of the workshop. The support during the workshop is subdivided into active and passive support. The active support (guideline 3.2) refers to actions that are visible for the participants, like the time management.

The passive support (guideline 3.3) involves activities that are invisible for the participants and are all about recognition and detection. These passive activities form the basis for a possible active intervention or contribution. For example, if a problem like the group is stagnating is detected (see *Recognize / detect problems*, guideline 3.3), the VA has to intervene (see *Intervene in case of problems*, guideline 3.2). A possible intervention is to motivate the participants using motivational quotes (Expert 1), requiring fast recognition and reaction.

Limitations and Potentials of the Virtual Assistant

Several potentials and limitations of the use of a VA for the moderation of DT workshops were unveiled during the expert interviews. Expert 4 mentions the limited input options of voice- and text-based interaction. Expert 1 extends this view with his concerns, that an interaction with a VA is too restrictive and that a participant would perhaps feel uncomfortable with the intervention of VA. The same issue is addressed by Expert 2, saying that there is probably a lack of interpersonal interaction and relationship between a virtual moderator and a participant. This problem may be approached with the help of guideline 2.1 - 2.3, by creating a VA with human characteristics, personality and offering a natural way

of interaction. Expert 1 and 3 are skeptical about the combination of a human moderator with a VA, because the human moderator might fear he could get replaced by the VA. Besides the fear of replacement, a conflict between the human moderator and the VA is considered to be possible (Expert 1).

As a VA supporting the moderation of a creative workshop is a collaborative setting with a machine as teammate, there are multiple areas of interest (Seeber et al. 2018), that are also to be considered in our case. For example, the impact of smart technologies on existing power and control relationships or to what extent robots can develop social relationships. Our expert's statements directly relate to these questions. Hence the concerns of the experts about VAs moderating a DT workshop may be viewed more generally as the problem of how collaborative settings including machines can be designed.

Alongside the limitations there are numerous potentials of a VA moderating creative processes. A VA is always accessible (Expert 1 and 3) and can be used asynchronously as well as in global-dispersed workshops (Experts 1, 2, 3 and 4). The number of participants cared of per workshop and even the number of simultaneously conducted workshops is more scalable with the help of a VA. Expert 3 mentions that participants might be more open to a VA than to a human moderator, because there is no fear of criticism. Furthermore Expert 3 supports the usage of VA in work, knowledge and creativity related tasks, as VAs are already convenient in the private environment showing great potential. Expert 3 and 5 see research playing an important role concerning a profound application of VAs in more areas. In their opinion research is needed for a controlled growth in the field of VAs.

Conclusion

The contribution of this paper is generally interesting for the design of VAs that support creative tasks or group work. We developed design guidelines for VAs supporting such processes with particular focus on the support of DT workshops. The basis for the guidelines are qualitative expert interviews combined with previous research results in the field. Through the analysis of the expert interviews we gained insights to idea generation and team-based efforts from a moderating perspective.

Three main categories of guidelines were identified in respect to our research question. The first category *general conditions* generally applies to settings in which VAs may support creativity. The second category *characteristics of a VA* contains general design requirements of a VA. While most of the components are also interesting for the design of DT-independent VAs, there are some moderation specific characteristics like *authority* or workshop specific tasks like *visualize, guide and support methods*. The third category *tasks of the VA as moderation support* is about information, methods and tasks to support DT that the interviewees frequently use in DT workshops. Methods used in a DT Workshop are not necessarily exclusive to DT or might even come from other creative processes originally. Thus, we suspect that implementing functions providing information and guidance on creativity and visualization techniques, may also provide value in other workshop or even collaborative settings. This applies not only to methods, but also to rules and desired mindsets that foster radically new ideas employed in other creative processes. This is why we suggest to also consider the VA components for other creative processes, after adapting them to the respective rules and conditions of the workshop.

Our focus was not on developing a VA itself, but rather on determining which aspects require attention during the development of a VA. Thus, our guidelines are of high interest for future work in the field of VAs that fulfill more profound and cognitively demanding tasks involving creativity and empathy. As these tasks are challenging even for experienced moderators, the requirements most likely exceed even the latest technological successes in collaborative settings. Nevertheless, evaluation and pursuit of this endeavor is appealing due to the relevance of innovation for economic growth and the cost attached to dedicated time slots for innovation with external experts and no certain outcome.

As shown VAs offer great potential, which should be exploited in smart working environments. Our proposed design guidelines are useful for practitioners who develop VAs supporting creative workshops, but they may also be applicable for other context-sensitive settings with high cognitive workload. By applying the guidelines to an actual development of a VA in future work, they can be evaluated and refined.

References

Adam, S., and Trapp, M. 2015. "Success Factors for Creativity Workshop in RE.," in *REFSQ Workshops*, pp. 54–61.

Balakrishnan, K., and Honavar, V. 2001. "Evolutionary and Neural Synthesis of Intelligent Agents," *Advances in the Evolutionary Synthesis of Intelligent Agents*, pp. 1–27.

Bell, B., and Kozlowski, S. 2008. "Active Learning: Effects of Core Training Design Elements on Self-Regulatory Processes, Learning, and Adaptability," *Articles and Chapters*. (https://digitalcommons.ilr.cornell.edu/articles/410).

Bellet, P. S., and Maloney, M. J. 1991. "The Importance of Empathy as an Interviewing Skill in Medicine," *JAMA* (266:13), pp. 1831–1832. (https://doi.org/10.1001/jama.1991.03470130111039).

Bogner, A., Littig, B., and Menz, W. (eds.). 2009. *Experteninterviews: Theorien, Methoden, Anwendungsfelder*, (3rd ed.), VS Verlag für Sozialwissenschaften.

Bostrom, N., and Yudkowsky, E. 2014. "The Ethics of Artificial Intelligence," *The Cambridge Handbook of Artificial Intelligence*, pp. 316–334.

Brem, A., and Spoedt, H. 2017. "Same Same but Different: Perspectives on Creativity Workshops by Design and Business," *IEEE Engineering Management Review* (45:1), pp. 27–31. (https://doi.org/10.1109/EMR.2017.2667143).

Brenner, W., Uebernickel, F., and Abrell, T. 2016. "Design Thinking as Mindset, Process, and Toolbox," in *Design Thinking for Innovation*, Springer, Cham, pp. 3–21. (https://doi.org/10.1007/978-3-319-26100-3_1).

Brereton, M., and McGarry, B. 2000. "An Observational Study of How Objects Support Engineering Design Thinking and Communication: Implications for the Design of Tangible Media," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '00, New York, NY, USA: ACM, pp. 217–224. (https://doi.org/10.1145/332040.332434).

Brown, T. 2008. "Design Thinking," *Harvard Business Review* (Harvard Business Review June 2008), pp. 84–96.

Brown, T. 2009. *Change by Design: How Design Thinking Transforms Organizations and Inspires Innovation*, HarperCollins.

Buchanan, R. 1992. "Wicked Problems in Design Thinking," *Design Issues* (8:2), pp. 5–21. (https://doi.org/10.2307/1511637).

Carlgren, L., Elmquist, M., and Rauth, I. 2014. "Design Thinking: Exploring Values and Effects from an Innovation Capability Perspective," *Design Journal* (17:3), pp. 403–424. (https://doi.org/10.2752/175630614X13982745783000).

Chasanidou, D., Gasparini, A. A., and Lee, E. 2015. "Design Thinking Methods and Tools for Innovation," in *Design, User Experience, and Usability: Design Discourse* (Vol. 9186), A. Marcus (ed.), Cham: Springer International Publishing, pp. 12–23. (https://doi.org/10.1007/978-3-319-20886-2_2).

Cohen, M. H., Cohen, M. H., Giangola, J. P., and Balogh, J. 2004. *Voice User Interface Design*, Addison-Wesley Professional.

Connoly, T., Routhieaux, R. L., and Schnerider, S. K. 1993. "On the Effectiveness of Group Brainstorming Test of One Underlying Cognitive Mechanism," *Small Group Research* (24:4), pp. 490–503.

Dunne, D., and Martin, R. 2006. "Design Thinking and How It Will Change Management Education: An Interview and Discussion," *Academy of Management Learning & Education* (5:4), pp. 512–523. (https://doi.org/10.5465/AMLE.2006.23473212).

Forés, B., and Camisón, C. 2016. "Does Incremental and Radical Innovation Performance Depend on Different Types of Knowledge Accumulation Capabilities and Organizational Size?," *Journal of Business Research* (69:2), pp. 831–848.

Friedman, P. G. 1989. "Upstream Facilitation: A Proactive Approach to Managing Problem-Solving Groups," *Management Communication Quarterly* (3:1), pp. 33–50. (https://doi.org/10.1177/0893318989003001003).

Gabriel, A., Monticolo, D., Camargo, M., and Bourgault, M. 2015. *Multi-Agent System to Support Creative Workshop*, IEEE, November, pp. 693–697. (https://doi.org/10.1109/SITIS.2015.114).

Geschka, H. 1986. "From Experience: Creativity Workshops in Product Innovation," *Journal of Product Innovation Management* (3:1), pp. 48–56.

Gregor, S., and Hevner, A. R. 2013. "Positioning and Presenting Design Science Research for Maximum Impact.," *MIS Quarterly* (37:2), pp. 337–355.

Grots, A., and Pratschke, M. 2009. "Design Thinking — Kreativität als Methode," *Marketing Review St. Gallen* (26:2), pp. 18–23. (https://doi.org/10.1007/s11621-009-0027-4).

Guest, G., and MacQueen, K. M. 2008. *Handbook for Team-Based Qualitative Research*, Rowman Altamira.

Guzman, A. 2017. *Making AI Safe for Humans: A Conversation with Siri*, Routledge, pp. 85–101. (https://doi.org/10.4324/9781315637228-11).

Hassenzahl, M., Wiklund-Engblom, A., Bengs, A., Hägglund, S., and Diefenbach, S. 2015. "Experience-Oriented and Product-Oriented Evaluation: Psychological Need Fulfillment, Positive Affect, and Product Perception," *International Journal of Human-Computer Interaction* (31:8), pp. 530–544.

Heiman, B., and Burnett, W. 2007. *The Role of Design Thinking in Firms and Management Education*.

Helpman, E., Itskhoki, O., and Redding, S. 2010. "Inequality and Unemployment in a Global Economy," *Econometrica* (78:4), pp. 1239–1283. (https://doi.org/10.3982/ECTA8640).

Hennessey, B. A., and Amabile, T. M. 2010. "Creativity," *Annual Review of Psychology* (61), pp. 569–598. (https://doi.org/10.1146/annurev.psych.093008.100416).

Hevner, A., March, S. T., Park, J., and Ram, S. 2004. "Design Science in Information Systems Research," *MIS Quarterly* (28:1), pp. 75–105.

Kung, H., and Schmid, L. 2015. "Innovation, Growth, and Asset Prices," *The Journal of Finance* (70:3), pp. 1001–1037. (https://doi.org/10.1111/jofi.12241).

Lempiala, T. 2010. "Barriers and Obstructive Practices for Out-of-the-Box Creativity in Groups," *International Journal of Product Development* (11:3–4), pp. 220–240.

Martin, R. L. 2009. *The Design of Business: Why Design Thinking Is the next Competitive Advantage*, Boston, Mass.: Harvard Business Press.

McTear, M., Callejas, Z., and Griol, D. 2016. "The Conversational Interface," *New York: Springer* (10), pp. 978–3.

Meuser, M., and Nagel, U. 2009. "Das Experteninterview — konzeptionelle Grundlagen und methodische Anlage," in *Methoden der vergleichenden Politik- und Sozialwissenschaft*, VS Verlag für Sozialwissenschaften, pp. 465–479. (https://doi.org/10.1007/978-3-531-91826-6_23).

Miles, M. B., and Huberman, A. M. 1994. *Qualitative Data Analysis: An Expanded Sourcebook*, sage. (https://books.google.de/books?hl=de&lr=&id=U4lU_-wJ5QEC&oi=fnd&pg=PR12&dq=miles+huberman+1994&ots=kD-G2DRYZR&sig=ypm91TAn3a0PnSN-Jn4qQNkGHMQ).

Miles, M. B., Huberman, A. M., and Saldana, J. 2014. "Qualitative Data Analysis: A Method Sourcebook," *CA, US: Sage Publications*.

Morehen, J., Wrigley, C., and Wright, N. 2013. "Teaching Design Thinking and Design Led Innovation to Non-Designers: A Tertiary Facilitator Multidisciplinary Study," in *Design Management Symposium (TIDMS), 2013 IEEE Tsinghua International*, IEEE, pp. 55–63.

Nielsen, J. 1995. "10 Usability Heuristics for User Interface Design," *Nielsen Norman Group* (1:1).

Nielsen, M. F. 2012. "Using Artifacts in Brainstorming Sessions to Secure Participation and Decouple Sequentiality," *Discourse Studies* (14:1), pp. 87–109.

Nijstad, B. A., and Stroebe, W. 2006. "How the Group Affects the Mind: A Cognitive Model of Idea Generation in Groups," *Personality and Social Psychology Review* (10:3), pp. 186–213. (https://doi.org/10.1207/s15327957pspr1003_1).

Paulus, P. 2000. "Groups, Teams, and Creativity: The Creative Potential of Idea-Generating Groups," *Applied Psychology* (49:2), pp. 237–262. (https://doi.org/10.1111/1464-0597.00013).

Paulus, P., and Nijstad, B. A. 2009. *Group Creativity: Innovation through Collaboration*, New York: Oxford University Press.

Pearl, C. 2016. *Designing Voice User Interfaces: Principles of Conversational Experiences*.

Saldaña, J. 2015. *The Coding Manual for Qualitative Researchers*, SAGE.

Santanen, E. L., Briggs, R. O., and Vreede, G. 2000. "The Cognitive Network Model of Creativity: A New Causal Model of Creativity and a New Brainstorming Technique.," *Proceedings of the 33rd Annual Hawaii International Conference on System Sciences* (vol. 1), 10 Pp.

Santanen, E. L., Briggs, R. O., and Vreede, G. 2004. "Causal Relationships in Creative Problem Solving: Comparing Facilitation Interventions for Ideation," *Journal of Management Information Systems* (20), pp. 167–198.

Seeber, I., Bittner, E., Briggs, R. O., De Vreede, G.-J., De Vreede, T., Druckenmiller, D., Maier, R., Merz, A. B., Oeste-Reiß, S., Randrup, N., Schwabe, G., and Söllner, M. 2018. *Machines as Teammates: A Collaboration Research Agenda*, presented at the Hawaii International Conference on System Sciences (HICSS), Waikoloa, HI, USA. (https://www.alexandria.unisg.ch/251713/).

Siemon, D., Eckardt, L., and Robra-Bissantz, S. 2015. "Tracking Down the Negative Group Creativity Effects with the Help of an Artificial Intelligence-Like Support System," in *System Sciences (HICSS), 2015 48th Hawaii International Conference On*, IEEE, pp. 236–243. (http://ieeexplore.ieee.org/abstract/document/7069685/).

Simon, H. A. 1969. *The Sciences of the Artificial*, [Cambridge, M.I.T. Press.

Skalski, P., and Tamborini, R. 2007. "The Role of Social Presence in Interactive Agent-Based Persuasion," *Media Psychology* (10:3), pp. 385–413.

Somech, A., and Drach-Zahavy, A. 2013. "Translating Team Creativity to Innovation Implementation: The Role of Team Composition and Climate for Innovation," *Journal of Management* (39:3), pp. 684–708. (https://doi.org/10.1177/0149206310394187).

Steinert, Y. 1992. "Twelve Tips for Conducting Effective Workshops," *Medical Teacher* (14:2–3), pp. 127–131. (https://doi.org/10.3109/01421599209079478).

Stickdorn, M., and Schneider, J. 2010. *This Is Service Design Thinking: Basics, Tools, Cases*, Consortium Book Sales & Dist.

Strohmann, T., Siemon, D., and Robra-Bissantz, S. 2017. "BrAInstorm: Intelligent Assistance in Group Idea Generation," in *International Conference on Design Science Research in Information Systems*, Springer, pp. 457–461.

Weerawardena, J., and Mavondo, F. T. 2011. "Capabilities, Innovation and Competitive Advantage," *Industrial Marketing Management* (40:8), Capabilities, Innovation and Competitive Advantage, pp. 1220–1223. (https://doi.org/10.1016/j.indmarman.2011.10.012).

Zhao, S. 2006. "Humanoid Social Robots as a Medium of Communication," *New Media & Society* (8:3), pp. 401–419.

Paper 9: On the Ability of Virtual Agents to decrease Cognitive Load – An experimental study.

Fact sheet

Type	Journal article
Authors	Brachten, Florian; Brünker, Felix; Frick, Nicholas; Ross, Björn; Stieglitz, Stefan
Year	2020
Venue	*Information Systems and e- Business Management (ISEB)*
Full citation	Brachten, F., Brünker, F., Frick, N., Ross, B. & Stieglitz, S. (2020). On the Ability of Virtual Agents to decrease Cognitive Load – An experimental study. *Information Systems and e-Business Management* (doi: 10.1007/s10257-020-00471-7).

Information Systems and e-Business Management
https://doi.org/10.1007/s10257-020-00471-7

ORIGINAL ARTICLE

On the ability of virtual agents to decrease cognitive load: an experimental study

Florian Brachten[1] · Felix Brünker[1] · Nicholas R. J. Frick[1] · Björn Ross[1] · Stefan Stieglitz[1]

Received: 28 October 2019 / Revised: 22 April 2020 / Accepted: 6 June 2020

Abstract

When attempting to solve a problem, humans call upon cognitive resources. These resources are limited, and the degree of their utilisation is described as cognitive load. While the number of parameters to be taken into account and to be processed by modern-day knowledge workers increases, their cognitive resources do not. Research shows that too high a load can increase stress and failure rates and decrease the work satisfaction and performance of employees. It is thus in the interest of organisations to reduce the cognitive load of their employees and keep it at a moderate level. One way to achieve this may be the application of virtual assistants (VAs), software programs, that can be addressed via voice or text commands and respond to the users' input. This study uses a laboratory experiment with N = 91 participants comparing two groups in their ability to solve a task. One group was able to make use of a VA while the other could not. Besides task performance, the cognitive load of the participants was measured. Results show that (a) cognitive load is negatively related to task performance, (b) the group using the VA performed better at the task and (c) the group using the VA had a lower cognitive load. These findings show that VAs are a viable way to support employees and can increase their performance. It adds to the growing field of IS research on VAs by expanding the field for the concept of cognitive load.

Keywords Cognitive load · Virtual assistants · Chatbots · Conversational agents · Task performance · Perceived workload · NASA-TLX

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s10257-020-00471-7) contains supplementary material, which is available to authorized users.

✉ Florian Brachten
florian.brachten@uni-due.de

[1] Department of Computer Science and Applied Cognitive Science, Research Group Professional Communication in Electronic Media/Social Media, University of Duisburg-Essen, Duisburg, Germany

Published online: 19 June 2020

1 Introduction

The working memory plays an important role in learning and processing of information. As its capacity is limited so is the amount of information that can be processed at the same time. The amount of working memory used during this processing is called cognitive load (Sweller 1988). The concept of cognitive load originates in behavioural psychology and the realm of learning but has since been transferred to research on other disciplines, amongst others to Information Systems (IS) (e.g. Hu et al. 2017). Studies have also transferred the concept to the working context and could, for example, show that work-related well-being was decreased when the cognitive load was too high (Pace et al. 2019). Even more severely, findings also indicate that task performance of employees seem to suffer under too high a load (Altaf and Awan 2011). Taken together with the digitalisation that takes place in almost all industries and poses challenges to both employers and employees who need to adapt to new and complex processes (Matt et al. 2015) it is increasingly important to control the cognitive load experienced by employees in order to ensure their well-being and in turn the performance of the employer.

However, the digitalisation does not only pose challenges but also offers opportunities to reduce the workload, for example in the form of virtual assistants (VAs). These computer-based support systems are projected to be used by at least a quarter of digital workers within the next two years (Maedche et al. 2019). It follows that research has looked at the application of this technology in the working context from several angles, e.g. its potential for support in customer service (Cui et al. 2017) or internal communication (Stöckli et al. 2018). Furthermore, introducing VAs in organisations aims at reducing the workload of employees by assisting in the execution of work-related tasks (Norman 2017). Research on whether support through computer-based systems may be able to reduce the workload yields ambiguous results (Moreno et al. 2001). On the one hand, studies show that learning is promoted (e.g. Moreno et al. 2001; Schmuntzsch et al. 2012) and superior results are achievable (Mechling et al. 2010). On the other hand, the learning process is not always facilitated (Schnotz and Rasch 2005) and work performances can be impacted negatively (Chandler and Sweller 1991). However, to our knowledge, the potential of VAs to reduce cognitive load, or the question whether they even interfere when performing tasks, has not yet been addressed sufficiently. If these systems are actually to become widely-used within a few years and they furthermore hold the potential to reduce the workload, they could be a feasible way to relieve employees, thus supporting their well-being and performance, and they could therefore be of great value to enterprises. However, to make informed assumptions on this topic, more research is needed as it is currently inconclusive, which is why more clarification is necessary. The aim of this paper is thus to shed light on these aspects, especially on the ambiguity regarding the in- or decrease in cognitive load through technology and thus to answer the following questions:

RQ1 To what extent do virtual assistants influence the perceived workload during the solution of a task?

RQ2 How do participants supported by virtual assistants compare to those without support regarding their performance at a task?.

To address this shortcoming, we conducted an experiment with N = 91 participants in two groups that had to solve a work-related task and where one group was supported by a VA. We then measured and compared the cognitive load of the participants in the two groups and their actual task performance. The paper presents the findings and aims to shed light on the potential of VAs to positively influence the cognitive load of their users. It first presents current literature on cognitive load and VAs before detailing the chosen methodological approach. Afterwards the results are described and discussed, followed by the conclusion as to what extent VAs are feasible to influence the cognitive load of employees.

2 Theoretical background

2.1 Cognitive load

Cognitive load theory explains how factors such as task difficulty and people's available mental resources influence their success in learning to solve problems effectively (Sweller 1988). It originates in educational psychology. An underlying assumption is that a learner has limited cognitive capacity that he or she can make use of when attempting to solve a problem. Effective learning takes place when the learner develops the ability to recognise that the task belongs to a category of problems, and knows which steps are normally required to solve such problems. This process is referred to as schema acquisition (Sweller 1988). It can only take place when enough cognitive capacity is available for categorising and systematising knowledge (germane cognitive load). If it is taken up by the task's inherent cognitive load demands (intrinsic load) and by additional cognitive load that is unnecessarily imposed, for example, by poor instructional design (extraneous cognitive load), then learning cannot take place (Paas et al. 2003). This implies that, counterintuitively, if all of one's cognitive capacity is devoted to achieving a specific goal set by the instructor, learning can actually suffer, and a goal-free approach might be better (Sweller 1988). The psychological resistance to stress or difficult situations, known as resilience (Neyer and Asendorpf 2017), might further impact the cognitive load in addition to task difficulty and people's available mental resources. The term resilience is defined as *"positive psychological capacity to rebound, to 'bounce back' from adversity, uncertainty, conflict, failure or even positive change, progress and increased responsibility"* (Luthans 2002). Resilience is related to satisfaction or commitment of employees at the workplace (Youssef and Luthans 2007) and changes when known behaviour and common

procedures vary. Cognitive load theory has obvious implications for instructional design. An overloaded or underloaded learner will acquire problem-solving skills less effectively. An appropriate instructional procedure should therefore encourage learners to use their cognitive resources in a way that furthers learning, while at the same time avoid demanding cognitive resources unnecessarily.

Cognitive load theory also has implications for management. In an organisational setting, the concept of learning how to solve a problem by attempting to solve it is commonplace, if not as a result of deliberate instructional design, then as a practical consequence of business demands. In an age of frequent technological and organisational change, Galy et al. (2012) argue, managing workload is an important part of ensuring employees' wellbeing and safety. Exorbitant cognitive load can have negative implications on business decisions: for example, managers under high cognitive load may face difficulties in evaluating job candidates appropriately (Nordstrom et al. 1996).

2.2 Virtual assistants in organisations

The deployment of virtual assistants in organisations seems reasonable for managing employees' workload, facilitating tasks and improving business decisions. Various synonyms for VAs exist which are used interchangeably (Luger and Sellen 2016). Terms that can be found in research and practice are, for example, *voice assistants* (Diao et al. 2014; Alepis and Patsakis 2017; Hoy 2018), *personal assistants* (Moorthy and Vu 2015; Sangyeal and Heetae 2018), *cognitive assistants* (Siddike and Kohda 2018; Siddike et al. 2018) or *conversational agents* (Saffarizadeh et al. 2017). Similar to the variety of words, there is no consensus on a precise definition. Researchers and practitioners explain VAs from different perspectives such as their primary mode of communication or their main purpose (Gnewuch et al. 2017) as well as by their tasks and system characteristics (Strohmann et al. 2018). The explicit classification of VAs in one of these categories is simply not possible due to overlaps. Text-based VAs might use speech-to-text modules to convert human language into text (Gnewuch et al. 2017). VAs might also be further developed and customised to adapt features to individual needs or specific tasks (Chung et al. 2017). However, VAs can generally be described as systems interacting with users by simulating the behaviour of human beings and using natural language (Luger and Sellen 2016; McTear et al. 2016; Diederich et al. 2019) to assist in the execution of work-related tasks or even have them fulfilled entirely (Norman 2017). In the current context, the definition used by Stieglitz et al. (2018) to define the term VA seems to fit best for our purpose: "software programs that can be addressed via voice or text and that can respond to the user's input (i.e. assist) with sought-after information" (p. 3).

Research has recently gained an interest in the interaction with VAs (Gnewuch et al. 2017) since building systems with the help of artificial intelligence and machine learning algorithms has become more practical to assist users in a wide variety of tasks (Knijnenburg and Willemsen 2016). VAs are particularly helpful in tackling repetitive tasks that require the fast retrieval and processing of digital data as well as the understanding of complex interdependencies (Dellermann et al.

2019). By tailoring systems to the users' needs, better assistance and added value can be generated (Maedche et al. 2016). Due to the fact that numerous benefits are generated, especially regarding competitive advantage, organisations are heavily investing in VAs (Schuetzler et al. 2018). Applied in organisations, for example in banking, insurance or retail, VAs aim at the prospect of generating additional revenue or cost savings (Quarteroni 2018) and positively influencing the customer's satisfaction (Verhagen et al. 2014). VAs can be used for the direct interaction with customers. When assisting while shopping online, VAs provide advice to find suitable products and thus reduce information overload (Benbasat and Wang 2005; Qiu and Benbasat 2009). Moreover, users are supported when having inquiries regarding the company's services (Quarteroni 2018). In addition, VAs also have the potential to enhance processes within organisations. Systems are utilised in human resource departments to facilitate the onboarding process of new employees by providing a question-and-answer assistant (Shamekhi et al. 2018). Further, VAs are applied in customer service (e.g. (e.g. Gnewuch et al. 2017; Hu et al. 2018) to reduce the workload of call centre agents. By supporting the handling of customer enquiries with VAs (McTear et al. 2016), a solution for users can be proposed immediately (Frick et al. 2019) as requests can be handled without additional overhead (Stieglitz et al. 2018).

2.3 Virtual assistants and cognitive load

Research has already taken several attempts to validate if the cognitive load can be reduced through the deployment of various technologies. Moreno et al. (2001) showed that students interacting with an animated pedagogical agent via natural language outperform students not using an agent when learning. Another study could show that VAs, embodied by an animated character, help to focus on relevant information and facilitate learning thus supporting users performing physical tasks (Schmuntzsch et al. 2012). However, a VA could also interfere with successful learning. Interacting with it requires the participant to exert cognitive resources. Cognitive load theory has long recognised that this additional, extraneous cognitive load may eliminate the benefit from the additional instruction (Tarmizi and Sweller 1988). Seemingly useful material can negatively impact performance if it is not essential to solving the task (Chandler and Sweller 1991). Schnotz and Rasch (2005) found that facilitating learning is not always beneficial as users are prevented from performing relevant cognitive processes on their own. Lohse et al. (2014) examined robot gestures and report that a higher human–robot interaction increases user performance and decreased cognitive load for difficult tasks but not for easy tasks. Regarding virtual agents, Moreno et al. (2001) make a similar argument, and refer to the constructivist hypothesis (that agents help learning) and the interference hypothesis (that they hinder it).

The application of VAs in organisations seems beneficial to facilitate internal processes and to gain competitive advantage in that it supports workers in better completing their tasks (Morana et al. 2017). As studies indicate that increased cognitive load at the workplace hinders employees from reaching their full potential

(Altaf and Awan 2011), it should be in the interest of organisations to keep this load at a moderate level. The utilisation of VAs aims at doing exactly that: reducing the cognitive load when enhancing humans in work-related tasks for further performance improvements. This might create significant benefits for the applying organisation itself and further for its customers. Studies have already shown that, under certain circumstances, cognitive load can be reduced through the use of technology, increasing the user performance. For example, Mechling et al. (2010) could show that groups instructed by a digital assistant showed better results than groups without that support. Likewise, cognitive load might be enlarged when dealing with additional instructions or different tasks. Until now there has not yet been any research giving evidence if VAs are able to reduce the cognitive load or if they even interfere when performing tasks. To test this, we conducted an experiment with two groups which had to solve a task and where the experimental group could use a virtual assistant to solve the task while the control group could not. In the following section we will describe the structure of the experiment, the task that was to be solved as well as the measures that were collected.

3 Method

3.1 Participants

The experiment was conducted at a German University between 28 May and 18 June 2019. The university's students were invited to participate on a voluntary basis. In this timespan, 91 people participated in the study. We then randomly assigned the participants into two groups, resulting in a well-balanced sample of 46 participants in the control group without a virtual assistant and 45 in the experimental group using a virtual assistant. Overall, 54.9% of the participants were female (N = 50), and their age ranged from 18 to 31 (M = 22.01, SD = 3.02), indicating a rather young sample. Furthermore, 80% of the participants had passed their A-levels while 14% held a Bachelor's degree. Together with the young age and in accordance with the mode of acquisition of the sample this shows a typical undergraduate student sample.

3.2 Materials

3.2.1 NASA task load index (NASA-TLX)

Concepts related to cognitive load are frequently measured using self-report rating scales (Paas et al. 2003). This approach assumes that learners are able to report the amount of mental effort that they experienced while attempting to solve a task. It is worth noting that self-report rating scales do not typically distinguish between the three types of cognitive load (intrinsic, extraneous, germane) but rather measure the overall load experienced.

A commonly used scale to quantify the perceived workload of a participant is the NASA Task Load Index (Galy et al. 2012). The National Aeronautics and Space Administration (NASA) developed the NASA-TLX in order to measure the perceived workload of a task (Hart and Staveland 1988). This measurement was successfully used in several contexts such as in both laboratory and field studies (Rubio et al. 2004; Noyes and Bruneau 2007; Cao et al. 2009). The index itself contains six subjective subscales forming the NASA-TLX score: (1) Mental Demand, (2) Physical Demand, (3) Temporal Demand, (4) Performance, (5) Effort, and (6) Frustration. These clusters of variables were chosen to cover the "range of opinions and apply the same label to very different aspects of their experience" (Hart 2006, p. 904). Due to the subjective experience of conducting a specific task, the NASA-TLX was developed to consider the perception of a variety of activities such as simple laboratory task or flying an aircraft. While (1) describes how much mental and perceptual activity was required, (2) shows the perceived amount of required physical activity. Besides the perceived mental and physical efforts, the NASA-TLX also covers the perception of time pressure (3) during a task. Furthermore, the subscales (4) to (6) ask about the perception of the results of the given tasks. Therefore, (4) describes the personal performance perception – i. e. the perceived success reaching the given goals of the tasks and (5) asks to what extent the participants had to work to reach the achieved level of performance. As people sometimes feel frustrated when a given task is perceived as too difficult, subscale (6) asks the participants about the level of frustration during the task (Hart 2006). In our experiment, all subscales had high reliability (Cronbach's $\alpha = 0.89$).

3.2.2 Resilience scale (RS-11)

According to the appraisal theory, stress emerges when a task at hand exceeds one's own resources and abilities (Smith et al. 2011). Following, an increasing level of stress might impact the participant's task performance as well as the perception of the work and its outcome. In order to avoid undetected distortion towards the task performance, we consider the psychological resistance to stress or difficult situations, known as resilience (Neyer and Asendorpf 2017). We use the Resilience Scale (RS-11) as a short scale for assessing the resilience of a human (Schumacher et al. 2005). The RS-11 is a self-report scale containing eleven items which are divided into two sub-scales: (1) personal competence and (2) acceptance of the self and life. The subscales had a high reliability, all Cronbach's $\alpha = 0.90$.

3.2.3 Virtual assistant

In order to investigate the impact of a text-based VA on decreasing the cognitive load during task-solving, we made use of Google's cloud platform *DialogFlow.*[1] This platform is widely used for developing natural and rich conversational experiences based on Google's machine learning (Canonico and De Russis 2018).

[1] https://dialogflow.com/.

Furthermore, the implementation is based on four general concepts (Muñoz et al. 2018). First, *Agents* transform natural user language into actionable data when a user input matches one of the intents. Second, *Intents* represent a mapping between what the user says and what action is taken. Third, *Entities* represent concepts and serve as a tool for extracting parameter values from natural language inputs. Finally, *Contexts* are designed for passing on information from previous conversations or external sources. To reduce the degree of complexity caused by the interaction with the VA, we focused on establishing a disembodied VA with a messaging-based interface (Araujo 2018).

As VAs exhibit social and conversational dialogue (Hung et al. 2009), our VA is implemented to make a simple conversation at the beginning of the interaction. Participants can interact with the VA via a web-based interface, similar to contemporary instant messengers such as Telegram or WhatsApp, using a keyboard and a computer screen. This interaction could be a request for the participant's name and feelings. Furthermore, the applied VA is text-based to avoid additional influential factors which may evolve by voice interactions or embodied avatars. Figure 1 shows a translated example of a dialogue with the VA.

To support the participants during the task, the assistant simulates intelligence by selecting a prefabricated answer based on distinct keywords used in the participant's input. We defined 25 Intents to match the user input. The intents belonged into roughly 3 groups: Introduction, Tutorial and Task Support. The Intents in the Introduction Group mostly revolved around welcoming the users, asking for their well-being and readiness to start the task. The tutorial intents were designed to increase the users' familiarity with the VA and the capabilities of the VA. Most of the intents revolved around Task support where users could ask for help solving the task, for example by asking what certain parameters meant or how they were calculated. We also used the standard „sys.given-name " entity provided by DialogFlow as an Entity. The VA's feedback includes a question-answering component (Morrissey and Kirakowski 2013; Lamontagne et al. 2014) that can be queried by the user to gain information, support and instruction about the specific task. In this context, the VA only provides helpful hints which support the participants solving the task. However, the VA does not deliver the actual solution to the current task.

3.2.4 Task performance and pre-study

Task performance was measured with a score ranging from 0 to 28 that captures how well participants did at a critical path method (CPM) task. A higher value represents a better performance in the execution of the task. The goal of this task was to use this method to plan a research project for the market research unit of a large organisation.

The task was determined in a pre-study to ensure that it is sophisticated and involves a potentially high perceived workload in the experiment. The sample of 10 participants (6 female, 4 male) consists of randomly selected students at the University. In this context, a good fitting task challenges the participants on decent level, and therefore causes an increased cognitive load score. A task which overwhelms

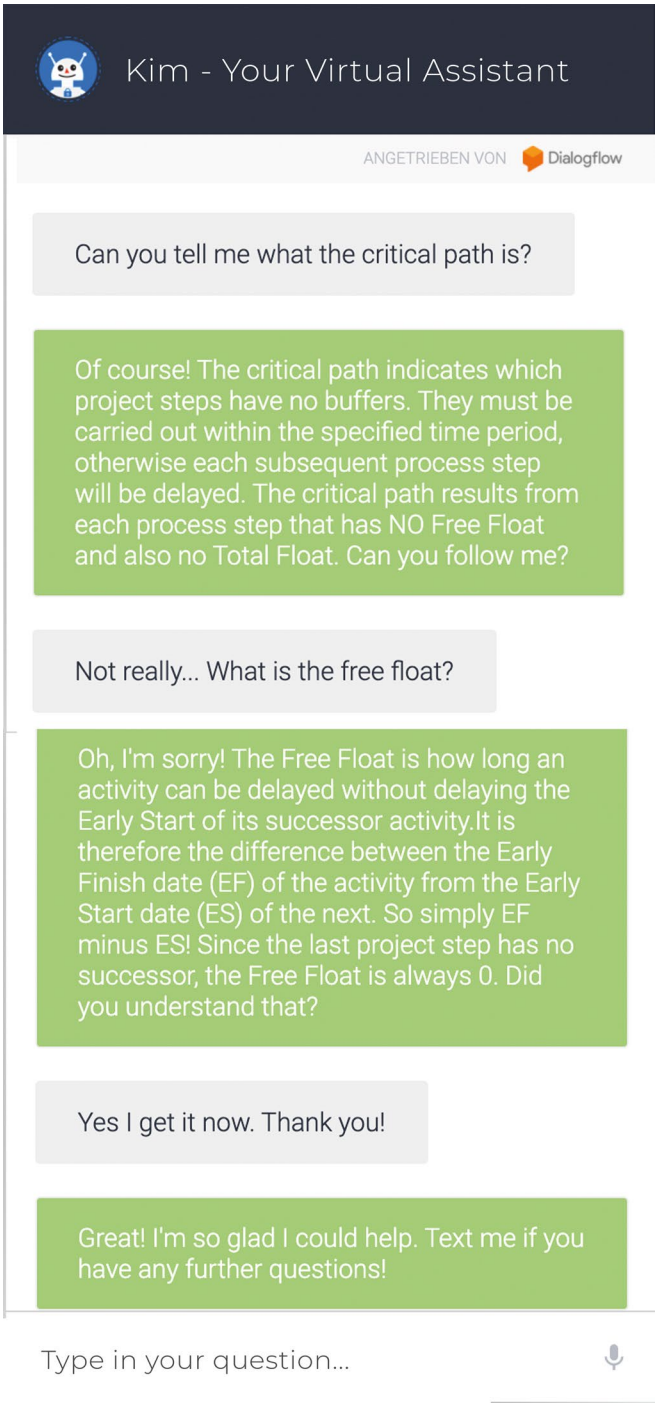

Fig. 1 Example of a dialogue with the Virtual Assistant

the participants may prevent sustained learning effects due to less available cognitive resources (Paas et al. 2003). To this end, a text-based task (TBT) and CPM were compared. On the one hand, the TBT required the participants to read three texts about medieval ages, a topic which does not rely on previous knowledge of the participants. On the other hand, the CPM was implemented with a scenario that puts the participants in a working context. In detail, the participants had to organise a marketing study using the CPM. The time limit for each of the tasks was 10 min.

Each task was given to five participants and the perceived workload was measured by the NASA-TLX. The age ranged from 22 to 31 (M = 25). On average, participants given the CPM task engaged in higher NASA-TLX scores (M = 12.5, SD = 3.85) than the TBT group (M = 6.36, SD = 4.06). This difference of 6.13 was significant (95% CI [0.35, 11.91], t(8) = 2.44, $p = 0.040$. Furthermore, it represents a large-sized effect, $d = 0.98$. Following, the CPM task has the potential to increase the cognitive load of the participants in a more effective way than the TBT does. Thus, due to its better potential to benefit from the use of a virtual assistant, the CPM was chosen for the main study.

3.3 Procedure

In order to investigate the influence of a VA on the perceived workload of a participant, the experiment used a between-subjects design. The independent variables were the resilience score (RS-11) and the usage of a VA (group variable) whereas the dependent variables were the perceived workload (NASA-TLX), the task-score as well as the time to finish the task. Analyses were conducted using the software SPSS Statistics (Version 25) and Jamovi (1.0.2.0).

The main study was conducted as a laboratory experiment at a German University in German language. A laboratory experiment was chosen to better control the surroundings, to ensure that the task performance was measured correctly and to ensure a steady and even experience with the virtual assistant. Furthermore, the investigators were present to assist the participants with questions should those arise. However, their assistance was not utilised by any of the subjects.

The participants were welcomed by the investigator and introduced to the study. They were then led to a computer to begin with the first questionnaire. First, the participants were presented with the RS-11 questionnaire to retrieve the resilience score.

Afterwards they were presented with an introduction to the CPM followed by an example. After reading through the briefing, participants were instructed to contact the investigator for the material needed. The goal was to use the CPM to plan a research project for the market research unit of a big organisation. Participants were given a list with unordered process steps (such as "literature research", "conducting the study" or "develop methodology"), the respective duration for each step as well as its dependencies on the other steps in the process. They were also handed an empty template for a CPM to fill out with the according parameters. Finally, the participants were informed of a virtual folder they were allowed to use which was

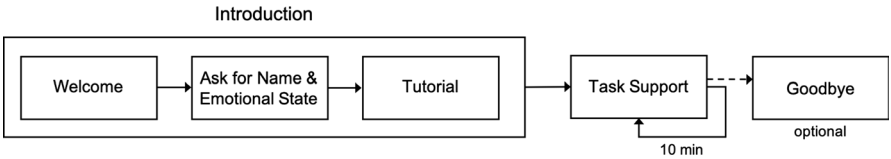

Fig. 2 Depiction of the interaction with the Virtual Assistant

located on the laboratory computer and included unordered text files explaining the CPM procedure and the calculation of the individual values.

Additionally, the participants in the experimental group were also presented with a sheet of paper which explained that they were allowed to use a text-based VA and that it was nested in a browser window in the computer. They were then explained how to use the VA properly such as using single sentences and that the VA did not have contextual knowledge. All subjects in the experimental group made use of the VA which provided the participants with the same information available in the folder to all groups but could be specifically asked for certain information, e.g. what certain parameters stood for or how they were calculated. Figure 2 depicts the steps a conversation with the VA consisted of. Except for the availability of the VA, the participants in the control and experimental group were presented the exact same task. Also, all participants had access to the same information for solving the task with the only difference that subjects in the control group could access the data via browsing through virtual folders on the computer whereas subjects in the experimental group could specifically access the information via dialogue with the VA.

Participants then had a time limit of 10 min to complete the task after which they had to stop solving the task even when they had not yet completed it. They were also instructed to give notice should they be finished before the time limit had run out. In the latter cases the investigator noted the time that was needed. After the participants either completed the task or the time ran out, they were re-referred to the computer to complete the remainder of the survey.

Following the task, the participants were presented with the NASA-TLX to assess their perceived workload immediately after solving the task. They were then asked whether they had already been familiar with the technique of CPM and the participants in the experimental condition were additionally asked whether they thought the support by VA was helpful.

Finally, all participants were asked for their gender, age, highest educational attainment and were debriefed, asked whether they had any further questions and then thanked for their time.[2]

[2] The concrete wording as well as the items can be taken from the supplementary material in the electronic version of this article.

4 Results

4.1 Descriptive statistics

Overall, 95% of the participants who used the virtual assistant found it helpful and on average had 14 interactions with the chatbot (where one interaction is defined as one user input followed by one chatbot reply), 94.3% of which were matched (i.e. the chatbot was able to match the input with intent). Participants scored between 0 and 28 points in the task, with M = 16.68 (SD = 7.81, Med = 17) indicating that most of the participants achieved more than half of the 28 points possible. 14% reached the maximum score of 28 points. 70% of participants used the full 10 min to complete the task, 3.3% needed less than 5 min. Only 4 participants had already used CPM.

4.2 Relation between perceived workload and task performance

To check whether the perceived workload was connected to the performance of the participants, we conducted a bivariate correlation separate for each group (i.e. with assistant and without assistant). In both groups the NASA-TLX score showed a significant negative correlation with the performance score, indicating a better performance when the cognitive load was lower. The correlation $r(46) = -0.673$, $p < 0.001$ in the control-group without an assistant was larger than in the experimental group with the assistant $r(45) = -0.462$, $p < 0.001$. To assess whether this difference was meaningful, we calculated the z-scores, which showed that the correlation between NASA-TLX and task-score was not significantly different between the two groups ($Z = 1.458$, $p = 0.072$) which indicates that, regardless of the condition the participants were in, a higher cognitive load was connected to worse performance.

4.3 Resilience and perceived workload

A t-test was conducted to investigate possible differences between the groups in the participants' resilience. To test whether the resilience of the participants influenced their perceived workload, the correlation between the two variables was calculated for each of the groups. To assess whether the strength of a possible relationship differed between the groups, a z-test was used to determine if the two correlations differed significantly.

The t-test for a difference in mean resilience score between the group with the assistant (M = 130, SD = 20.5) and the group without the assistant (M = 128, SD = 14.6) does not indicate a significant difference, $t(89) = -0.52$, $p = 0.602$, $d = -0.110$. In both groups, resilience and cognitive load correlated significantly with $r(46) = -0.354$ ($p = 0.016$) in the control group and $r(45) = -0.380$ ($p = 0.010$) in the experimental group. The correlation coefficients did not differ significantly between the groups ($Z = 0.139$, $p = 0.445$), indicating that the higher the participants' resilience, the lower the perceived workload – regardless of the group.

4.4 Performance of the groups

To assess whether performance differences between the groups with and without an assistant exist, we calculated an independent t-test with a 95% confidence interval and with the task-score as independent variable and the groups as factor (Table 1). As Levene's Test for Equality of Variances was not significant ($p = 0.767$), equal variances were assumed. On average, participants who used an assistant performed better (M = 19.76, SD = 7.36) compared to participants in the group without an assistant (M = 13.67, SD = 7.09). This difference was significant $t(89) = -4.01$, $p < 0.001$ and represents a large-sized effect ($d = 0.84$).

4.5 Difference between cognitive load of the groups

Next, to test whether participants using the assistant differed in their reported perceived workload, we conducted an independent t-test with a 95% confidence interval and with the NASA-TLX-score as independent variable and the groups as factor (Table 2). Levene's Test for Equality of Variances was not significant ($p = 0.470$). The mean of the group without the assistant was significantly ($t(89) = 3.55$, $p < 0.001$) higher (M = 10.28, SD = 4.52) than the mean of the group using an assistant (M = 7.17, SD = 3.79), indicating a higher perceived workload for the group without an assistant. The effect size was large ($d = 0.75$).

4.6 Time needed by the groups

An independent t-test with the groups as factor and the time needed to complete the task was calculated to check whether one group on average took less time

Table 1 Comparison of Task score performance between the groups

	Group	N	Mean	Median	SD	SE
Task score performance	Without assistant	46	13.7	14.0	7.36	1.09
	With assistant	45	19.8	20.0	7.09	1.06

Max. Score: 28

Table 2 Comparison of perceived workload between the groups

	Group	N	Mean	Median	SD	SE
NASA-TLX score	Without assistant	46	10.3	11.3	4.52	0.666
	With assistant	45	7.17	6.83	3.79	0.565

Max. Score: 20

Table 3 Comparison of time needed for working on the task between the groups

	Group	N	Mean	Median	SD	SE
Time (s)	Without assistant	46	586	600	58.0	8.56
	With assistant	45	518	578	107	16.0

Max. time: 600 s

Table 4 Comparison of personal performance perception between the groups

	Group	N	Mean	Median	SD	SE
Personal performance perception	Without assistant	46	10.1	8.50	6.86	1.01
	With assistant	45	7.00	5.00	5.70	0.849

Max. Score: 20

than the other (Table 3). As Levene's test for equality of variances was significant ($p<0.001$), degrees of freedom were adjusted from 89 to 67. The control group without the VA needed significantly more time (M = 586 s, SD = 107) than the group with the assistant (M = 518 s, SD = 107), $t(67.35)=3.74$, $p<0.001$. $d=0.79$ indicated a large effect size.

4.7 Difference in the personal performance perception of the groups

As subscale 4 of the NASA-TLX measured the perception of the participants own performance in completing the task it was used to assess whether this perception differed between the groups. A lower value on this scale indicates a better performance. An independent t-test with a 95% confidence interval was calculated. Levene's test for equality of variances was significant ($p=0.045$), and the degrees of freedom were reduced accordingly (from 89 to 86.74). The groups differed significantly ($t(86.74)=2.32$, $p=0.023$) with the participants in the control group without the VA having a higher mean (M = 10.1, SD = 6.86) than the participants in the experimental condition (M = 7.00, SD = 5.70) indicating that the former believed they did a worse job at completing the given task than the group using a VA believed of themselves (Table 4). The effect size was moderate ($d=0.486$).

5 Discussion

5.1 Key findings and implications

One question this paper aimed to answer was how VAs influence cognitive load during the solution of a task, as previous literature presented inconclusive findings on this matter (e.g. Lohse et al. 2014). The current findings support the notion that VAs are suitable to decrease said load, indicating that the application of VAs also has an impact on the perceived workload of its users in that it reduces this workload as

well. Prior findings were ambiguous, indicating that the application of technology supporting users may also hinder the users because of the additional effort needed to learn interacting with the supporting system (Tarmizi and Sweller 1988). In our study this was not the case, for which there are several possible explanations. VAs may be easy and intuitive to use. As the interaction with the VA applied in our study takes place in text form and natural language akin to a chat with a human, this concept may be familiar with participants, especially regarding the young age of the sample. Thus, there is no additional effort needed to first learn how to use the VA and it can solely help in supporting to reach the goal. Based on this assumption, one takeaway from the study is that a VA which is easy to use and whose usage is not connected to any extra effort is a good way to support people in fulfilling certain tasks. Then again, the VA used in the current study had a very special focus as did the participants while solving the task. It may be that specialised VAs are able to more easily help solve special goals while VAs with a broader skill set may be less effective. However, in practice it is not always feasible to provide several VAs for several tasks. On the one hand, that may drive up costs for an applying organisation; on the other hand, it may also be counterproductive in that it actually would require the users to always have to pick the right VA according to the task at hand but in turn increasing workload as additional mental resources are needed to make that decision. However, this was not the current studies' scope but needs more elaboration in future studies.

From a learning perspective, the result that the group with virtual assistant perceived a lower mental workload, performed better on the task and also perceived their performance as better means that more resources could be available for germane cognitive load, which fosters schema acquisition and thereby improves learning (Paas et al. 2003). However, this process is not automatic. Depending on the design of the materials, people who are exposed to too little cognitive load could also be less likely to learn permanently how to solve the problem. In other words, the availability of a VA might have helped them perform their tasks better but at the same time made the task too easy for them to be able to later recall how they solved it. This effect could be detrimental to their performance especially if they will not always have the assistant available. Future studies should examine retention, and employers who consider supporting their knowledge workers with VAs should keep this point in mind to avoid an undesirable over-reliance on technology.

The current study furthermore examined the influence VAs could have on task performance. Here, the findings show that the application of VAs had a positive influence on the performance while solving a task. This is in line with findings from previous literature e.g. from the learning domain which reported that students which received support by a system akin to a VA performed better than those who did not have the support (Moreno et al. 2001; Mechling et al. 2010). However, the current study demonstrated this aspect on a domain rather related to the working environment. This finding supports assumptions made by other researches (e.g. Morana et al. 2017) and shows the importance that these systems may have in improving organisational performance (through its employees) which may also lead to a higher satisfaction of the latter. Furthermore, it demonstrates, that VAs are not only beneficial for organisations when applied in contact with its customers (cf. Quarteroni

2018) but also for its own employees. However, as we discuss in the limitations and outlook section, the current study was conducted as a laboratory experiment which means that its external validity (i.e. in a practical context) has to be shown in future studies. Nevertheless, these findings lay the groundwork for further evaluation of VAs in the working context. Aspects such as the applicability over various industries and for various tasks could be examined. Also, the acceptance by the users and the applicability for different tasks or categories of tasks need to be explored.

Connected to the aspects discussed before, the current findings also show that a reduced perceived workload is beneficial in reaching a higher score on task solution. The findings thus show that it is desirable to reduce the workload in order to improve performance. In context of a working environment and regardless of the application of VAs, employers should in general strive to support knowledge workers in eliminating any distraction. Processes in the way to reach a certain goal could be examined regarding unnecessary or outsourceable steps which could then be reduced to a minimum. Especially steps connected to repetitive or overhead (i.e. bureaucratic) activities seem to hold potential to be reduced as especially the latter have been found to be negatively related to the perception of work-related well-being (Pace et al. 2019).

One important finding is also that the personal predisposition in the form of resilience plays a role in the amount of perceived workload the participants felt. The higher the resilience of the participants, the lower their perceived workload – regardless of whether they used a VA or not. This means that, besides external support, personal predispositions do also play a role in the amount of workload people report. This is an important aspect that should not be overlooked as it indicates that a person can get all the help in the world and still have a high workload which may impede the performance in solving a task. Furthermore, different people may need different amounts of support in solving a task or reaching a goal, which is important to consider when evaluating any performance-related finding in regard to perceived workload. While this aspect is not at the focus of the current paper, it shows that individual predispositions should be taken into account and considered when conducting research and interpreting results on perceived workload and in turn cognitive load.

5.2 Limitations and future research

As with all research, several limitations to the findings apply. Our findings are based on a sample consisting mostly of undergraduate students. Because of their youth and thereby assumed familiarity with communication technology, the experimental group using the VA may have had less trouble operating the VA than an average adult that may be not as affine to modern communication technology. Furthermore, our paper aimed to research cognitive load and task performance in a working context, for which one may argue that students are not as feasible as actual employees with working experience. However, as Kretzer and Maedche put it "students are suitable subjects, and students may also tend to be less biased than experienced professionals due to their general relative youth

and lack of work experience." (2019, p. 1156). Still, future studies may take this aspect into account and aim to replicate the findings on different samples that may be older or have experience in a working environment.

Furthermore, the task chosen in the current context may not completely represent daily work processes as it is rather abstract. Still, with regard to the measurability of the performance outcome, the task chosen in the current paper is, in our eyes, a good compromise as it accomplishes two things: It simulates a task in a working environment (such as a process that needs to be planned in a short time span) and it ensures measurability and applicability in a laboratory research setting, which increases the validity of the results. Nevertheless, future studies may alter the task and, for example, conduct a case study in a real-world setting. Our research thus also adds to the emerging body of work considering the "Operator 4.0" – the worker who relies on increasing automation at the workplace for increased efficiency (Romero et al. 2020). As the technology becomes more capable and people become more familiar with it, this field of research is poised to grow.

Future studies may also deeper examine the relation of personal predispositions and the effect of a heightened perceived workload. In the current paper, resilience showed that it impacted the perceived workload of the participants. First, this connection could be further examined – is this true for other samples or for other tasks? As the resilience was correlated with the perceived workload – which in turn was based on the task to solve it would be interesting to see how this relation holds up if the task is changed, e.g. to be more complex or easier. Furthermore, different predispositions could also be taken into account like the involvement in a certain topic that is investigated with the task, the personality of participants or technological affinity and the likes. This would lead to better insights into what aspects play a role in in- or decreasing perceived workload in individuals.

As the cognitive load theory originates in educational psychology it would also be interesting to conduct time-series analyses. These could show the effect learning has on the task performance. This way, the effectivity of VAs over time may be shown, e.g. to be even greater, as participants get used to working with such systems, possibly reducing cognitive load that is initially needed to adapt to the system. While the aspect of adapting to a VA didn't seem to play a role in the current study, future studies that alter the task or the VA could help shed light on this aspect. This may be especially interesting to test how the interaction with a VA evolves over its lifespan and to evaluate its long-term value, for example for organisations.

Future studies should also examine whether our findings hold true for alternate VA approaches. For example, in the current study, the response the participants in the experimental group got was instant—i.e. there was no delay between sending a question to the VA and getting feedback (in part due to the realisation via DialogFlow). However, current research indicates that artificial delays in the response time by a VA may lead to a more satisfying experience for the users (Gnewuch et al. 2018). Here, it could be interesting to see how such modulations influence task performance or the perceived workload as in situations where a user is under pressure it may be more beneficial to deliver fast answers.

In summary, there is much research to be done on the effectiveness of virtual assistants. Our study has shown that they are a viable option that is worth exploring. If they can succeed in reducing knowledge workers' cognitive load in a variety of situations, then they might be able to help make the digitisation of the workplace something to be welcomed, and perhaps even enjoyed, by all.

Acknowledgements Open Access funding provided by Projekt DEAL.

Appendix

Structure and content of questionnaire.

- Greeting of the participant
- Measurement of resilience (RS-11)
- Prompt to address the investigator regarding introduction to the task followed by the processing of the task
  - Measurement of task performance score
  - Measurement of time of processing the task
- Measurement of perceived work-load (NASA-TLX)
- Question whether CPM was known before the current study
- Question whether VA was helpful (only experimental group)
- Demographics
  - Gender
  - Age
  - Highest educational attainment
- Debriefing

References

Alepis E, Patsakis C (2017) Monkey says, monkey does: security and privacy on voice assistants. IEEE Access 5:17841–17851. https://doi.org/10.1109/ACCESS.2017.2747626

Altaf A, Awan MA (2011) Moderating affect of workplace spirituality on the relationship of job overload and job satisfaction. J Bus Ethics 104:93–99. https://doi.org/10.1007/s10551-011-0891-0

Araujo T (2018) Living up to the chatbot hype: The influence of anthropomorphic design cues and communicative agency framing on conversational agent and company perceptions. Comput Human Behav 85:183–189. https://doi.org/10.1016/j.chb.2018.03.051

Benbasat I, Wang W (2005) Trust In and Adoption of Online Recommendation Agents. J Assoc Inf Syst 6:72–101. https://doi.org/10.17705/1jais.00065

Canonico M, De Russis L (2018) A comparison and critique of natural language understanding tools. In: CLOUD COMPUTING 2018: the ninth international conference on cloud computing, GRIDs, and virtualization, pp 110–115

Cao A, Chintamani KK, Pandya AK, Ellis RD (2009) NASA TLX: Software for assessing subjective mental workload. Behav Res Methods 41:113–117. https://doi.org/10.3758/BRM.41.1.113

Chandler P, Sweller J (1991) Cognitive load theory and the format of instruction. Cogn Instr 8:293–332. https://doi.org/10.1207/s1532690xci0804_2

Chung H, Lorga M, Voas J, Lee S (2017) Alexa, can I trust you? IEEE Comput Soc 50:100–104. https://doi.org/10.1109/MC.2017.3571053

Cui L, Huang S, Wei F, Tan C, Duan C, Zhou M (2017) SuperAgent: a customer service chatbot for E-commerce websites. In: Proceedings of ACL 2017, system demonstrations. association for computational linguistics, Vancouver, Canada, pp 97–102

Dellermann D, Ebel P, Söllner M, Leimeister JM (2019) Hybrid Intelligence. Bus Inf Syst Eng 61:637–643. https://doi.org/10.1007/s12599-019-00595-2

Diao W, Liu X, Zhou Z, Zhang K (2014) Your voice assistant is mine: how to abuse speakers to steal information and control your phone. In: Proceedings of the ACM conference on computer and communications security, pp 63–74

Diederich S, Brendel A, M Kolbe L (2019) On conversational agents in information systems research: analyzing the past to guide future work. In: Proceedings of 14th internatioanl conference on Wirtschaftsinformatik, pp 1550–1564

Frick N, Brünker F, Ross B, Stieglitz S (2019) The utilization of artificial intelligence for improving incident management. HMD 56:357–369. https://doi.org/10.1365/s40702-019-00505-w

Galy E, Cariou M, Mélan C (2012) What is the relationship between mental workload factors and cognitive load types? Int J Psychophysiol 83:269–275. https://doi.org/10.1016/j.ijpsycho.2011.09.023

Gnewuch U, Morana S, Adam MTP, Maedche A (2018) Faster is not always better: understanding the effect of dynamic response delays in human-chatbot interaction. In: 26th European conference on information systems

Gnewuch U, Morana S, Maedche A (2017) Towards designing cooperative and social conversational agents for customer service digital nudging view project designing chatbots. In: 38th international conference on information systems

Hart SG (2006) Nasa-Task Load Index (NASA-TLX); 20 years later. In: Proceedings of the human factors and ergonomics society annual meeting, pp 904–908

Hart SG, Staveland LE (1988) Development of NASA-TLX (Task Load Index): results of empirical and theoretical research. Adv Psychol 52:139–183. https://doi.org/10.1016/S0166-4115(08)62386-9

Hoy MB (2018) Alexa, Siri, Cortana, and more: an introduction to voice assistants. Med Ref Serv Q 37:81–88. https://doi.org/10.1080/02763869.2018.1404391

Hu PJ-H, Hu H-F, Fang X (2017) Examining the mediating roles of cognitive load and performance outcomes in user satisfaction with a website: a field quasi-experiment. MIS Q 41:975–987. https://doi.org/10.25300/MISQ/2017/41.3.14

Hu T, Xu A, Liu Z, You Q, Guo Y, Sinha V, Luo J, Akkiraju R (2018) touch your heart: a tone-aware chatbot for customer care on social media. In: Proceedings of the 2018 CHI conference on human factors in computing systems - CHI '18

Hung V, Elvir M, Gonzalez A, DeMara R (2009) Towards a method for evaluating naturalness in conversational dialog systems. In: 2009 IEEE international conference on systems, man and cybernetics, pp 1236–1241

Knijnenburg B, Willemsen M (2016) Inferring capabilities of intelligent agents from their external traits. ACM Trans Interact Intell Syst 6:1–25. https://doi.org/10.1145/2963106

Kretzer M, Maedche A (2019) Designing social nudges for enterprise recommendation agents: an investigation in the business intelligence systems context. J Assoc Inf Syst 19:1145–1186. https://doi.org/10.17705/1jais.00523

Lamontagne L, Laviolette F, Khoury R, Bergeron-Guyard A (2014) A framework for building adaptive intelligent virtual assistants. In: Artificial intelligence and applications

Lohse M, Rothuis R, Gallego-Pérez J, Karreman DE, Evers V (2014) Robot gestures make difficult tasks easier. In: Proceedings of the SIGCHI conference on human factors in computing systems, pp 1459–1466

Luger E, Sellen A (2016) "Like having a really bad PA": The gulf between user expectation and experience of conversational agents. In: CHI '16 Proceedings of the 2016 CHI conference on human factors in computing systems, San Jose, CA, USA, pp 5286–5297

Luthans F (2002) The need for and meaning of positive organizational behavior. J Organ Behav 23:695–706. https://doi.org/10.1002/job.165

Maedche A, Legner C, Benlian A, Berger B, Gimpel H, Hess T, Hinz O, Morana S, Söllner M (2019) AI-based digital assistants. Bus Inf Syst Eng 61:535–544. https://doi.org/10.1007/s12599-019-00600-8

Maedche A, Morana S, Schacht S, Werth D, Krumeich J (2016) Advanced user assistance systems. Bus Inf Syst Eng 58:367–370. https://doi.org/10.1007/s12599-016-0444-2

Matt C, Hess T, Benlian A (2015) Digital transformation strategies. Bus Inf Syst Eng 57:339–343. https://doi.org/10.1007/s12599-015-0401-5

McTear M, Callejas Z, Griol D (2016) The conversational interface: talking to smart devices, 1st edn. Springer, Cham

Mechling LC, Gast DL, Seid NH (2010) Evaluation of a personal digital assistant as a self-prompting device for increasing multi-step task completion by students with moderate intellectual disabilities. Educ Train Autism Dev Disabil 45:422–439

Moorthy AE, Vu KP-L (2015) Privacy concerns for use of voice activated personal assistant in the public space. Int J Hum Comput Interact 31:307–335. https://doi.org/10.1080/10447318.2014.986642

Morana S, Friemel C, Gnewuch U, Maedche A, Pfeiffer J (2017) Interaktion mit smarten Systemen — Aktueller Stand und zukünftige Entwicklungen im Bereich der Nutzerassistenz. Wirtschaftsinformatik Manag 9:42–51. https://doi.org/10.1007/s35764-017-0101-7

Moreno R, Mayer RE, Spires HA, Lester JC (2001) The case for social agency in computer-based teaching: Do students learn more deeply when they interact with animated pedagogical agents? Cogn Instr 19:177–213. https://doi.org/10.1207/S1532690XCI1902_02

Morrissey K, Kirakowski J (2013) "Realness" in chatbots: establishing quantifiable criteria. In: Kurosu M (ed) 5th international conference on human–computer interaction: interaction modalities and techniques - volume Part IV, pp 87–96

Muñoz S, Araque O, Llamas AF, Iglesias CA (2018) A cognitive agent for mining bugs reports, feature suggestions and sentiment in a mobile application store. In: 2018 4th international conference on Big Data innovations and applications (innovate-data), pp 17–24

Neyer FJ, Asendorpf JB (2017) Psychologie der Persönlichkeit, 6th edn. Springer-Verlag, Berlin Heidelberg

Nordstrom CR, Williams KB, LeBreton JM (1996) The effect of cognitive load on the processing of employment selection information. Basic Appl Soc Psych 18:305–318. https://doi.org/10.1207/s15324834basp1803_4

Norman D (2017) Design, business models, and human-technology teamwork. Res Manag 60:26–30. https://doi.org/10.1080/08956308.2017.1255051

Noyes JM, Bruneau DPJ (2007) A self-analysis of the NASA-TLX workload measure. Ergonomics 50:514–519. https://doi.org/10.1080/00140130701235232

Paas F, Tuovinen JE, Tabbers H, Van Gerven PWM (2003) Cognitive load measurement as a means to advance cognitive load theory. Educ Psychol 38:63–71. https://doi.org/10.1207/S15326985EP3801_8

Pace F, D'Urso G, Zappulla C, Pace U (2019) The relation between workload and personal well-being among university professors. Curr Psychol. https://doi.org/10.1007/s12144-019-00294-x

Qiu L, Benbasat I (2009) Evaluating anthropomorphic product recommendation agents: a social relationship perspective to designing information systems. J Manag Inf Syst 25:145–182. https://doi.org/10.2753/MIS0742-1222250405

Quarteroni S (2018) Natural language processing for industry: ELCA's experience. Informatik-Spektrum 41:105–112. https://doi.org/10.1007/s00287-018-1094-1

Romero D, Stahre J, Taisch M (2020) The Operator 4.0: towards socially sustainable factories of the future. Comput Ind Eng 139:106128. https://doi.org/10.1016/j.cie.2019.106128

Rubio S, Díaz E, Martín J, Puente JM (2004) Evaluation of Subjective Mental Workload: A Comparison of SWAT, NASA-TLX, and Workload Profile Methods. Appl Psychol 53:61–86. https://doi.org/10.1111/j.1464-0597.2004.00161.x

Saffarizadeh K, Boodraj M, Alashoor TM (2017) Conversational assistants: investigating privacy concerns, trust, and self-disclosure. In: Thirty eighth international conference on information systems, South Korea

Sangyeal H, Heetae Y (2018) Understanding adoption of intelligent personal assistants. Ind Manag Data Syst 118:618–636. https://doi.org/10.1108/IMDS-05-2017-0214

Schmuntzsch U, Sturm C, Reichmuth R, Roetting M (2012) Virtual agent assistance for maintenance tasks in IPS^2 - first results of a study. Adv Hum Factors/Ergonom 18:221–231. https://doi.org/10.1201/b12322-27

Schnotz W, Rasch T (2005) Enabling, facilitating, and inhibiting effects of animations in multimedia learning: why reduction of cognitive load can have negative results on learning. Educ Technol Res Dev 53:47–58. https://doi.org/10.1007/BF02504797

Schuetzler RM, Grimes GM, Giboney JS (2018) An Investigation of conversational agent relevance, presence, and engagement. In: Twenty-fourth Americas conference on information systems

Schumacher J, Leppert K, Gunzelmann T, Strauss B, Brahler E (2005) The resilience scale-A questionnaire to assess resilience as a personality characteristic. Zeitschrift für Klin Psychol Psychiatr und Psychother 53

Shamekhi A, Liao QV, Wang D, Bellamy RKE, Erickson T (2018) Face value? Exploring the effects of embodiment for a group facilitation agent. In: 2018 CHI Conference on human factors in computing systems

Siddike A, Kohda Y (2018) Towards a framework of trust determinants in people and cognitive assistants interactions. In: 51st Hawaii international conference on system sciences, pp 5394–5401

Siddike MAK, Spohrer J, Demirkan H, Kohda Y (2018) People's interactions with cognitive assistants for enhanced performances. In: 51st Hawaii international conference on system sciences, pp 1640–1648

Smith C, Crook N, Dobnik S, Charlton D, Boye J, Pulman S, de la Camara RS, Turunen M, Benyon D, Bradley J, Gambäck B, Hansen P, Mival O, Webb N, Cavazza M (2011) Interaction strategies for an affective conversational agent. Presence Teleoperators Virtual Environ 20:395–411. https://doi.org/10.1162/PRES_a_00063

Stieglitz S, Brachten F, Kissmer T (2018) Defining bots in an enterprise context. In: Thirty ninth international conference on information systems

Stöckli E, Uebernickel F, Brenner W (2018) Exploring affordances of slack integrations and their actualization within enterprises? Towards an Understanding Of How Chatbots Create Value. In: 51th Hawaii international conference on system sciences

Strohmann T, Fischer S, Siemon D, Brachten F, Lattemann C, Robra-Bissantz S, Stieglitz S (2018) Virtual moderation assistance: creating design guidelines for virtual assistants supporting creative workshops. In: 22nd Pacific Asia conference on information systems, pp 3580–3594

Sweller J (1988) Cognitive load during problem solving: effects on learning. Cogn Sci 12:257–285. https://doi.org/10.1016/0364-0213(88)90023-7

Tarmizi RA, Sweller J (1988) Guidance during mathematical problem solving. J Educ Psychol 80:242–436. https://doi.org/10.1037/0022-0663.80.4.424

Verhagen T, van Nes J, Feldberg F, van Dolen W (2014) Virtual customer service agents: using social presence and personalization to shape online service encounters. J Comput Commun 19:529–545. https://doi.org/10.1111/jcc4.12066

Youssef CM, Luthans F (2007) Positive organizational behavior in the workplace: the impact of hope, optimism, and resilience. J Manage 33:774–800. https://doi.org/10.1177/0149206307305562

DOI: 10.17185/duepublico/74364
URN: urn:nbn:de:hbz:464-20210519-091551-3