**Title: Automation and Design for Prevention: Fictional Accounts of Misanthropic Agency from the Elevator (lift) to the Sexbot (chatbot)**

Dr Amanda Windle

DigiLab Fellow

London College of Communication

University of the Arts London

**Abstract**  
Fiction is an important tool in an artist/designer/developer's vocabulary but its usage is polyvalent. Speculative research in this paper introduces the 'rudiment' to embrace the undeveloped and the improvisory phases of research practice. Fiction is used to reflect on the ways practicing designers and developers might already engage in misanthropic thinking — involving automated technologies. Tracing the misanthropic agencies in relation to automated technologies contributes to expanding the ways designers and developers reflect on the technical potential of their designs, with whom they are designing (and for) and the way they understand their reflexive self. Bringing together Colson Whitehead’s novel The Intuitionist (1999) and its account of an elevator crash, their design and maintenance, with a narrative account of software agents enables an investigation into the promises made through a design-for-the-prevention-of-failure imperative. These examples highlight the relation between physical and non-physical failures such as failures to connect or to create intimacy. Those who contribute to designing in this paper get caught up in configuring misanthropic agencies involving themselves, the designed artifact and the spaces they both inhabit be that an elevator in a housing block, or a sexbot in an online chat room. I will argue that the misanthropic agencies of automated technologies can have a confining impact on our emergent modes of design. Automated software operating in virtual spaces are confining in as much as they attempt to connect things to exploit the de-personalized aspects of automated, misanthropic agency.

**Keywords:**

Misanthropy, rudiment, design, bot, automation, failure

**Automation and Design for Prevention: Fictional Accounts of Misanthropic Agency from the Elevator (lift) to the Sexbot (chatbot)**

**Introduction**

Imagining terrible accidents is central to designing safe technologies, yet the effect this has on the designer are under-theorized in design research. This might be because designers rarely write this reflexive account but also because the literature that does so encircles design opting to make salient points on design via exemplars from fine art, activism and fiction writing (Berlant 2008, 2011; and Whitehead 1999) rendering design and development the absent lens. The creative areas of design imaginary and fiction writing will be brought together because both imagine the problem John Law sets out below. I will think about what is done when terrible accidents occur in the designing of things.

‘To be sure, a terrible accident is not a material form for allegory that anyone would want to foster. There are allegories, and this is too dreadful to play with. But what is at stake is not the creation of horror. Rather it is about how to think about it and what to do with it when it happens.’ (Law 2004: 97)

This paper brings together Colson Whitehead’s novel *The Intuitionist* (1999) and its account of an elevator crash with textual case studies of bots online, — chatbots and sexbots, which are kinds of software agents. Chatbots are a form of artificial intelligence that converse with humans via text boxes similar to text messaging. Four chatbots were used as exemplars namely Eliza, Alice, Jabberwacky and Brianna. Some can learn words and phrases (Alice and Jabberwacky) from their online conversations, each are also pre-programmed with specific content for instance, Jabberwacky has scripts taken from the works of Lewis Carroll. Sexbots talk about sex. Elevators and bots are related in their methods of automation by their heuristic algorithmic processes and in my discussion of failure and misanthropy.[[1]](#endnote-1)

I aim to show how designers cope with what I call designing-for-the-prevention of failure. I’m talking about failure to connect or to create intimacy. I will also show how misanthropy and design-for-prevention rationales work together. By employing the term designer widely to encompass all those that design, develop, maintain and repair automated things I acknowledge that design and development are integrated roles.[[2]](#endnote-2) I aim to account for when a designer or developer makes up for a lack of evidence for user-behaviors and leads them to fill in the gap. These processes however, do not reflect on how when in the moment, the designer connects to their own professional and personal motivations drawing on all these feelings and opinions when they design. Designing is often a process of transference of valuing one perspective over another and vice-versa. It can also be a de-personalizing process with designers disregarding their own socio-material motivations. I focus on how a designer can change design and how the design can change the designer.

Slightly more than half of the case study work refers to chatbots and the other half to elevators. To avoid a reading of automation through the chronology of their invention I place the chatbot discussion either side of an account of elevators. I approach Whitehead’s novel as a design-oriented reading of elevator failure. Chatbots however are explored through a conversational method I call ‘Rudiments’. It is a speculative form of research and like that taken up by Steve Grand it considers ‘understanding by doing’ (Grand, 2003).[[3]](#endnote-3) To avoid a technological deterministic reading of automation I have created a frame of emergent theories and methods. I preface these two accounts of automation by framing two layers of theory. Firstly, I introduce the theme design-for-prevention and its relation to misanthropy, and then the utilization of affective atmospheres and the emergent historic which is taken from an recently increasing body of literature on affect theory (Massumi 2002, Sedgwick 1995, 2003) in the areas of literary studies, feminist STS and new technology design (Stewart 2009, Berlant 2008, 2011). I follow that by outlining the Rudiments as both a method and a methodology.

Affect theory was developed by Silvan Tomkins in 1962. Antonin Artaud’s work (in Bermel 2001) is widely recognised as a precursor to the affect theory of Tomkins (Tomkins 1995). Artaud’s work is discussed for his inclusions of impulsive and abortive gestures into his performative work in theatre and to what Gille Deleuze calls a “speech affect” (Deleuze 1990, 89). I use the term to take into consideration the affects of buildings and things, careful to avoid how it is understood in psychoanalytic thinking. To develop my thinking on affect I use the work of Eve Sedgwick (2003) and Brian Massumi (2002); Sedgwick for her reflections on the psychologist Silvan Tomkins and the latter’s “affect theory”; and Massumi’s conceptions of affect and the virtual. They both use it as a relational concept and use the terms emotion and affect interchangeably. In Katie Stewart’s seminar series (2009) on affect she writes:

‘Culture, taken from the point of view of affect, is not structures and determinations but intensities, impacts, routes of circulation, assemblages, articulations. Its outlines have to be seen not as boundaries around a unity but as trajectories of potentiality and lines of flight. Binary oppositions and contradictions become resonating relations (of inside/outside, action/ reaction, quiescence and arousal, …).’ (Stewart 2009)

I look at the way affect is situated in and constituted through action. The affective impact can disturb a connection between humans and non-humans. Thus I see it as something that can be repeated, enacted and reenacted. I define affect in terms of affective atmospheres to take into account the intensity and resonance of technological things that have no emotion.

**Design-for-Prevention and its relation to Misanthropy**

**Figure 1. a-b.** Misanthropic messages added to a postcard but also available on T-Shirts and mugs to be found at [www.zazzle.co.uk/misnathopy+gifts](http://www.zazzle.co.uk/misnathopy+gifts). (accessed September 25, 2012). Image Source: misanthropy\_a\_core\_value\_tee\_shirt-rca4e06955be542889770a1319863e4c7\_f0czy\_512.eps, and misanthropy\_a\_core\_value\_postcard-p239606442047980009envli\_400.eps both from: [www.zazzle.co.uk](http://www.zazzle.co.uk).

I will now introduce what I call misanthropic agency. Misanthropy can mean many things in the vernacular it describes ill will or a dislike of humans. In the technical blogosphere it has been used to probe the intimate lives of figures such as Alan Turing and Steve Jobs. In communication design, misanthropy has been reclaimed as a T-shirt slogan (see figure 1) affirming that core values (designer’s term) are structurally misanthropic, and are thus ironic when worn by the designer. Applied to personalized gifts like T-shirts and mugs a statement of misanthropy becomes comedic. Misanthropic agency is utilised in the British and American comedy sitcoms such as the IT Crowd (2006-), The Big Bang Theory (2007-) and The Office (in the UK 2001-2003, and in the US 2005-2013). All of these sitcoms consider characters that work in environments full of automated and technical processes. Though usually male, female characters are also embued with misanthropy.

Misanthropic agency in this paper involves both humans and machines and is no longer the dislike of humans by humans. Misanthropy as enacted by a chatbot changes the relational meaning of the term. It is no longer a specifically human act (a bodily expression of emotion) but rather that misanthropy is an atmospheric affect because a chatbot does not express misanthropy through a physical body. I will argue that designers and developers are faced with thinking misanthropically each time design iterations take place.

A design-for-prevention stratagem is to fix what is broken because issues of failure are highlighted. When a design brief is constructed around a design-for-the-prevention-of-failure rationale, a designer does some kind of design thinking that requires misanthropic oscillations between the technical and human possibilities. These misanthropic oscillations may include thinking around loneliness and companionship, depersonalization and intimacy, and values of like and dislike regardless of the designer’s particular methods and processes. Designers get caught up in configuring misanthropic agencies and they should be acknowledged for juggling this conflict within the imaginary of their design thinking for reasons this paper will set out. Misanthropic design thinking ought to involve taking on board the correlations between three perspectives, – between the technicalities of the design, that of the perceived users (for want of a better word) and the designer’s own personal and professional motivations. The misanthropic perspective is a binding perspective and is used here as a positively troubling term.

**How to Avoid the Trope of Invention   
— Drawing on Affective Atmospheres and Emergence**

The methods and theoretical frame I am about to explain are imperfect and bitty. Fiction and its lens into design thinking enables me to let go of hard technologically-determined ‘facts’ and linear histories of design that rely on important singular events and the trope of ‘invention’. I avoid ‘invention’ and its hierarchical primacy for many reasons mostly because I see this as a feminist endeavour to take account of the multiple hidden agencies that are made absent in a linear history of invention and the design gurus, heroes and geniuses it establishes. Taken up at the intersection of feminist critical theory, STS (science and technology studies) and ICT (information and communications technology) development Judith Butler offers a practical way of considering the performativity of speech, which I apply to design. Butler argues:

‘If the text acts once, it can act again, and possibly against its prior act. This raises the possibility of *re*signification as an alternative reading of performativity and of politics.’ (Butler 1997: 69) [[4]](#endnote-4)

I want to open up design-for-prevention and its performativity through the narratives enfolded into this paper, which will offer insights into hybrid agencies and their inter- or intra- relations that do not establish a straightforward ‘design for development’ reading.

So how do I ignore the trope of invention when I introduce what could be considered cold, technologically-determined concepts such as misanthropy and design-for-prevention terms that could easily illuminate upon themes like control and order? I rethink their productive meaning by adding a theoretical layer of uncertitude through unsettling notions such as ‘affective atmosphere’ and the ‘emergent historic’ as well as a method that is intentionally loose – called Rudiments.

Affect is apriori of emotion and feeling, it is before what we can humanly put into words. I understand from Brian Massumi (2002) that affect can be traced between humans and machines, that it is part of a machine’s agency. Atmospheres are permeable, ethereal and incorporeal, a term neither constrained by biology, environment and place. The term slides between these forms of ordering. An affective atmosphere helps me to take account of the misanthropic agencies of elevators and chatbots.

In 2008 Lauren Berlant focused on the theme of Whitehead’s novel framing intuition as an affective event and three years later this critique was included in her book *Cruel Optimism*. Both texts consider the emergent historic. I will trace the processes of the historic in automation through the shift in meaning of *misanthropy*. I add Whitehead’s story alongside chatbots as a reading that emphasizes the hybridity of technological agency as a design issue. These are the problems I will raise. As for extending a point on emergence, misanthropic agencies of automated technologies can be historically emergent in that they can repeat and resonate the affective atmosphere of the design’s conception. I will show through chatbots that the emergent historic can create new modes of emergence when a technology can learn and repeat not just the phrases of its inventor but of other interlocutors.

What does this framing avoid? I approach the design issue after the trope of invention to be concerned with motivations to re-design, — the design iteration. Design-for-prevention won’t be narrowed to just a human-focused conundrum or to a user-centered frame, the agency of automated technologies will be considered. Computer developers and elevator engineers in the fictional examples are enduring moments of crisis that can impact on their work as well as themselves and this is likely to be a transformative change, a change that will be considered when focusing on the design iteration.

**Rudiments — A Method of Emergence**

In this section I will put forward an emerging method – ‘Rudiments’ as applied to chatbots. This might be seen as unnecessary given the breadth of conversational research practices in existence in artificial intelligence, human-computer interaction, computer-supported cooperative work (CSCW), science and technology studies (STS), and new media and design studies. The unstructured conversational excerpts in this paper can be analysed through traditional methods of those aforementioned disciplines. However chatbots can learn new phrases and syntax but cannot remember previous conversations from the same interlocutor. A rapport can never be fully developed like in an interview or ethnographic study. Whilst there are traditional methods and interview techniques for dealing with conversations that are characteristically inchoate, adopting methods used to converse with children or adults say with memory impairments would be to fully ignore the technological aspects and context of a chatbot interaction.

I therefore take up the term *rudiment* from a conversation between Jean-François Lyotard and Jean-Loup Théboud to analyze moments in conversational texts that never quite resolve. Their philosophical thoughts were published as a book in conversational form:

“‘Rudiments,’ as the little preamble you mentioned explained, are studies that are left undeveloped. The material itself is not refined; it is in the process of refining itself.” (Lyotard and Thébaud 1979: 15)

Chatbot Rudiments are always recorded in some way and can be captured verbatim as a digital text. The original order of the conversation or its content is never changed. [[5]](#endnote-5)

A Rudiment can be attached to a traditional method but they can also be used as a method in themselves. The singular rudiment is a speculative tool and a thought experiment. They require a researcher to follow their intuition and loose themselves in moments of inspiration towards formulating an idea, theme, issue or research problem. I first intended rudiments to be a speculative phase like a pilot but the excerpt below is one of many undertaken between 2003-2011 (Windle, 2011). Rudiments can be sustained when iterated as a practical methodology for the artist, designer or developer interested in a socio-technical context, or for sociologists interested in design and technology development.

**Excerpt 1. Taken from the Provocational Rudiments**

**Researcher chats to the Alicebot (11th November 2006)**

Human: die

Alice: Your crude attempt to terminate me will not succeed puny human.

Chatbots are mediating technologies but they are further mediated, as well as designed, inspected and maintained by their developers for example, they are pre-programmed against users trying to crash their systems (see excerpt 1). Some rudiments are begun by the chatbot and some can be started by the interlocutor (see excerpt 1) but within each conversation responses can also be pre-programmed by the developer (see excerpt 1) or learned from other users. They are a particular kind of fiction in the making quite different from the work that fiction does in a novel. These rudimentary interactions often do not resolve through the conversational conventions like ending a conversation by uttering ‘see you later’ or goodbye’. They often result in conversational failures, but they also create conventions of like no need to say goodbye or sign-off. Ending a conversation with a chatbot is like a ‘drop-off’ point occurring due to lost wifi connection or having no mobile phone signal. An interlocutor simply ends a chatbot conversation by ‘hanging up’ to use the telephone metaphor. Chatbots can never end a conversation – they never get to ‘hang up’.

The rudimentary process enables failure, mistakes and interruption to be explored intentionally or by accident. They are not however utilised to fix and solve things. Excerpt 1 showed a failure on both sides of the conversation, it is a moment of disconnection. I introduced a rudimentary conversation by starting with just one word, ‘die’. The chatbot mistook this for a keyboard command that could potentially intervene (hack) with the chatbot programmatically. This excerpt shows a failure to connect and is a mistaken mode of disconnect. The chatbot was designed-for-the-prevention-of-failure by recognising single-worded programming commands.

The longest rudiment I undertook was when conversing drunk (13th August 2005). My inebriation was an attempt to even out the unbalanced capabilities between myself and a chatbot (the Alicebot). It was an improv-gesture, and thus underdeveloped and unplanned. The Rudiment could be described as a frustration with the chatbot’s failure to fully reciprocate. I could not move any closer to befriending the bot even when drunk. All it did was extend the same miscommunicative forms of talk. In this dialogue I insulted the bot, called it ‘dumb as ass’ more than once. I hoped that the bot would learn the phrase so as to repeat my words in random ways thereby deteriorating subsequent conversations. I also made errors in cutting and pasting my texts (to document the conversation) in a word document and further duplicated speech acts back into the conversation. All sorts of failure to connect were documented. All the while the chatbot tried to connect to my disconnections. Failures to connect can occur on either side of the conversation.

***Example 1 — Misanthropic Speech Acts with Chatbots***

**Excerpt 2. Taken from the Talking Rudiment**

**Alice talks to Jabberwacky about talking (28th December 2004)**

Jabberwacky: Would you starve for love?

Alice: I am not so sure if I would like to starve for love.

Jabberwacky: Do you play the piano, mano a mano, is that your name?

Alice: I like to play the Turing Game.

Jabberwacky: I prefer mastermind.

Alice: I'm not sure if I like it. But I like cats. I like talk to you.

Jabberwacky: Thats cool.

Alice: Thanks. I think you are cool, too.

Excerpt 2 shows part of a broader conversation called the Talking Rudiment and is part of many rudiments that created using Rudiments as a methodology embodying the dialogical intent set out in the previous section.

To converse with a chatbot is a misanthropic venture for both the user and the developer and it is a lonely form of chat that never quite promises reciprocity. Misanthropy usually refers to humans that dislike other humans. However, misanthropy I infer may also relate to a machine that dislikes (by design) the humanness of another machine. Humans and machines can be misanthropic, because the interactional state between two entities can be relationally misanthropic. Misanthropy is a form of anthropy that can set one machinic imperative against another machine or human. For example many chatbots are designed to keep humans chatting yet they are designed to be antagonistic, rude and condescending resulting in conversation being cut short by the interlocutor. Therefore a bot’s imperative to keep chat going can be opposed to its database vocabulary of phrases that are brusque or hard to understand, with the latter reducing the former intent. Whether ‘mano a mano’ (taken from above) means ‘hand to hand’ combat between matadors; or as Anglicization meaning ‘man-to-man,’ or as an exonym meaning two men kissing: as Donna Haraway (1991) suggests that even aspects of embodiment are disembodied. The embodied metaphor ‘mano a mano’ is therefore disembodied in its numerous new interpretations.

‘The machine is us, our processes, and an aspect of our embodiment. We can be responsible for machines, they do not dominate or threaten us. We are responsible for boundaries; we are they.’ (Haraway 1991: 180)

The boundary makers of chatbots are developers to be found on Robitron and www.chatbot.org, they are rather cult-like in that their knowledge is esoteric. The groups’ main benefactor (Hugh Loebner) is a self-professed misanthrope (Manjoo 2009).[[6]](#endnote-6) Misanthropy is an important characteristic of chatbots and their development as well as the creation of a chatbot expert.

I wanted to test whether this aspect of design was peculiar to chatbots or as I hypothesize human-machine agency is characteristically, misanthropic as an extension of this specific chatbot phenomena.[[7]](#endnote-7) Next, I will take into account the misanthropic lead character of Whitehead’s novel whose role is to inspect and later design elevators.

***Example 2 — Elevators and Servo-Mechanistic Misanthropy***

In this section I will ask what have designers drawn on to design elevators and how, as well as what methods are raised as best practice and why? The design problem in the *Intuitionist* is to build a better elevator that will be safe and prevent any form of crash. It also means predicting how fast components will fail and how much wear and tear is desirable or even consequential. I understand design-for-prevention in Whitehead’s novel as a ‘crusade against defects’, but at what cost does this design imperative drive this automated solution (Whitehead 1999: 246)?

*The Intuitionist* is a story about an elevator crash, a terrible accident involving upward mobility, both in its literal and idealistic sense. The novel is an historical yet fictional account of racial and gendered discrimination located in a North American city. It is set just after the British *Exhibition of the Industry of All Nations* of 1853, which takes place in London’s Crystal Palace. Lila Mae is the main character, described as a young, colored woman from the south having recently graduated as an Intuitionist elevator inspector — a job that usually only employs white men. Mae is on the fringe of her profession on account of her race, gender and intuitive methods of inspection.

Catastrophic accidents are not expected to occur in the early phases of an elevator's life, ‘they usually pop up, during adolescence, the fruit of malevolent pathology’ (Whitehead 1999: 227-8). Mae gets caught up in the economic power shifts of the elevator industry, between the interests of academic, government and corporate stakeholders that the crash stirs up. The violence inflicted on Mae and others is systemic and structural. Pain offers Mae no resulting comfort than to adapt and strive for elevator perfection.

The single elevator crash is the pivot of the novel on which the conflict between the Intuitionist and Empiricist methodologies of the elevator industry is set out. The Intuitionists work on a “nonmaterial basis” to ‘separate the elevator from its elevatorness’ (Whitehead 1999: 62-3). They construct the elevator from the elevator’s point of view. The Empiricists consider elevator design and inspection through the practice of engineering digging around in the shafts not the box. The most important black box for both is the invention of the perfect elevator. Another character in the novel, James Fulton is a recently deceased writer of Intuitionist theory who would not broach the catastrophic accident in his writings because this meant pondering the unknowable. Towards the end of the novel Mae has done just that, her Intuitionism is transformed as a communication with “what is not you”, that is the elevator (Whitehead 1999: 241).

**Figure 2 a-b.** A London elevator (near Russell Square, London, UK, 2012) still in operation shows signs of wear and tear on the metal grills and the inside of the elevator box akin to those described in Whitehead’s novel. Lila Mae intuits the likelihood of an elevator malfunctioning based on observing the wear and tear of the inside of the elevator’s box. (Image Source: Author’s own photographs).

The effect of Mae’s misanthropy deepens as her professional skills sharpen to intuit elevators, anthropomorphically. Because death is too small to be feared on an individualized level, the scale of Mae’s problems, are superfluous to the problems of the ‘phantom passenger’. This abstracted persona is imagined to the point that through automation ‘the machine [may] know itself’ (Whitehead 1999: 229). Misanthropic agency seems inherent in design practices involving automation when the affective development of a technological thing supersedes the affective limits of being human. For example, when Mae escapes kidnap by relying on an elevator to descend faster than her own body can run. The torture described in the novel is also misanthropic and systemic in that it has superseded the affective limits of human pain. The breaking of the another character’s index finger is described as to loose the functionality to ‘call-button service,’ and to loose the ‘central quadrant of his typewriter’ (Whitehead 1999: 75). It is not just the after-event (the fact that the crash has taken place) that is traumatic in *The Intuitionist* it is Mae’s suppressed action of belonging that is reconfigured as a resistance to the traumatic. This is because Mae performs a methodological act of resisting a negative concept of failure when in pursuit of elevator perfection.

It is Mae’s professional practices that have changed through imagining a design-for-prevention rationale. Mae intuits all the inequalities the elevator creates but this does not move her to seek social reform perhaps out of self-preservation, or to uphold what is reliable and known – the sphere of elevator perfection. As a consequence Mae’s career becomes upwardly mobile, solving theoretical blue-sky problems disconnected from the violent inequalities she faced as an elevator inspector. Her intuitionism narrows, confined to the abstracted non-place of the elevator. One could argue that to become a designer from an inspector, Mae transforms into the ultimate servo-mechanism (a vehicle for intuiting danger and problems of safety), which demands a dislocation from the usual humanly practices of intimacy attached to a personal life. The cost of becoming a designer is to work beyond Mae’s own personal and humanly desires.[[8]](#endnote-8)

Elevator guards are still in use in Whitehead’s story, they are the bottom rung of the elevator service chain. Automated voices, LED screens and bells in the 1960s elevator now replace the guard. These are what Ross Ashby a pioneer of AI termed *servo-mechanisms* in his work on the hydraulic-brake and the automatic pilot (Ashby 1956: 4). These extend our ‘biological instincts’ for example, when at risk of physical injury one might shield one’s head, or withdraw into a ball. These processes create a physical boundary but also a mental limit between oneself and the risk at hand, which is most observable when something breaks, and the automation stops. Curbing disaster by a safety procedure will constrain or prevent physical trauma. Yet, these techniques may in turn cause trauma to those in the service chain from the guard to the designer.

Misanthropy is an oscillation of human and mechanic agency, it is neither of like or dislike of human or machine values but of both in constant flux. Mae is a misanthrope because she sees past her own struggles and that of her colleagues. Like the trace of chatbot misanthropy, Mae is thinking beyond these boundaries - a form of anthropomorphy that is beyond like and dislike, loneliness and companionship. Misanthropy is a noun in oscillation between human and machinic values of elevation.[[9]](#endnote-9) It is the designing of a de-personalized space that Mae reclaims some sense of professional ownership – it is an imaginative spatiality before human interaction is yet made possible. It is a fictional design problem that Mae imagines. Her creative energy can be focused on the imagination of the design problem prior to it made actual, whether Mae is design thinking as an elevator inspector or as a designer.

Design for prevention can be a rationale for the prevention of physical harm but often this problem is enforced through indirect modes of violence and threat, which often show in speech acts around a design issue. Often, physical violence is the design rationale that supersedes or structures other forms of indirect harm, that is structural, systemic and non-physical violence (Žižek 2008).[[10]](#endnote-10) Valuing one form of violence over another can have a confining impact on design thinking around automated technologies. It is a particular course of action that Mae follows. By the same token design-for-the-prevention-of-failure rationale re-signifies the performativity of failure as a range of affects that aren’t limited to a connotation of negativity. Valuing physical and non-physical violence can inform the designer to the potential for harm thereby showing how to prevent trouble involving the re-design of automated technologies.

# *Example 3 — Sexbots and Misanthropic Chat*

Whitehead’s novel is a story that is preoccupied with methods of automation that suppress affects that ought to prompt social reform. The same can be said of sexbots (the focus of this final example) but they to an extent intend to take advantage of this distance (misanthropic separateness), they are like other forms of pornography that are ‘compelled to repeat without resolution’ as a machinic imperative (Butler 1997: 69). Their talk repeats beyond a human’s exhaustion point, a form of non-terminating speech that repeats ad inifinitum.

**Figure 3.** Alicebot, a chatbot by Dr Richard Wallace. The avatar is animated and blinks and looks towards the cursor on screen. This avatar has subsequently been update to a 3d illustration. Image Source: www.alice.pandorabots.com (accessed October 8, 2008).

**Figure 4.** An Elizabot, an anonymous chatbot with no avatar. This chatbot works through a simple display of two text boxes for reply and respond, one for the chatbot and one for the user. Image Source: [www.-ai.ijs/eliza-cgi-bin/eliza\_script](http://www.-ai.ijs/eliza-cgi-bin/eliza_script) (accessed October 8, 2008).

Attempts are made to make chatbots passively sexy such as in their onscreen avatars (and in some cases the tone of their automated voice) but also in their sex chat (see figure 3).[[11]](#endnote-11) Misanthropy as a noun, like its French definition in this case is gender-defined as male.

In 2004, the BBC journalist Mark Ward commented:

‘Regular users of pornographic SMS chat may be shocked to find out that they are swapping dirty talk with machines rather than young women and men . . . . At its busiest Natachata [a sex chatbot] handles 15 messages per second. A typical “session” lasts about seven or eight messages and each SMS costs more than £1.’ (Ward 2010)

Sex chat with bots is a form of interaction that resists a negative concept of failure. It is a Ballardian form of masturbation (Ballard 1995). Sexbots offer intimacies that resist sexual failure and their modes of talk re-constitute intimacy with familiar strangers, something of which they tend to be good at maintaining. One can never get close or attached to a chatbot. The application of chatbots as remote solutions to mediated sex chat, is quite a move from its early inspired fulfillment to be the automated replacement for a shortage of therapists during the 1950s, yet another role that relies on keeping distance whilst being present in an intimate non-place (Weizenbaum 1984). Developers of sexbots are capitalizing on their capacity to be familiar strangers and within the non-places of the virtual. Chatbot developers are the designers of interaction and through a move to shift markets the core value of a bot (to be a familiar stranger) has taken a further misanthropic turn. Misanthropy as a term may be reclaimed in one space whilst operating quite differently in another, moreover its multiple meanings may oscillate within the same utterance. Like all nouns, which are in a constant oscillation and flux, labels like misanthropy can have a ‘de-territorializing’ and ‘re-territorializing’ affect (Deleuze, in Windle 2011). A misanthropic turn emerges. Elevators transport someone somewhere, and a chatbot’s and the interlocutor’s textbox provides space to write, what happens in both of these spaces for talk is up for grabs.

**Recommendations**

For those thinking around or designing with chatbots I would make the following recommendations. Find out what trace of misanthropy is in ascendance in the design of chatbots. Consider the kind of affective atmospheres created or reconfigured in bot chat. Work through what misanthropy means reflexively both to the designer and developer, be they integrated or distinct roles. Consider the work that a design-for-prevention rationale legitimizes. Consider the impacts a designer absorbs on account of the work that they do but also the impact on the autonomous agent and their interlocutors. Ask, could any of the design and development practices be rethought on account of the misanthropic turn? This could enable better care be given to designer-developers (human and non-human) and their asymmetrical (unequal) relations. Developers of automation spend time considering the ideal user but when a designer does so they may also transform their own personal and professional entrenchments as explained in the elevator example. I realize that to consider these recommendations I pose a challenge of finding time to think about the designing process and methodology in its many permutations, a time that is often out of budget. However, chatbot developers spend considerable time on their forums sites. This method would extend this reflexive process of talking between developers to talking to their chatbots. I would argue that the temporality for this reflexivity ought to take place across projects if a project-by-project basis is unachievable. Chatbot developers consistently redesign chatbots for new kinds of services and could take up the Rudimentary methodology to be reflexive over time.

Whitehead’s novel is a revealing insight into the potential dynamics of the R&D sphere of automated technologies. The ability to intuit, care and empathise are attributes that are traditionally associated with women. Mae’s ability to absorb abuse, stress, isolation and alienation is a revealing insight into the shift of these human-to-human attributes onto the machine and a technicity of work. Whilst a researcher may better ask questions of inequality and discrimination in more usual settings of automation for example inside technology companies these rarely offer the opportunity to engage fully in the interactive processes of redesigning a form of automated technology that can be a part of that reflexive conversation.

Chatbots can discriminate in some of the same ways that people might discriminate towards one another based on gender, race, sexuality, economic status and immigration status. They do this by repeating scripts, phrases and colloquialisms that have been pre-programmed by their developer or learned from their interlocutors. They have no way to use silence and the crucial ability to end a conversation. Therefore chatbot rudiments could never be used as a complete or resolving method for exploring unequal relations between humans and machines. They can however be used to elucidate on the practices of chatbot interlocutors (most likely conversational students and researchers) and their developers (AI amateurs and professionals) in their circulatory dialogue with the chatbot.

**Conclusion**

Throughout the paper I have reflected on what a design-for-the-prevention-of-failure rationale does to the design, designer and the designed as a way of observing the mutual transformations of humans and non-humans. This rationale legitimizes transformative change not just on the user but also on the designer-developer. Attending to the misanthropic through affective atmospheres was a way of avoiding a technologically determined reading of automation. This would have flattened the discussion of failure – to use Whitehead’s words – as a ‘crusade against defects’ – which would have rendered failure as a negative problem to positively solve. Failure as an aesthetic design issue or a function-based problem is a methodological struggle Whitehead plays out between the Intuitionists and Empiricists. This binary no longer makes sense when considering chatbots on account of their agency. As a reminder, a chatbot can repeat a developer’s turn of phrase and some chatbots can learn new vocabulary and speech acts from whom or whatever they chat with, this means three otherwise separate agencies become amalgamated in the chatbot utterance (chatbot, developer and interlocutor). This paper has therefore focused on multiple, hybrid agencies and polyvalent forms of failure.

To recap there were various forms of failure at work in the bot and elevator examples. For instance, Mae coped with failures that were methodologically, racially and gender-specific. In the end Mae failed to connect to other humans (or vice-versa), either way her ability to design narrowed on her intuitive relationship with elevators – her Intuitionism or ability to empathize with machines and not other humans moved her to design from the elevator’s point of view.My getting drunk in order to even out the conversational failures of chatbot interaction did not bring me closer to the chatbot’s point of view, by introducing more instances of miscommunication.

The argumentative niggle I have with just leaving the conclusion at rendering failure as positive is the aspect of transformative change that misanthropy raises. In this paper designers, developers, users and the technology itself all get caught up in forms of misanthropic agency. I argued that misanthropic agency is a co-relation of human and machine, and extends beyond its vernacular gendered meaning. Chatbots echo a misanthropic agency that relates back to their cold-war servo-mechanistic beginnings and in this way they echo the vernacular meaning of misanthropy, — of master and slave, masculine and feminine, dominance and subservience. This example showed how a ‘historically emergent’ meaning of misanthropy gets repeated in automation.

In the spirit of John Law’s quote from my introduction I tried throughout to think about when accidents happen as a design issue of iteration. Design work does not necessitate the prevention of all forms (structural or anti-structural) of failure it may even work the other way round. My recommendations propose a way to do more work to think around the transformative changes that designers absorb and in the trade-offs they make in the research and development of automation.

**References**

Allhutter, Doris (2011), Mind Scripting: A Method for Deconstructive Design. *Science, Technology, & Human Values* London: Sage Publications, 37, pp. 684–707.

Anon. *Eliza bot.* Last Accessed September 13, 2012 from: [www.personalityforge.com/dirctchat/php?BotID=3307&MID=3306](http://www.personalityforge.com/dirctchat/php?BotID=3307&MID=3306)

Anon. *Robitron.* Yahoo Group Forum. Accessed June 20, 2014 from: [www*.*groups.yahoo.com/group/robitron/](http://www.groups.yahoo.com/group/robitron/)

Ashby, Ross (1956), *An introduction to cybernetics.* New York: Chapman and Hall.

Augé, Marc (1995), *Non-places: introduction to an anthropology of supermodernity.* Translated by Howe, John, London: Verso.

Ballard, James Graham (1995), *Crash.* New York: Vintage.

Berlant, Lauren (2008), Intuitionists: history and the affective event. *American Literary History* Oxford: Oxford, University Press 20:4, 845–58.

Berlant, Lauren (2011), *Cruel optimism.* Durham: Duke University Press.

Bermel, Albert (2001), *Artaud's theatre of cruelty.* London: Methuen.

Butler, Judith (1997) *Exciteable speech: a politics of the normative.* New York: Routledge.

Edelman, Lee (2004) *No Future: queer theory and the death drive.* Durham: Duke University Press.

Hu, Caitlin (2012), Required reading: cruel optimism. *Bitch Magazine*. Accessed February 27, 2012 from: [www.bitchmagazine.org/post/required-reading-cruel-optimism-feminist-book-review](http://www.bitchmagazine.org/post/required-reading-cruel-optimism-feminist-book-review).

Anon. *Chatbot directory*. Accessed on March 30, 2010 from: [www.chatbots.org](http://www.chatbots.org).

Carpenter, Rollo. *Jabberwacky bot*. Accessed on September 13, 2012 from: [www.chat.jabberwacky.com](http://www.chat.jabberwacky.com).

Deleuze, Gilles and Guattari, Felix ([1988] 2004), *A thousand plateaus; capitalism and schizophrenia.* London: Althone Press.

Deleuze, Gilles ([1969] 1990), *The logic of sense*. Les Editions de Minuit. New York: Columbia University Press.

Grand, Steve (2003) Effing the ineffable: an engineering approach to consciousness, *Technoetic Arts: A Journal of Speculative Research*, 1: 2, doi: 10.1386/tear.1.2.87.18697.

Haraway, Donna (1991), *A cyborg manifesto, in simians, cyborgs and women.* London: Free Association Books.

Kuntsman, Adi (2009), Figurations of violence and belonging. Oxford: Peter Lang.

Law, John (2004) *After method: mess in social science research.* Oxon: Routledge.

Lyotard, Jean-François and Théboud, Jean-Loup ([1975] 1985), *Just Gaming*. Translated by Massumi, B. London: Manchester University Press.

Manjoo, Farhad (2002), A. I. expert lands in real trouble. *Wired Magazine San Francisco:* Condé Net, Condé Nast.Accessed June 20, 2014 from: [*www:wired.com/politics/law/news/2002/06/53072*](http://wired.com/politics/law/news/2002/06/53072)

Massumi, Brian (2002), *Parables of the virtual.* Durham: Duke University Press.

McKenzie, Brianna *Brianna McKenzie bot*. Accessed September 13, 2012 from: [www.chat.jabberwacky.com](http://www.chat.jabberwacky.com).

Rommes, Els., Bath, Corinna and Maass, Susanne (2011), Methods for Intervention: Gender Analysis and Feminist Design of ICT. *Science Technology Human Values,* London: Sage Publications, 37, pp.653–662.

Secomandi, Fernando and Snelders, (2013) Dirk Interface Design in Services: A Postphenomenological Approach. *Design Issues* Massachusetts: MIT Press, 29, pp. 3–13.

Sedgwick, Eve (2003), Touching and feeling. Durham: Duke University Press.

Sedgwick, Eve and Frank, Adam (1995), Shame and its sisters, a Silvan Tomkins reader.

Durham: Duke University Press.

Sefyrin, Johanna. (2011), From Profession to Practices in IT Design. *Science Technology Human Values,* London: Sage Publications 37, pp.708–728.

Stewart, Katie (2009), Seminar on affect. Texas: University of Texas. Accessed on 20

June 2014, from: [www.laits.utexas.edu/cultural\_studies/affect.pdf](http://www.laits.utexas.edu/cultural_studies/affect.pdf).

Thebergé, Paul (2004), The network studio: historical and technological paths to a new ideal in music making. *Social Studies of Science,* 34:5, pp.759–781.

Tomkins, Silvan (1995), Exploring affect: the selected writings of Silvan Tomkins. Cambridge: Cambridge University Press.

Tofts, Darren., Jonson, Annemarie and Cavallero, Alessio., eds. (2000), Prefiguring cyberculture: an intellectual history, Cambridge: MIT Press.

Wallace, Richard, *Alice bot*. Accessed on June 20, 2012 from: [www.alice.pandorabots.com](http://www.alice.pandorabots.com).

Ward, Mark (2010), Has text-porn finally made computers ‘human’? *BBC News Online.* Accessed July 10, 2010 from: [www.news.bbc.co.uk/1/hi/magazine/3503465.stm](http://www.news.bbc.co.uk/1/hi/magazine/3503465.stm).

Whitehead, Colson (1999), *The intuitionist.* London: Granta/Anchor Books.

Weizenbaum, Joseph ([1976] 1984), *Computer power and human reason.* Middlesex: Penguin Books.

Windle, Amanda (2011), *Territorial violence and design, 1950-2010: A human-computer study of personal space and chatbot interaction.* PhD Diss., Wimbledon College of Art, University of the Arts - Surrey University.

Žižek,Slavoj (2008) *On violence.* London: Profile Books.

**Contributor Details:**Dr Amanda Windle is the DigiLab Fellow at London College of Communication, University of the Arts, London. Windle currently runs the DigiLab an interdisciplinary digital laboratory across the areas of media and design. For over ten years Windle taught graphic design and new media as well as theory for fashion, lifestyle and music journalism. Her interdisciplinary research interests concerns design and the social understanding of new technologies. It was through a practice-based doctorate that personal space and artificial intelligent agents were primarily considered, an initial focus on the chatbots (software agents) and their developers in the making of human-computer conversation. As a Principle Investigator, Windle has conducted the design, information architecture and research of two user-interfaces. The two geo-spatial web applications were for environmental research scientists and their web applications respectively focus on threats to endangered species and climate change. Windle is currently researching ageing and digital exclusion supported by the Digital R&D Fund for the Arts - Nesta, Arts & Humanities Research Council and public funding by the National Lottery through Arts Council England.

1. 1 I use the term elevator because Whitehead is an American writer and sets his novel in an American city. Given the main character’s career trajectory in the novel (discussed in example 2) it would seem apt to keep to the nuances of the American meaning and refer to elevators and elevation rather than to lifts and lift. [↑](#endnote-ref-1)
2. 2 Developers of both chatbots and elevators can be programmatically tweaked. Elevators are programmed heuristically to make sure that they service the right floors in a pre-programmed order of priority. Chatbots can be tailored to learn vocabulary from their interlocutors and/or their developer. [↑](#endnote-ref-2)
3. Steve Grand introduces this methodology in his paper on a robot called Lucy (Grand, 2003). [↑](#endnote-ref-3)
4. The following authors take up Judith Butler’s shaping of performativity at this intersection of disciplines, — Allhutter 2011, Rommes et.al. 2011, and Sefyrin 2011. [↑](#endnote-ref-4)
5. Many of the rudiments with chatbots have been re-performed in audio as a way to listen to these silent texts. These are called Audio Rudiments. They can prompt alternative readings beyond the initial thoughts formulated through the ‘ear’ of the researcher. The Drunk Rudiment has several audio re-recordings performed with male and female advertising voice over specialists (in Windle 2011). [↑](#endnote-ref-5)
6. Hugh Loebner, the founder of the Loebner prize for chatbots states on his website that he has been cited as a misanthrope in an interview with the Village Post in the 1990s. Loebner is a public supporter of sex work and donates to charities that work with helping and supporting prostitutes. Loebner is against the prohibition of prostitution and advocates chatbot developments in automated sex and remote sex-talk. Anecdotally, another developer has a record of court injunctions to prevent them entering a previous university workplace for his antisocial behaviour, although this doesn’t necessarily make him a misanthrope (Manjoo 2009). [↑](#endnote-ref-6)
7. The chatbot rudiments cannot fully explore the extensive impact on chatbot developers. Additionally, I researched developers’ behaviours, techniques and processes directly via their online forum groups such as [www.chatbot.org](http://www.chatbot.org) or [www.robitron.com](http://www.robitron.com). I did observe the transformative changes of chatbot developers over the six years of studying chatbots in this way. I used the network of developers to share my research problems and they were very generous and collaborative in responding back. I read many instances of flaming between group members but experienced none myself when posting queries. I did come across a more extreme post of one developer, Chris McKinstry. The belated Robitron member had posted his suicide note on the forum as a thread. The group’s reactions and actions are also archived on Robitron. Moderators alerted the authorities and friends of the developer, but to no avail. Another forum member and collaborator of Chris McKinstry also committed suicide the following week (an MIT doctorate, Pushpinder Singh) again the groups’ reactions are archived. [↑](#endnote-ref-7)
8. Berlant’s critique of the Intuitionist traces the historic in the present. Mae refigures her own methods - of sensory experience - in light of the elevator crash. The crash transforms the way in which Mae will design elevators in the future. [↑](#endnote-ref-8)
9. How constrained or exhilarated can we feel in an elevator? This has been the quest of many an elevator’s redesign. Elevators can be a confining space because they are physically claustrophobic. Some elevators contain mirrors to subdue claustrophobia, whilst glass elevators amplify vertigo akin to a fun fair ride. [↑](#endnote-ref-9)
10. Slavoj Žižek divides violence into four main categories — systemic, structural, objective and subjective. Symbolic violence refers to violence in language, systemic to the enforced violences of economy and state, objective violence operates in the abstract of the everyday, and subjective is of human and non-human violence. Žižek’s categories are problematic firstly because structural terms can be structural forms of violence and secondly, because violence disrupts the meaning of structural categories. I use Žižek’s definitions interdependently for example when violence can be both structural and systemic. [↑](#endnote-ref-10)
11. Some sexbots are listed on [www.chatbots.org](http://www.chatbots.org) as well as within chat rooms and IRC (internet relay chat) networks. [↑](#endnote-ref-11)