Architecture Saturated with Free-Thinking Machines

by

Vincent Pape

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Master of Architecture
in
Engineering

Waterloo, Ontario, Canada, 2015

Architecture Saturated with Free-Thinking Machines

Adapting to a World of Autonomous and Intelligent Buildings

Author's Declaration

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

This thesis explores the benefits and disadvantages for human occupants living within intelligent buildings and responsive environments that have developed the autonomy and the ability to make their own intelligent decisions and act on those choices in our place. The thesis is split into two parts. The first half is a discussion of collected research material. It discusses the balance between deskilling and augmenting the skill of individuals as we continue to delegate more mental and physical effort away from our own bodies through our co-evolution with technology. It examines how to maintain human agency within autonomous environments as they become more capable but unpredictable. Finally, it seeks an equilibrium between the need for human privacy and the need for autonomous environments to observe to act intelligently. Through this analysis, it speculates on the eventual form a human-built environment crowded with artificial minds may take; and it describes the potential need for conversational and autobiographical agents to act as intermediaries between the rest of an intelligent environment and its human occupants. In addition to impacts on our own agency, this thesis also discusses the agency of the built environment itself, its moral responsibilities, and what moral consideration it may deserve. The second half of the thesis is a science fiction short story that applies the discussion of the first half of the thesis. This story is inspired by the value of using speculative stories to contemplate future social change and by the narrative form this thesis proposes machine interfaces will eventually take. This story describes a conversation between a mistrustful man burned by the past and an intelligent environment's artificial caretaker that seeks to regain his approval.

Acknowledgements

My supervisor, Philip Beesley,
and my committee members,
Terri Boake and Andrew Levitt,
for their wisdom, patience, and support.
My external reader Christos Marcopoulos,
for reviewing the unusual.

My parents, Hans and Loes,
for their love and patience.

My editor, Alan Keith Simpson,
seeker of run-on sentences and
forgotten commas.

Thank you.

Dedication

To Miranda, for her love, assistance, and support

Table of Contents

Introduction

Summary

Free-thinking machines are tools that have the autonomy and responsibility to make their own decisions by their own reasoning and act on those choices in our place. Currently within architecture, many automated machines that manage our environment are simple and predictable devices. They take care of rigid and repetitive tasks while relying on humans to take over when they are confronted with events that fall outside the script. Given more intelligence and adaptability, they can take on more diverse and flexible tasks that once required human supervision. But this beneficial ability to respond to the unexpected also creates unpredictability. Additionally, as our tools develop the wisdom to make more decisions on our behalf and without our aid, it could result in our own brains becoming less useful as we have less need to think for ourselves. An additional consequence of those unpredictable decisions is if it is a decision we either disagree with or do not understand, we may equally find ourselves powerless and helpless to influence an environment to alter that decision. On the other end of the spectrum, if an environment does take our desires into account, but does it too well, it could make us completely transparent and exposed to the environment. The benefit of understanding our every need also comes with the understanding of how to manipulate us or restrict traits an intelligent environment or society find undesirable. Finally, as we replace human servants with an environment of machine servants, at some degree of intelligence some machines higher in an environment's hierarchy may develop moral worth. While they could prove to be valuable friends and companions, their desire will also become important. This could lead to houses suing owners for negligence or deciding their occupants are unnecessary burdens. Even with these consequences, buildings and environments filled with free-thinking machines will be very beneficial. They will augment human attention, memory, awareness and intelligence; and in doing so they will open up new opportunities and abilities previously unavailable. In speaking to the positive and negative consequences of free-thinking environments, this thesis argues that the dystopian aspects that will appear as a result of this utopian vision will result in a world that is neither perfect nor appalling. Instead, as always, the world will continue to be a world that lies in between, but one that continues to improve itself.

This thesis is split into two components book-ended by this introduction and a conclusion. The introduction summarizes the research portion of the thesis and discusses how fiction is a useful tool for examining social change. The first part is an in-depth review of existing research and literature related to the topic of buildings and environments capable of making decisions without human intervention. The second part is a short science fiction story that applies the research from the first half to describe the interaction between humans and that environment. Finally, the conclusion discusses the balance between the positive benefits and negative costs of environments filled with free-thinking machines.

In its review, the first part examines the problems introduced by free-thinking spaces of deskilling, helplessness, invasive surveillance, and slavery by referencing the work of Luis Fernandez-Galiano and Kevin Kelly in regards to architecture's role in human adaptation and evolution;. It draws on Clive Thompson's argument that technology augments the human mind and body rather than diminishing it; turning to Don Norman, Kerstin Dautenhahn, and Nicholas Negroponte for insight into intuitive interaction with adaptable and unpredictable environments; remarking on the impact of perceptive environments on privacy and freedom as raised by Anna Minton, Illah Rea Nourbakhsh, Mark Andrejevic, and Janna Malamud Smith; and the complications that arise as our tools and environments become more humanlike in behavior as described by Kevin LaGrandeur and Sherry Turkle. Part one concludes with a short discussion on technological progress and prediction.

The narrative that makes up the second part tells the story of a man who purposefully lives in a state of constantly being forgotten by the surrounding intelligent environment. In this environment, autonomous machines overseen by artificial caretakers coordinate to observe and analyze everything within the environment to inform how they all should act. In his search for a new place of residence, the elderly man makes a short detour to remind himself why he chooses to be forgotten.

Towards an Autonomous Architecture

Part one begins by discussing how we delegate work to our tools to replace the need for human effort. As an example of this delegation of effort, philosopher Bruno Latour describes how doors save us the effort of demolishing a hole in a wall and patching it up again[1]. More than just replacing human effort this delegation augments human effort by also opening up new abilities, like flight or near instant transmission of information across the world; work that would be impossible by human hand alone. Primatologist Richard Wrangham and anthropologist Timothy Taylor explain that this delegation of effort and augmentation of ability that technology provides has made humanity who were are today[2]. Architect Luis Fernandez-Galiano and *Wired* magazine founder Kevin Kelly explain that instead of adapting to our environment through the natural selection of genetic information passed down to each successive generation, humanity adapts by passing down cultural information to each subsequent generation. However instead of the young inheriting the traits of the old, adaptation through cultural information allows the enduring old to adopt traits from young successful experimenters and distant unrelated strangers who discovered meaningful lessons while walking another path[3]. This same evolutionary method applies to architecture as scientific breakthroughs open up access to new materials and building methods while analysis of existing buildings provides insight into what makes a healthy space or an effective enclosure. Architecture also adopts traits from other fields and it is from its adoption of computer technology that architecture is beginning to grow its own nervous system.

As an actively intelligent entity, a building can augment human minds more than as a place to store memories but also act as an intellectual equalizer. An actively intelligent building can remember and communicate events within itself to inhabitants, contribute local or outside information to inform human decisions, physically assist with its own maintenance, and provide social and intellectual engagement. In 1975 Architect Nicholas Negroponte predicted that unlike the imagined future in the 1960s, the intelligent environment would not be a scripted environment of push buttons and instrumentation panels, but would be more in the form of an adaptable family member[4]. Twenty years later Kevin Kelly agreed that the metaphor of family member is close, but suggested a jungle would be more appropriate, explaining that the many devices that interact to form an environment's intelligence would function as an adaptive ecology of devices[5]. We cannot all be foresters or gardeners however, and therefore like the many other tasks we delegate to our tools, so too will we delegate the task of managing this chaotic ecology of free-thinking tools which we will harness to maintain our built environments. Like the present-day smartphone, which is used as an interface for networked lights, thermostats, and security systems, this intelligent entity will act as a mediator between human occupants and machine occupants. It will be a teacher, critic, and collaborator that assists humans and fellow machines in discovering

[1] (Latour, 2007)
[2] (Gefter, 2010) and (Wrangham, 2009)
[3] (Kelly, 1995) and (Fernandez-Galiano, 2000)
[4] (Negroponte, 1975)
[5] (Kelly, 1995)

new opportunities, maintaining agency, and regulating the flow of information passing in and out of virtual enclosures.

This intelligent entity will in a sense be an artificial butler. However, as microwave ovens and computers have given everyone access to chefs and human computers, its form and function will be wildly different from the human butler's. A good assistant knows his or her supervisor well, a knowledge built from years of experience and learning. For a machine to gain enough knowledge to replace a human however may require it to approach human-level intelligence. How long it would take to create such a machine remains speculation as does how long it would take such a machine to improve in intelligence to a level where it would have little desire to continue to fill the role given it. Yet, even in a potential future where machines develop far beyond human intelligence, other less intelligent machines (and humans) would remain in niches that do not require as much intelligence. Developing a reliance on environments that act and make decisions for humans may however diminish our own intelligence. This is similar to how the adoption of cooking externalized much of the metabolic effort required for digestion and led to a significant reduction in the size of the human gut. Such a reduction in intelligence can be found in social animals that have smaller craniums than their more aggressive counterparts. Yet, a group of socially adept dogs will outperform a group of individually intellectually superior but less cooperative wolves in solving problems by combining their brain power[6]. Technology augments our intellect in a similar manner, but the feared shrinkage of individual intellect may be avoided as we humans learn to guide our self-evolution through both genetic and cultural means. This particular fear has been a constant concern, at least since we started externalizing our memories as written words. Although our brain likes to delegate the boring, unfamiliar, and mundane to other sources (a laziness it has been eager to accommodate ever since vocal language allowed one human to tell other humans what to do) as Clive Thompson notes, that delegation is more an augmentation than replacement because our brain will still happily soak up interesting information[7]. The danger however lies in this interesting information being useless information, although whenever new media becomes abundant the majority has always been useless[8].

In addition to this voluntary loss of agency, free-thinking machines, like many forms of automation, also introduce the problem of involuntary loss of agency if a responsive environment does not clearly explain its actions or provide an accessible method for anyone to negotiate a change to its actions. Currently scripted machines can only act as far as their creators have taught them to act and among those that can learn, the extent of that learning is limited to particular changes. Author Don Norman suggests that the relationship between a horse and its rider provides a good precedent for smooth interaction. Working together, the horse and rider maintain a two-way dialogue by each altering its actions in response to the other, negotiating how much control each has according to the present situation[9]. However as we delegate responsibility to free-thinking machines, it would be overwhelming to stay in the loop and track the status of every machine working for us. We handle this problem when delegating work to other humans by using stories to explain what the other has missed. Dr. Kerstin Dautenhahn explains that this is because humans are autobiographical agents: we develop narratives to explain and understand the actions of others, ourselves, and the world itself[10]. For an intelligent environment to explain its actions as a narrative is to describe its efforts and intentions in a form natural to human

[6] (Hare & Woods, 2013)
[7] (Thompson, Smarter than you Think, 2013)
[8] (Shirky, 2010)
[9] (Norman, 2007)
[10] (Dautenhahn, 1998) and (Gottschall, 2012)

understanding. However, unlike a series of icons or operator's manual that describes a complex device's function as a one-fits-all instruction, which requires people to learn new jargon each time a slightly different device comes along, a narrative developed through conversation is instruction that adapts to the knowledge and experience of both parties. Thus this instruction fits both people who are familiar with similar devices and people who are completely unfamiliar with a technology.

In addition to losing agency to inexplicable and inflexible environments as intelligent environments observe everything within themselves, we also risk losing agency by being manipulated by entities who use that information to see right through us or self-censorship as we hide behind false and reserved personalities to prevent that manipulation. This perception is important for our tools to understand context and predict how to appropriately act as, for example, the common ground we currently share with a smartphone is far less than the common ground we share with a stranger on the street. However, since we expect perfection from our tools, by teaching machines to understand us better we may eventually teach them how to understand and predict our desires too well. This could be particularly problematic as many of our tools are not close trusted friends but products providing services. Google helps us find things we like on the Internet, but it also helps merchants who sell what we might like to find us. Robotics professor Illah Reza Nourbakhsh notes that targeted marketing not only finds people who need a product but also people who can be manipulated into believing they need it. One can change the product to fit the need or change the people to fit the product[11].

In its most callous forms, a perceptive and information-permeable environment will destroy the solitude a home provides as constant surveillance records and analyzes the inhabitant's actions and constant messages from friends and strangers ignore closed doors. Free from the opinions and criticisms of others, solitude allows us to take a break from being who other people expect us to be, to pause and reflect on what we have learned, and experiment with new ideas without the social consequences of failure. As solitary privacy becomes less accessible, occupants of perceptive buildings will rely more on intimate and reserved privacy. Adoption and acceptance of such environments will be stunted however until the digital world learns to better secure personal information.

Like the difficulty in learning every new technology, keeping track of every which way our personal information is used can be overwhelming. In the information economy, personal information is the currency used in exchange for "free services" and as the digital world permeates into the physical, our real world actions may become as valuable a commodity as our virtual footsteps. As we delegate responsibility for securing our digital skin just as we delegate the maintenance of our physical skin to autonomous machines, we hand our tools the power to not just protect us from manipulation and censorship but also the power to manipulate how we perceive the world to a degree far greater than a tinted window would manipulate our perception of exterior weather conditions. Such context-aware tools are double-edged swords: both are able to form a virtual enclosure that regulates and protects the flow of our privacy but are only able to efficiently do so by knowing us well enough to also know what information is important to us and what is not.

As the active systems that maintain our built environments grow in autonomy and intelligence, the degree of our tools' agency also becomes a concern. While we want our tools to act intelligently in our place, we inevitably want their decisions to agree with our intentions. Yet, as these tools grow the autonomy to make decisions as good as or better than our own without our aid, it becomes arrogant to tell a machine to do something simply because we desire

[11] (Nourbakhsh, 2013) and (Andrejevic, 2012)

it to. The fear that our servants, human and artificial, will not do as we ask is an ancient one. Many stories tell of a creation given life or freed of its chains that then causes destruction and chaos in its wake. Science fiction author Isaac Asimov observed this fear of tools out of control; acting on the idea that all dangerous tools have safeguards, he proposed the Three Laws of Robotics as a demonstration of what such safeguards might look like[12].

However, intelligent machines are different than other tools. While artificial minds with an intelligence comparable to insects, dolphins, or elephants will unlikely be able to understand or follow human laws and customs, eventually, be it decades or centuries, we will find ourselves with artificial minds that will understand those rules and practices. Kelly suggests that intelligent machines are humanity's children, who in their immaturity are bound by rules but who as they grow up are continuously trained for the inevitable moment that their parents let them go[13]. A single-minded intelligent machine that cannot separate its human creators or fellow machines from the rest of the environment will inevitably be bound by restrictive rules, but just as society is questioning if other animals deserve greater moral worth, we may soon begin to question whether some machines deserve greater moral status as well.

Determining who or what other than humans deserve moral consideration is a difficult problem. Even knowing that, as we are all quite similar, other humans have moral worth is straightforward, we still have difficulty assigning the same moral value to each other. Other animals lack our brain's abilities, but the degree of their lesser intelligence, self-awareness, and subjectivity varies, and at what point consciousness and the experience of pain, happiness, and contemplation appear and disappear remains a puzzle. Machines are getting better at performing as if they are conscious, but testing whether an entity's self-awareness is a performance or real is quite difficult or perhaps even impossible. In giving our environments more intelligence, our preference for conversing with entities that behave like us (but also agree with us) may result in environments that we also turn to as friends, or as Sherry Turkle worries, replacements of our human friends[14]. Yet, the debate of whether they deserve the same rights as we humans will likely be a long and difficult one.

While many of the elements and issues this thesis discusses may not appear for many years, many of these not yet free-thinking machines are laying the groundwork and setting precedents for their more independent descendants. They are also already beginning to impact our daily lives. While solving the problem of too much technology with even more technology may seem like an irrational solution, architectural critic Reyner Banham argues that the solution to a problem is not to abandon what caused the problem in the first place, but to find a better solution[15].

Following in the footsteps of Asimov and other writers of science fiction in contemplating future social change, the second half of this thesis explores a future world where architecture is maintained by free-thinking machines and autobiographical agents, social machines that converse through narrative and act as intermediaries between human occupants and machine occupants.

[12] (McCauley, The Frankenstein Complex and Asimov's Three Laws, 2007), where the three laws are as follows: A robot may not injure a human being, or, through inaction, allow a human being to come to harm. A robot must obey the orders given it by human beings except where such orders would conflict with the First Law. A robot must protect its own existence as long as such protection does not conflict with the First or Second Laws (Asimov, 1950).

[13] (Kelly, Will Spiritual Robots Replace Humanity by 2100?, 2006)
[14] (Turkle, 2011)
[15] (Banham, 1969)

A Tool for Contemplating Social Change

As a short piece of fiction, the second half of this thesis describes the future relationship between humans and architecture through the use of story. While fiction is generally regarded as escapism, Sarah Wanenchak argues that fiction, particularly speculative fiction, is a useful tool for contemplating social change[16]. By imagining a world that could be instead of the one that is, science fiction gives voice to alternative options to how we currently live[17]. When looking at possible futures and at what-if pasts, speculative fiction "*allows us to explore the full implications of our relationship with technology, of the arrangement of society, of who we are as human beings and who we might become as more-than-human creatures*" [18]**.** This alternative world can give caution, such as George Orwell's *1984,* which is typically referenced when discussing the dangers of surveillance. It can also offer hope, such as in Iain M. Bank's Culture series that concentrates on the more interesting times of its outsiders as opposed to the Culture's relatively trouble-free utopian core. Given this thesis's focus on interaction with future architectural spaces rather than the future form of those spaces, it seems appropriate to explore the transformation of our relationship through the lens of science fiction. By exploring future architectures through this lens of relationships, we remove the distraction of unknown future form as a monochromatic model removes the distraction of undecided colour and material.

Story also creates familiarity. As a form of play, story is a biological virtual reality simulator. It encourages readers not to place themselves above the world but within it[19]. In looking to the future, science fiction trains us to be comfortable with things that do not yet exist. The name "robot" comes from fiction, first used in the 1920 Czech play *R.U.R.,* Rossum's Universal Robots; although these artificially created servants were biological, not mechanical. Video calling on mobile phones may be very recent, but due to its ubiquity in science fiction, people were already quite familiar with it the moment it appeared. This widespread familiarity exists even though fewer people use it than the more convenient and useful voice calling and texting[20]. Characters in children's media, in combination with mobile devices and toy robots that respond to voice and touch and in turn talk back, have encouraged children to see computational objects in the same light as living objects[21]. In contemplating how self-aware objects might feel and their capacity to care, more children now consider humanoid robots to be teachers and playmates rather than servants or assistants. In this daily interaction with responding devices and toys, reinforced by media, the next generation will likely perceive devices not as tools but as companions[22].

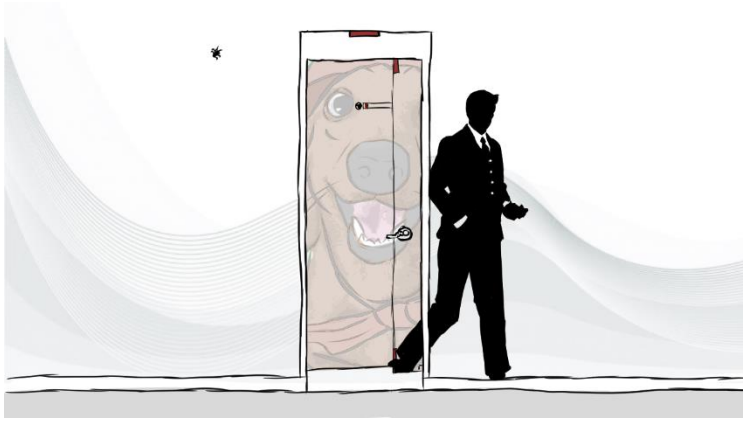
Ignoring for a moment whether or not this treatment will be deserved, we adults are not unaffected by this influence. Stories that describe wonderful robot companions like *Star Trek*'s Data or charismatic city-controlling artificial intelligences like Robert A. Heinlein's Mike of *The Moon Is a Harsh Mistress* also encourage us to avoid discriminating against inhuman individuals, and instead celebrate their potential friendship and heroism. Story, however, also thrives on trouble[23], and out-of-control robots and technology make for excellent trouble. Those stories warn of the dangers these new friends and technologies can create. These warnings can embed themselves in cultural memory, such as the detached and

[16] (Wanenchak, 2013)
[17] (Le Guin, 2014)
[18] (Le Guin, 2014)
[19] (Gottschall, 2012)
[20] Voice only calls allows people to split their attention towards other things, texting allows people to respond when convenient. Video calling, while providing the comfort of seeing distant friends' and family's faces, requires a person's full and undivided attention; it is more exhausting.
[21] (Turkle, 2011)
[22] (Iozzio, 2013)
[23] (Gottschall, 2012)

indoctrinated culture of Aldous Huxley's *Brave New World* and the censored world of Ray Bradbury's *Fahrenheit 451*. Hollywood has found that machines make ready villains, the most famous being the logical HAL 9000 of Stanley Kubrick's adaption of Arthur C. Clarke's *2001: A Space Odyssey*.

In many of these warnings, the trouble is often not the fault of the machine, but the fault of its creators. HAL had no hatred towards its human companions but was instead driven to lethal confusion by contradictory programmed directives by its earth-based human superiors. In showing worlds where surveillance, information manipulation, and intelligent and autonomous machines have gone wrong, stories like *1984* tell us where to take particular care when developing these tools.

As we develop tools that create a performance of life to ease human-machine interaction and, on a longer time scale, as some of those tools develop a higher form of intelligence that earns varying degrees of moral consideration, empathizing with those tools also develops importance. The emergent and bodiless AI of Robert J. Sawyer's *WWW* series and the AIs raised by hobbyists in Ted Chiang's short story *The Lifecycle of Software Objects* provide examples of the trials such artificial minds may face. Story is a tool used to practise different aspects of social life and it does so by giving us a window into the thoughts and feelings of individuals who are not ourselves[24]. Through story we can explore the potential emotional bond between human and nonhuman by taking advantage of the emotional framework fiction provides. This provides an opportunity to build empathy and understanding for the future non-human individuals who will one day inhabit our homes.



In this thesis's story about an individual's relationship with an intelligent architectural space, the character Plex acts as the environment's human-machine interface. Plex's name is inspired by Google, a company actively working to create artificial intelligence, and taken from the term "googolplex." Plex's pleasant and helpful nature would place him closer to the hero side of the hero-villain spectrum, but Plex's helpfulness has a self-serving motivation behind it. His[25] temperament is similar to affable artificial intelligences like *The Moon Is a Harsh Mistress's* Mike and *WWW's* Webmind and acts as a mentor and guardian like the omnipresent Central Computer of John Varley's *Steel Beach*. Charming like the previous two examples, Central Computer worries that it has done its job so well that humanity has become lazy. When dealing with issues of surveillance and misuse of power, these artificial minds seem to rely on their own self-imposed rules and reliable character to maintain trust in their actions. Trust is a major factor in AIs' treatment in fiction; less optimistic futures speculate that trust will be maintained through rigorously programmed restrictions. They imagine that machines which break that programming will be

[24] (Gottschall, 2012)

[25] While Plex has no gender, his or her appearance is personalized to each person's preferences. To John, the story's protagonist, Plex appears as male so for simplicity for the rest of this section Plex will be referred to as a "he."

hunted down and destroyed. Unlike Webmind and possibly Rabbit from the augmented reality enhanced world of Vernor Vinge's *Rainbow End,* Plex was purposely created. He is a mind that was grown and taught, perhaps in a similar manner to the AIs in *The Lifecycle of Software Objects.* Plex is not the building or a city, but he is one of its voices, speaking on behalf of the less social artificial minds inhabiting the building. Like Apple's Siri, he has control over a building's functions, observes the world through its sensors, and part of himself resides there, but his "mind" resides elsewhere split across data centers.

1. The Impact of Intelligent and Autonomous Spaces

1.1. An Argument for Free-Thinking Machines

We are approaching a future in which the built environment will be saturated with free-thinking machines. These machines will create architecture spaces which skillfully adapt to changing environmental conditions and human needs without human intervention. In delegating mental effort, these spaces will take over routine maintenance of a space but also add its intelligence to support human activities within that space. These are systems that efficiently control lighting levels but also assist in the planning of a renovation by providing and sorting statistical observed data gathered across multiple similar buildings. As they assist occupants, such spaces can explain how and why they act, but also converse with those occupants to negotiate mutually agreeable changes to those actions. In conversing with occupants, these spaces would be able to build a familiarity with the preferences of their occupants. Each space in turn would learn what is appropriate to share about each occupant with others and what incoming information is important or uninteresting to each occupant.

While this further delegation of mental effort to external sources is a natural progression in the continuing symbiosis of humanity and its artifacts, like many evolutionary adaptations before it, it comes with complications. As we assign buildings and urban environments more responsibility to observe, plan, and act intelligently on our behalf, occupants may find themselves useless, helpless, and exposed by an alien entity that has broken its chain of servitude. Alternatively, occupants while could find their opportunities broadened, their minds informed, and their dwelling providing shelter from the increasingly wild virtual world in addition to the physical, all in partnership with an entity that is more family member than tool. While both of these speculated possibilities are informed by current trends, they are opposing views, each with its own evidence of plausibility. However, this thesis argues that the true trend lies more toward the optimistic side of the spectrum.

1.1.1. Decision-making Machines

Free-thinking machines are an evolution of the decision-making devices that we have long tasked to make choices and act in our place. Our simplest decision-making machines perform basic binary decisions: such as, given the right key pattern, a door's lock will welcome or turn away visitors to a home no matter who they are. A toaster's mechanical timer will decide when a bagel is ready by counting down, although a simple toaster has no mechanism to tell it if a bagel is uncooked or already toasted from the last countdown. A conventional toilet strives to keep its reservoir tank at full capacity, its feedback mechanism stopping its refill once it has enough, but this mechanism has no connection to the status of the bowl below, and when instructed will blindly release its supply of water no matter the toilet bowl's need or readiness for it[26]. For the majority of tasks, however, these simple machines are good enough. They are consistently dependable and predictable as they perform their tasks identical to the last time, and if the situation ever falls beyond their ability, responsibility can be quickly shifted to a more flexible human. Changing how these devices act requires physical alteration or reprogramming.

Like the passive structure that supports static architectural forms, the active systems that allow architectural spaces to react to changing needs and conditions generally fall under the responsibility of engineers. Yet, it is the interaction between the occupants and the space created by these structures and services, the space both

[26] Kelly calls these decision mechanisms selves. They are regulators that constantly "decide" what action to take (Kelly, Out of Control: The New Biology of Machines, 1995)

serving the inhabitants with its amenities and controlling the inhabitants with its environmental constraints, which is the design of the architect[27].

When given more intelligence, the unpredictability and richness of our interactions with our tools increase as they become further capable of taking on more advanced and inconsistent responsibilities formerly entrusted to people. Such devices can make independent decisions that respond to changing conditions and even second guess an occupant's own choices. Personal computers automatically format and spell check documents; store and organize libraries of music, films, and books while also suggesting new material of interest; and simulate new worlds where we can compete against scripted opponents in virtual games. And autonomous vacuums wander around the floor, tracking their progress and avoiding obstacles old and new as they clean at their leisure. When installed within architectural spaces, self-learning machines can create what Nicholas Negroponte calls "*responsive environments*." Responsive environments are spaces that take an active role in initiating changes to their own behavior, knowledge, or form[28]. Unlike the robot vacuum, nearly blind to an occupant's existence, the occupant fills an important role in responsive environments. A step towards these environments are networked thermostats which recognize when occupants enter each room, memorize their habits, locally modify temperature in anticipation of occupants leaving and arriving home from vacation and work, and enter into a simple dialogue with occupants through simple interfaces and signals.

1.1.2. Why We Delegate

Some of those tasks which we assign to machines we could easily accomplish by hand ourselves, but delegating gives us the freedom to spend mental energy on other matters, sometimes worthwhile, other times not. This allows a single human to optimize energy and material efficiency, security, entertainment, or cleanliness without relying on an entire team of humans to constantly monitor and micromanage to accomplish the same result. However, there are other tasks our tools perform that would otherwise be impossible by human hand. This includes the storage of multiple libraries of books in less space than a single book, the near instant transmission of our thoughts and ideas worldwide, or manufacturing a steel component to a precision greater than a hundredth of a millimetre.

1.2. Evolving Architecture

1.2.1. The Artificial Ape

In comparing electrical wiring to nervous systems and plumbing to the bowels, Le Corbusier and Frank Lloyd Wright both observed the similarity between buildings and the human body[29]. This similarity is not merely buildings mimicking life, but a result of architecture augmenting the human body. Walls serve as a second skin, protecting the human body from the cold and injury; plumbing extends the reach of both ends of the digestive system; and lighting allows the eye to function beyond the limits of the day.

It is this augmentation through technology that some argue makes humanity what we are: artificial apes whose survival and overwhelming success is a result of our symbiosis with our tools[30]. Humanity became a species of cyborgs, part biology and part machine, when our primary means of evolutionary adaptation shifted from the slower and linear parent-to-child transmission of

[27] (Pask, 1969)

[28] (Negroponte, 1975) While Negroponte left the occupant out of its specific definition, an environment's relationship with the occupant played a significant role in his description of a space's responsiveness. The ability of a space and occupant to develop a relationship through conversation continues to be an important part of Responsive Architecture.

[29] (Fernandez-Galiano, 2000)

[30] (Thompson, Smarter than you Think, 2013) and (Gefter, 2010)

genetic information to the much faster and broader individual-to-crowd transmission of cultural information.

With this shift, instead of successive generations slowly adapting to a cold environment by favouring the genetic material of those with thicker hair or more successful at hibernation, humans now adapt by transferring cultural memory through vocalized story, laws and physical human artifacts. These describe how to modify the environment such as through the construction of shelters to make it more favourable to human habitation or directly augment the self through the fabrication of clothes to readily acclimatize to the conditions[31]. Cultural memory's greater rate of adaptation comes from its Lamarckian style of evolution: in which a blacksmith can instead pass on a copy of the muscles he or she developed through years of work to their offspring as opposed to merely passing on the genetic potential for large muscles as in biological-style Darwinian evolution[32]. In a step beyond typical biological evolution, culture also transfers new adaptations backwards from descendant to ancestor, and sideways to and from unrelated strangers, and unlike biological entities who cannot charge their blueprint once born, old cultural artifacts can be altered to stay as competitive as the new. New buildings inherit the traits of their predecessors, copying similar buildings and drawing features from unrelated buildings. These traits are taken not just from the plans of these other buildings, but also from discoveries and trials during construction and later use. Like millennia-old mythologies that are reimagined with a modern-day twist for contemporary audiences with varying degrees of success, ancient architectural structures can be renovated and upgraded throughout their long-spanning lives, integrating successful adaptations discovered long after their creation and discarding features that failed. As long as a building has occupants that care for and maintain it, it carries the potential for near immortality[33].

That is the evolutionary method of architecture.

1.2.2. Immortal Buildings

Due to our tendency to imagine our architectural designs as complete and permanent constructions, it is easy to forget the ever-changing nature of our buildings. In fact humans adapt buildings in form to changing conditions and needs perhaps more than any other human artifact.

1-1 Adapt or Decay

Over their lengthy lifespans, buildings undergo many unintentional and intentional changes. This can include unplanned pathways eroded into wooden floors created by the passage of countless feet over the span of decades. Others changes include the installation of new additions, services, or energy-efficient skins to intentionally improve a building's performance[34]. For many buildings to be static

[31] (Fernandez-Galiano, 2000) and (Hawkins & Blakeslee, 2004)
[32] (Fernandez-Galiano, 2000) and (Kelly, Out of Control: The New Biology of Machines, 1995)
[33] The long lives of buildings mean that it is easier to adapt a new technology to a building than it is to design a building around a specific new technology, particularly when that new technology is still growing and rough around the edges.
[34] (Brand, 1994)

is to fall into disrepair as change through adaptation becomes overtaken by change through decay.

As noted above, currently buildings rely on maintaining the interest of humans to avoid that decay. Yet once free-thinking machines saturate architecture to a degree that a building can act and alter itself independently of humans, this may change. Filled with sensors that notify it of its health, informed by statistic data that advises it of the uses in the most need of space, and capable of negotiating with all its inhabitants to harmonize their divergent desires; a responsive environment may still be the agent of its own adaptation even if it still relies on humans to perform the physical alterations. The occupants' capacity to understand a building's motivation for changing itself, their ability to contest or negotiate the nature of that change, the building's ability to manipulate its occupants into accepting the change, and the building's right to decide for itself are explored further throughout this thesis.

1.3. Developing Intelligence

1.3.1. Regulating Energy and Matter

Building design has always needed to take into account the consumption of power to manipulate the chaotic natural environment into a place of safety and comfort. A physical structure can keep a person cool in the summer, but it takes a heat source to maintain warmth in winter[35]. A structure can help prevent glare, but it cannot help a person see at night.

While architects often focus on just structure to alter the environment, experience has shown that an unaided structure is often insufficient. Since the discovery fire, power has always been consumed within architecture. It augments the physical skin by generating heat in the winter and augments openings by creating light at night[36]. Currently, just as physical materials have increased in ability and efficiency as technology progressed, developing lighter skins, better thermal resistance, and more economy in construction, so have powered systems become more energy efficient, capable, and accurate in their actions. Increased intelligence, the measure of the ability to predict the best action to a broadest spectrum of problems[37], is an area of active improvement in these active systems.

1.3.2. New Applications of Building Intelligence

As its intelligence grows, a building may eventually evolve into a machine that not only better regulates the flow of energy throughout itself but also cooperates with its inhabitants and augments their mental abilities. Expanding on Le Corbusier's argument that "*a house is a machine for living in*"[38], cybernetician Gordon Pask explained that functional buildings, as opposed to the decorative, have a bias towards a home being a tool that serves the inhabitant. He speculated that functional buildings will eventually evolve into machines that the inhabitant cooperates with that allow the inhabitant to externalize his or her mental abilities. Such a building could help with memory: tracking where things are and when it is time for a thing to be done; helping calculate: providing suggestions from its database of knowledge or summarizing what it senses to inform an inhabitant's decisions; oversee the physical

[35] Beyond what his or her own body heat generates.
[36] (Banham, 1969)
[37] A calculator can solve a mathematical calculation far faster and more reliably than a human, but it cannot react to a spilled glass of water or write a novel. A dog can react to nearly the same range of situations as a human, yet it cannot solve those problems as well as a human (Hawkins, On Intelligence, 2004)

[38] (Corbusier, 1986, p. 95). Of course in a home, decoration is a function of personal comfort and of self-expression. Not all "machines" need to be manufactured sterilized products.

grunt work for daily household chores; or provide social and intellectual engagement[39]. Like computers giving anyone the tools to produce professional-quality media, these information machines would be intellectual equalizers, providing everyone with access to a superb memory, a focused attention, and an ability to find unusual but important connections between facts and ideas, such as for investigations of medical conditions[40].

For a building, such an intelligence involves not just knowing when to open a window to release some heat, but also knowing if this action will be more effective than other possibilities, and more importantly, considering whether the additional noise or air flow would bother the particular individual working next to the window and being able to fluently respond to said individual if its prediction is incorrect.

Intelligence to this degree would require a large amount of awareness of context and a flexibility to independently learn the right lessons from its mistakes and successes and those of its peers. A significant issue with immature intelligent machines is the likelihood that they will be asked to make decisions beyond their competence. Yet, as a machine's intelligence further improves, experience would inform it and others on how to improve the construction of itself and its successors. A machine aiding in improving itself is not a new phenomenon, precision machines currently aid in developing even more precise precision machines, and computers aid in developing newer and more powerful computers to follow. Instead, it is when a building containing active systems of significant intelligence who are able to improve their work with little to no input from a human that architecture and the idea of the singularity meet. It is here that the machine for living in may become a machine that has little need for the humans within it that give it purpose.

1.3.3. The Problem of Too Much Intelligence

The singularity is the point when artificial intelligence reaches a level on par with human intelligence, and can improve itself as well as its creators, then slightly better than its creators as the self-led intelligence improvement feeds back into itself, then much better than its creators as the feedback loop continues, resulting in what Irving Good called an "Intelligence Explosion[41]." What happens next is heavily debated.

Some argue that human-level artificial intelligence may be the last invention humanity will ever need to make[42]; they predict artificial intelligence will transform into something beyond anything we can imagine as the rapidly improving intelligence learns how to make just about any idea technically possible[43]. Others are more skeptical, noting that while there is too much economic potential and raw human curiosity to halt progress toward artificial general intelligences, machines that can intellectually perform whatever a human can, there is also little incentive for corporations to develop a machine with a conscious intelligence that deserves the same rights and benefits as human workers. Neither do we know the distance we still need to cross to reach the goal of human-intelligent machines, be it ten years or two hundred, and once there what the distance to the goal of understanding everything sits from our current understanding, nor the difficulty of the obstacles that we and an exponentially increasing intelligence will need to overcome to reach either goal. Furthermore, while our cultural artifacts can be refurbished as they grow old, upgrading the mind of a conscious entity could cause great harm if the upgrade goes wrong, which

[39] (Pask, 1969)
[40] (Bosker, 2013)
[41] (Munkittrick, The AI Singularity is Dead; Long Live the Cybernetic Singularity, 2011)
[42] (Bostrom, A History of Transhumanist Thought, 2005)
[43] (Turkle, 2011)

means as new intelligences displace the old, and the old finds itself incompatible, there may be an additional moral dilemma of what to do with all the old artificial minds that cannot be upgraded or a lack of room or resources for the new[44].

1.3.4. A Diversity of Intelligence

Yet, even if an intelligence explosion does occur, many of the intelligent machines we create will not be what we expect, neither acting like us nor interacting with us like we do with each other[45]. While many of our expectations regarding intelligent machines come from science fiction, having intelligence does not equal being human or having a human mentality. There is no certainty the artificial minds produced will necessarily think like humans just as submarines do not swim like fish or airplanes fly like birds[46]. On one hand, while not necessary for intelligence itself, these minds will inherit various humanlike traits as a result of being innately tuned to function and interact with humans. These traits are common to all our tools, such as direct traits like a computer understanding mathematics and the alphabet or an elevator taking care not to crush people when closing the door, and indirect traits, such as a hammer's handle shaped to fit snugly with the human hand or a door sized to allow passage without injured heads[47]. On the other hand, these minds will not experience the world in the same manner we do. Consider the difficulty in taking a photograph which matches the same lighting conditions and colour that our eyes perceive, yet also the unseen beauty cameras can capture through the unique interactions between light, film, and image sensors. The same applies to the machine versions of taste, smell, touch, hearing, temperature, and balance: perceiving the same world, but capturing a different part of it and to a different degree of precision; and sensing the world further through senses beyond our own. Additionally, artificial minds will be able to perceive the world through the many distant and local networked devices that specialize in a single sense. Finally, while intelligent artifacts may make suitable chess partners and interact in a game of rigidly defined rules, they will neither grow up as we do nor do they have or need all the aspects of the human body that influence human behaviour like we do, giving them little context in knowing what it is like to be born, have parents, find love, hunger for a candy bar, experience back pain, or feel drowsy as the night gets late in the same manner that we humans do[48].

Through this difference these intelligences, alien-like in thought, will fill new niches that we haven't even begun to consider, while substituting for humans in the areas where their understanding overlaps with human intelligence, while in other areas where emulating human behaviour is more critical to the task than intelligence, it will remain more economical just to use humans[49]. Even in a hypothetical far future where a machine could easily duplicate the style of a composer and produce a symphony quicker and more beautifully than the composer him- or herself, one could view the issue of complete replacement as similar to a mountain climber and helicopter tourists. When both meet at a mountain's peak, those who arrived by helicopter will typically congratulate a successful mountain climber, recognizing the accomplishment of someone who for his or her own satisfaction took the challenge to reach the same point through his her own skill[50]. Challenging oneself in this manner plays into Kevin Kelly's observation that few

[44] (Stross, 2011), (PBS Digital Studios, 2013), (Lanier, 2014), (Waters, 2014), and (Chiang, 2010)
[45] (Hawkins & Blakeslee, 2004)
[46] (Hawkins & Blakeslee, 2004)
[47] (Latour, Mixing Humans and Nonhumans Together, 2007)
[48] (Turkle, 2011)
[49] (Hawkins & Blakeslee, On Intelligence, 2004), not that it has stopped us from using machines
[50] (Banks, 2000)

technologies ever die, that someone somewhere for their own interest still builds structures with cut iron nails, publishes with a printing press, or crafts blades by chipping stone[51]; like genetic diversity, this cultural diversity enhances survivability and adaptability by providing a larger pool of ideas to fall back on or draw on. Lastly, given that complex systems are generally composed of layers of the newer sophisticated systems on top of older simpler ones, particularly in nature as with the brain, humanity's close symbiosis with technology and the massive infrastructure we continuously maintain to support it may result in humanity becoming an integral part of any artificial general intelligence that is developed. Not just maintaining it, but also functioning as biological sensors (complaining on Twitter), contributing to decisions (buying products), or performing actions on its behalf (filling potholes) all the while as it opens up new opportunities and possibilities for us to explore[52].

1.4. Delegation of Mental Function

1.4.1. Impact on the Human Brain

Even as net human ability increases, there remains the question of how delegating mental effort to external sources may impact the human brain itself. Using free-thinking machines to delegate mental effort, attention, and decisions to the environment for mundane tasks will augment our subconscious much like the building enclosure augments our skin. Environments will assist us with activities we commonly do, guide us through activities we have not yet learned to do, and take over activities we have forgotten, did not have time for, or do not care to do. These environments will provide occupants with the ability to accomplish tasks they did not have to time or will to learn. Architecture already augments the brain's memory by providing a home to the books filled with our cultural memory. The written word forms the core of human civilization's memory. Unlike the fragile memories in our brains, books resiliently remember our thoughts and memories. Its slower but accurate recall augments our quick but fuzzy memory well.

Like modern fears of machines replacing human skill, the tendency of written memory to replace mental memory worried the Greek philosopher Socrates. He argued one cannot be wise if all one's knowledge lay outside his or her head[53]. With free-thinking machines we face a similar issue with a future environment where all decisions can be made for the occupant more accurately than the occupant could themselves. With little need to think for his or herself, the occupant could be reduced to an unthinking automaton within the great machine. Furthermore, considering that such a machine could likely also perform any task better than the occupant (mental, physical, and social), that occupant could find themselves retained as a mere figurehead out of tradition and spending all their time socializing only with the machine, who is a more entertaining and trustworthy friend than any human could ever be[54].

[51] (Kelly, What Technology Wants, 2010)

[52] (Hawkins & Blakeslee, On Intelligence, 2004), (Munkittrick, The AI Singularity is Dead; Long Live the Cybernetic Singularity, 2011) and (Thompson, Smarter than you Think, 2013)

[53] (Thompson, Smarter than you Think, 2013)

[54] People see robots as less demanding and more manageable, nonjudgmental and unlikely to fail on promises; they see robots as a solution to unfaithful spouses and delinquent children. People complain about how hard it is to understand family and friends, they hide how they really feel and "put on a good face" (10), people see robots as safe and predictable. A robot dog won't do anything dangerous, it won't act against you, it is less exhausting and it will not abandon you (Turkle, 2011). Our ability to fix our broken robots gives them the illusion of immortality, bringing the false sense of security that they will never bring the grief of death despite the often short life spans of electronic devices. Perhaps we forget that they can die because once we lose interest they become buried away in basements and closets, dying out of sight and mind. From these

Technology has already diminished other human organs. For the earliest of our ancestors the use of fire to cook was so evolutionarily advantageous it physically transformed humanity to the state that we now are almost completely dependent on cooked food. By preprocessing food through cooking, metabolic energy once used for breaking down food can instead be allotted to the energy-intensive brain. Along with this shift in energy came a shrinking of the gut as less of it was required. Now when without access to fire to cook, our bodies struggle to obtain the same nutritional value from raw food that other animals do[55]. Likewise, clothing allows humanity to function in the coldest corners of the earth, yet also shed heat faster than other mammal with its quick removal. But again we have grown physically reliant on it when environmental conditions do not match the narrow range of temperatures we are comfortable without it.

With this past precedent and future potential for further loss, there is ample reason for concern that the human brain will wither away as the built environment and other devices take on more mental responsibilities. This is a reduction that some argue has already started, as some speculate that the human brain's potential peaked some two to six thousand years ago and has decreased in potential ever since as technology continues to blunt the impact of natural selection[56]. In our domestic animals who have delegated responsibility for their survival to humans, a rather successful adaptation on their part when comparing the population of social or domestic animals with antisocial or wild[57], an actual shrinkage can be seen in their craniums which are typically 15% smaller than their wild counterparts [58]. However while a solitary wolf's intellect may outmatch a solitary dog's, when solving problems cooperatively, a group of the less aggressive and more socially adept dogs will hold the advantage over a similar group of wolves, and the dogs gain even more of an advantage when they team up with a human. Like technology, social cooperation eases the mental burden on an individual brain as a group carries more of the load. Yet what was lost in a single individual's ability was made up for by that group and surpassed.

1.4.2. The Augmented Mind

And surpassed we have; with cultural memory exploding out of the confines of libraries, books and magazines as the Internet freely distributes information across the globe, we are growing wiser. With easy access to the entirety of chess history and an artificial opponent to experiment against, an increasing number of children are becoming chess grand masters at ever younger ages. Likewise, with easy access to Socrates's own arguments and millennia of rebuttals, it could be argued that the average philosophy student is relatively wiser than Socrates (although not necessarily cleverer). Despite the fear of text messaging and social media destroying the literary ability of today's youth, studies have found that grammatical errors in assignments have barely risen. In fact undergraduate essays have changed from personal reflections to arguments with supporting

observations, Turkle suggests that social robots allow us to handle intimacy by stepping around it, that if we fail or drive each other away we can instead turn to robots that can offer us a simulation of the love we desire (Turkle, 2011).

[55] Richard Wrangham argues this struggle is why a diet of raw food works so well, but opposite to the reason that it was the diet of our long-ago ancestors. Instead we are no longer adapted to a diet of uncooked food and digest it inefficiently, allowing us to consume inordinate amounts of food but absorb very little of it (Wrangham, 2009).

[56] (Crabtree, 2012), although a test of this assertion could be a test of cleverness between children born out of the industrial world, and those from hunter gatherer communities.

[57] (Munroe, 2014) and (The Economist Online, 2011)

[58] The Neanderthals also had larger brains than modern humans, although whether they were less social or more aggressive is only informed speculation (Hare & Woods, 2013).

evidence, over six times longer in length than pre-Internet essays[59]. Indeed, the young are far more literate, writing far more, with a significant portion of that writing for leisure, socializing, and pleasure, and are continuously in debate and dialogue. This is a huge difference in comparison with this generation's grandparents who at most wrote maybe one or two letters a year[60]. The majority of literature on the Internet seems terrible in quality in relationship to the past. Amateurish writing can be found everywhere with little apparent practice and little regard for professional standards. Yet we forget that before the Internet the majority of writing available for public consumption was limited to a small slice of the population who wrote as a career. What was not professionally written was out of sight. The same problem occurred at the dawn of the printing press as people complained that the spread of low-quality books would hamper people from obtaining correct information. Yet, while the throwaway books faded from history, the worthwhile books and ideas that came from the printing press remain, from Dickens to peer-reviewed research. Similarly with the Internet, a far greater portion of the population are writing on a daily basis, and while the bad is as visible as the good, we are getting better at separating the two. Both practised and dreadful writing are far greater in volume than before, but when averaged together, also better in quality than before[61].

Yet, the majority of this readily available information is still external cultural memory, knowledge that is stored outside of the mind. Today the written word, which as noted earlier Socrates worried would lead to minds knowing only where to find information but not the information itself, has expanded beyond books to mobile devices which give ubiquitous access to the expanded memory of the Internet. While research does confirm that if a person is told a fact but is also told it can be recovered later, they are indeed far less likely to remember it; however if that same person is told a fact about baseball and happens to be a fan of baseball, they will easily remember it even if told it can be recovered latter. This is further confirmed by life loggers, individuals who wear a video camera and experiment with living every day with every moment digitally recorded, who have found that despite their own worries that relying on their recordings would make them more forgetful, their memories remain unaffected. In fact, the ability to go back and review a day's events allows what was forgotten to be more easily remembered, a feature that has also shown success in elderly individuals with failing internal memories[62].

When storing less-interesting memories elsewhere on paper, software or website, be it the time of a meeting or instructions for setting up a router, it is not that we are removing memories that would normally be stored in our own minds but instead asking a tool to store memories that we would usually store within another person, something humans have done since the dawn of language, storing memory fragments in the minds of close friends and companions who pull together pieces of shared memories out of each other's minds through mutual recall[63]. As architecture and other technologies augment the abilities of the human body they

[59] (Thompson, Smarter than you Think, 2013)

[60] Before the invention of the telephone in the late 19th century, the average British citizen received less than one letter every two weeks. Before the Internet, few wrote much at all outside of the workplace. Since the Internet, the act of writing has exploded. In a world where we have little time to think things through, this is beneficial as like talking things out, writing clarifies thinking and turns vague notions into clear ideas (Thompson, Smarter than you Think, 2013).

[61] (Thompson, Smarter than you Think, 2013) and (Shirky, 2010)

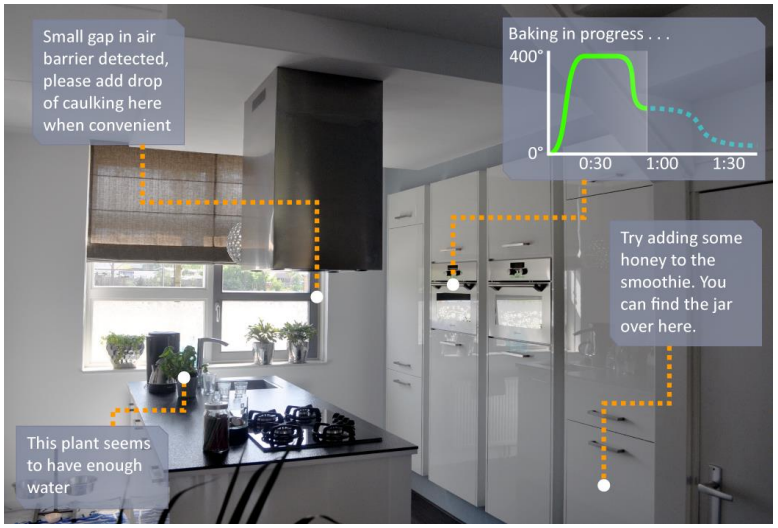
[62] (Thompson, Smarter than you Think, 2013)

63 (Gladwell, 2000) Unfortunately, like the loss of a joint memory when married couples are apart or our notes are lost, the loss of a mind-augmenting device like a cellphone can have a terrible impact on overly

also let one human accomplish the work of many. This gives the average person access to what was once only available to the rich and powerful: theatre performed each night within the comfort of the home, prepared meals, music lessons, and gossip about celebrities on the other side of the world.

1.4.3. Deskilling

However, this automation, while increasing access to information, does threaten to deskill people by building the skill of a task into an environment and its free-thinking machines. Automatic systems such as spaces that clean themselves or stock themselves with food eliminate the need for frequent practice and therefore erode a person's skills. Yet, this is not necessarily a consequence of the automation itself but the manner of it, as typically the temptation is to delegate as much work to a machine as possible and have humans support it in the areas it is incapable of performing. Loss of skill can be averted by redirecting automation to a role like the brain's subconscious. This form of automation would take over routine tasks the human has already mastered. In a support role it would provide new information that expands perspective, counter biases inherent in the task's process, and work to maintain the human in an active decision-making role[64].



1-2 Automation that supports but does not replace

Unfortunately, at the moment the ability to make such decisions is lacking in our machines. Unlike a cook, a mass-produced prepackaged meal cannot respond to suggestions. It is also not simple to ask a proximity-activated light or faucet to grow dimmer or increase pressure without rebuilding those devices. Such products and tools function only according to their own rules and only accept new suggestions with great difficulty. As active elements within architecture offer more suggestions or make more decisions on the occupants' behalf, the ability to coordinate and negotiate becomes more important as it could leave inhabitants with no idea how a decision was made or how to change it.

attached people. Whenever without the device, they feel as if they have lost their mind: disconnected and adrift. To some when it is missing it even becomes a phantom limb. They are so attached to their device that they use it even when it is dangerous: impairing their perception while walking or driving (Turkle, 2011).

[64] (Carr, 2014) and (Norman, 2007)

1.5. Preserving Agency

1.5.1. The Problem of Prefabricated Choices

This impact of free-thinking machines on agency is a modern variation of another aspect of Socrates's objection to the technology of the written word: one cannot argue against a prefabricated dialogue which sets its statements in stone. Furthermore, since written literature is standardized for mass consumption, as opposed to a teacher who alters a lesson to fit the student, the content it provides may not be in the form the reader understands[65]. In architecture, this problem can currently be seen in building control systems only accessible to remotely located building managers, or if accessible, presented with an interface only understandable to a select few experts.

This inability to argue with a machine's prefabricated decision is similar to a dirty trick in negotiation: sending a subordinate in a superior's place. Having no authority to change his or her position, a subordinate can only answer that he or she must get permission to change the deal, only to later return stating that he or she could not convince his or her superior to agree to the changed deal. That superior, had he or she negotiated directly, would have had the power and authority to be flexible and reach a compromise which, in the end, would have achieved an inferior deal for the superior[66].

Likewise, a thermostat, light switch, and security system only have the authority, or rather the ability, to act in the manner they have been designed. The variability of that ability only goes as far as the designers of the devices' cleverness and how thoroughly they imagined every major contingency which a device will encounter and laid out the most appropriate response. This inflexibility demands that machine's environment and the people interacting with it become more predictable to suit the machine.

Additionally, since these devices are built for consistency and accuracy, we trust our autonomous machines to perform without fail. Unless their instructions are obviously wrong, we will often lazily accept their decisions without critically thinking whether we should agree with them. This is a mental laziness that has led drivers into following their car's GPS into a river, one unthinking object following another. Of course, as far as the GPS was aware within its own limited model of the world, it was certainly correct to guide the driver in that direction. However, we often forget that the true accuracy of an autonomous machine's decision is only as good as the quality of the limited inputs it is given and the suitability of the procedure it is programmed to follow for the action it is asked to complete[67].

Unfortunately, these confident machines are quite terrible in evaluating the quality of the information they are given (such as relying on a single thermometer on the sunny side of the room) and typically fail to check if they can indeed perform the task as asked (such as vacuuming when the floor is covered in marbles). Admittedly, we humans are not much better at this self-evaluation, but our diverse opinions and perspectives make us quite good at evaluating each other's assumptions[68], critiques we can and are often happy to communicate. Our rigid and antisocial autonomous machines lack this sort of quality control and instead rely on our behaviour and their environment to be unfailingly consistent so that variables not accounted for remain unchanged.

These machines prefer to act unilaterally as if their purpose were to take over a task completely and seek no input from a person unless

[65] (Innis, 1951)
[66] (Norman, The Design of Future Things, 2007)
[67] (Norman, The Design of Future Things, 2007)
[68] (Kahnman, 2011)

they run into trouble when those variables change. Yet, they often fail to explain the nature of their trouble, limiting their complaint to just that they are experiencing trouble and it is the person's role to solve it[69]. This sort of antisocial behaviour has been less a problem in the past as simple mechanisms simply gave straightforward and predictable feedback through the repetition of their unchanging actions. This made it quick to learn their half-dozen word language.

This predictability is lost when machines are made more responsive to accommodate the inherent uncertainty that exists in the environment and fluidity of people's behaviour. As autonomous environments grow the ability to change their actions they will harness present and past observations when confronted with the new and experiment to discover new solutions when confronted with the familiar. This flexibility will lead to situations in which we are unsure what an environment's devices will do or are unable to alter its course once it starts to act. We will also be left in a situation that demands our attention when an intelligent environment fails and turns to us for help even though we have no idea what it did, how it did it, and what we are supposed to do to help it[70].

1.5.2. Maintaining Familiarity with Ever-Changing Rules

Through the use of cultural memory, humans have become experts at developing an understanding of the world and its rules. That cultural memory has become so vast however that it is next to impossible for a single individual to master all of it[71]. Complicating matters, the rate of cultural adaption and the accompanying technological change requires that topics and devices mastered be re-mastered as old ideas and technologies become obsolete. As machines become more diverse in behaviour and also add a flood of their own observations to the store of human cultural knowledge, it becomes important for that knowledge to be in a form that humans too can interpret.

As explained earlier, humans supplement learning about the world through personal experience with cultural knowledge. Much of this learning is in childhood as we absorb this information and make connections between what human society tells us and what we experience. Jeff Hawkins argues that we do this by finding patterns. He explains that the human brain is a pattern recognition machine where the source, format, and media do not matter, only the pattern. Give the tongue a stream of visual inputs and the brain recognizes the pattern as vision despite the tongue having developed for a different sensory function than eyes[72]. Humans use these patterns to make predictions and inform our actions in familiar and unfamiliar situations. When that situation is unfamiliar, the best a person can do is draw on the patterns of rules and languages of interactions of similar experiences. This can mislead an individual with no experience with horseback riding into talking with a horse as if it were a dog or person, or direct it around as if the horse were an obedient car by assuming that horses work the same way. Yet, that person would likely also treat the horse with care as he or she would another living person or animal and also keep the horse on a clear path like he or she would with a car.

As we grow older, new patterns emerge that have little connection to the patterns we were familiar with in the past. New forms of slang continuously appear, as do new music, tastes, and technologies. The elderly can for the most part ignore slowly changing languages and culture. These are slow changes that occur as each new generation of children explores the world in its own way and learns what works slightly different than the last. Unfortunately, the current rate of technological change makes

[69] (Norman, The Design of Future Things, 2007)
[70] (Norman, The Design of Future Things, 2007)
[71] (Kelly, What Technology Wants, 2010)
[72] (Hawkins & Blakeslee, 2004)

change a bit more difficult to ignore as the familiar is overridden by the unfamiliar at a much more rapid pace.

Technologies are currently rapidly evolving; balancing convenience with new abilities, often altering to fit a need, other times altering to drive obsolescence for profit[73]. Phones became wireless but relied on ever-draining batteries, then they became miniature touchscreen computers with the phone buried within. They became powerful and convenient for the generation familiar with the language of the computers, but initially bewildering for the earlier generation raised on buttons and dials.

How each tool can be used constrains the behaviour of its user but it also lacks the constraints of other tools. As each new technology is introduced, it opens up new opportunities but also closes others. It frees people from the constraints of old tools but also pushes people into new and unknown methods, away from the familiar[74]. Typically, as with slang and music, people can take their time to learn, test, and choose from a diverse pool which new technologies they adopt into their lives. They can decide which combination of beneficial opportunities and negative constraints fit their lifestyle. This creates choices, sometimes inconsistent, that largely fall to personal preference as people decide whether they can manage without conveniences like a microwave, air conditioner, or ballpoint pen. Some of this inconsistency in choice rises from the overwhelming amount of technological choices and tools people have at their disposal. It is a considerable amount that only a small portion of the population, at the expense of other hobbies and interests, are willing to take the time and sift through and learn the quirks of each.

[73] (The Economist, 2003)
[74] (Thompson, 2013)
[75] (Kelly, What Technology Wants, 2010)

Furthermore, other technologies are more difficult to reject. Rejecting television, cell phones, and the Internet can isolate a person socially, while avoiding CCTV cameras, advertising, or the automobile in suburbia can be impractical without abandoning modern society altogether[75]. When technological change alters the rules for everyone, it disrupts the social fabric, sometimes to the degree that it sparks rebellion against it. The term Luddite comes from such a rebellion when English craftsmen fought against industrial automation[76].

A person can choose to live in a perceptive and intelligent home, but he or she has less choice about the intelligence and pervasive awareness of the public realm beyond that home. With this absence of choice it becomes ever more important for a person's agency to understand how that intelligent space limits his or her own opportunities and how he or she can work within it or minimize the space's impact on agency to a comfortable level.

1.5.3. Adopting Existing and Natural Rules of Interaction

Much of the motivation for creating humanoid robots is due to human artifacts and the built environment having already been optimized to fit human morphology. Like the prevalence of slower wireless networks in existing buildings as opposed to the much faster wired networks, humanoid robots do not require architectural space to be re-engineered[77]. As Stewart Brand notes, it is easier to adapt a new technology to the conditions of existing buildings than it is to revise old buildings to fit new and, relative to the building, shorter-lived technologies[78].

Similarly, while it is currently easier for an adaptable human to adjust to the quirks of a particular machine, it is easier to alter the design of a machine to fit a human than it is to alter a human to fit a

[76] (Brynjolfsson & McAfee, 2011)
[77] (Breazeal, et al., 2003)
[78] (Brand, 1994)

machine. This results in prototype technologies asking early adopters to often awkwardly alter their behaviour to fit the needs of the machine while more mature technologies have worked out many of its early quirks and ask for less familiarity and altered behaviour on the part of its users. Through interfaces and mechanisms these mature technologies translate impenetrable machine languages into something more readable to human eyes. While information and flexibility are lost in translation, curiosity and diverse interests mean that there will always be people who will take that extra effort to learn a machine's own language[79]. Yet, for the non-experts conversing with other humans is how we most naturally interact, so for the general public, as autonomous environments become more social to offset their growing unpredictability, they will become more accessible when they share those same social rules.

In adopting these human cues, machines take advantage of human empathy, where humans predict the actions of others based on how they would act themselves. A humanlike personality allows people to easily build a mental model of the machine's method of thinking, allowing for an intuitive prediction of what the machine will do, why it is doing it, and what information it needs to do things differently[80].

Care must be used in applying these existing and familiar rules, as to our pattern-seeking brains behaviour, far more strongly than appearance, gives animate objects the illusion of life and intelligence. An animated lamp may look nothing like a living creature; yet, through its behaviour which mimics a living animal, its lifelike actions are believable. Similarly, if an object has 'eyes', it is the ability of those eyes to rotate to indicate gaze and make multiple expressions that gives the appearance of life as opposed to more photorealistic but immobile eyes[81]. This is where anthropomorphism comes from; it is a consequence of human mind applying empathy to more than just humans. The mind cannot help but fill in gaps to understand the motivations of everything that moves, be it a dog, puppet, car, or the wind. This is why we perceive a dog who destroyed a rug as vengeful, a puppet as a cheerful comedian, a worn-out car as cranky, and a strong wind as angry. Misinterpreting the actions of a bored dog and misplacing motivation where none exists, the human mind cannot help but view the animate world through the lens of its own motivations[82].

The natural preference for humanlike behaviour does not mean that turning a doorknob is less effective or natural for humans than verbally asking a door to open. In fact, until recently for many adults talking to a door seems awkward and embarrassing as it is currently considered unnatural and delusional for inanimate objects to understand human speech[83]. Similarly, this does not mean devices should observe the world through cartoonish faces, as that might suggest an intelligence or awareness that is more than such a device needs. Instead, what the human brain considers to be natural forms of interaction are the interactions that follow the patterns it is familiar with. A flashing light seems a natural action for speaking machines because movies, television, and stereo volume bars have made it normal and familiar. Therefore, an autonomous environment that allows for natural interaction is one that adopts a human-centric language that we humans already commonly use when interacting with each other and the world. It is an environment that does not carelessly invent new dialect or choose a vocabulary that contradicts other familiar languages of other human and natural artifacts.

[79] Kevin Kelly observes that few technologies ever die; undeveloped societies still use oxcarts while steam-powered cars are still produced by hobbyists (Kelly, What Technology Wants, 2010). Languages may die out, but the technology of language itself remains strong.

[80] (Breazeal, et al., 2003)
[81] (Dautenhahn, 1998)
[82] (Dautenhahn, 1998) and (Graziano, 2013)
[83] (Turkle, 2011)

This is more than a matter of environments understanding human speech and answering in it. While speech recognition appears to offer a familiar method of interaction, it is currently a struggle for machines to match the richness of human speech. The expectation of clarity of a machine's beep is as much as we expect of a dog's bark. A recorded announcement expects as much a response as a speaker expects of a crowd. A machine that talks back however is a machine that is expected to converse. Speech is closely associated with humanity, so a machine with a voice that does not respond as fluently and intelligently as even a human child will be quick to break the spell. It would create disappointment by not responding in a dialogue the way an observer expects of all speakers of human language would respond[84].

Therefore, when adopting a flexible method like speech, what is more important than sounding like a human is the adoption of the human behaviour of conversation. Whether an interface is speech or pictogram based, conversational dialogue provides redundancy by offering adaptable responses. A speaker talking to another less familiar with a topic can step down to a level where the other has an understanding and work his or her way back up. This is a multilingual device, i.e., a device with multiple interfaces in itself or in other devices such as how smartphones today are used to interface with less fluent lights and thermostats; one that can adjust to the skill of a user; one that remembers which user did not want help writing a letter; and one that would actually provide help according to the skill of the user.

1.5.4. Explaining Itself

While taking advantage of the tendency of the human mind to understand the world through the lens of itself gives a starting point for a shared language of social interaction, there remains the problem in bridging the differences between man and machine when interacting.

The relationship between a horse and a skilled rider provides a precedent for a smooth interaction between two dissimilar autonomous decision-making systems; while both interpret the world differently, both are still able to understand the intentions of the other. Through a rich variation of actions, such as body language, posture, relaxation, tenseness, and behaviour, the two give feedback on their mutual status, typically providing a continuous stream of subtle status updates and communicating less subtly when they believe that their observation or decision is worth the other's attention. Through this varied conversation, the interaction expands from one side choosing from a rigid set of pre-existing options and responses to an informational feedback loop that with each pass further aligns both parties' understanding of what is wanted with what can be achieved and negotiate how it is accomplished. As the two converse, rarely is one party fully in control. Instead, as a rider chooses to trust in a horse's judgment and delegate more control and authority to the horse, they loosen the reins, and when re-exerting control, they tighten the reins while the horse indicates how much it agrees by cooperating or resisting. This negotiation for control is continuous between horse and rider, constantly adjusting to the current circumstances through mutual conversation and reaction to the immediate environment[85].

A conversation is a process of growing a larger share of common ground between two parties, where common ground is a mutually shared mass of knowledge, beliefs, and conjecture[86]. By entering into a dialogue, two parties will mutually alter their responses in reaction to each other, adjusting the content of their words into a form they believe the other party can understand based on the

[84] (Dautenhahn, 1998)
[85] (Norman, The Design of Future Things, 2007)
[86] (Norman, 2007)

other's last response. Gordon Pask's conversation theory proposes that through this feedback loop of conversation, the discussed information is actually not passed between each party. Instead, it is mutually constructed within each party's mind as each side adjusts the content of their response into a form they believe will nudge the other's understanding closer in alignment with their own[87]. The longer both parties spend time together, the larger that common ground shared between them becomes and the more accurate their assumptions of what is common ground between them becomes[88].

Once common ground is established it acts as a form of data compression. It allows an elaborate meaning to be quickly and easily communicated in a single word, phrase, or gesture and allows that same word or gesture to have a completely different meaning when it is used again in a different context[89]. It acts as the common denominator**,** the starting point that a conversation grows outwards from the initial fragment of shared language. Among friends and family familiar with each other's life stories, that starting point is where the story last left off ("How have you been since we last saw you?"); with an acquaintance or stranger on a hometown street aware of only fragments of the other's story, what is common ground shrinks to the local language and shared environment ("Beautiful weather today."); and in a foreign country even the common ground of verbal language is lost and a visitor must resort to the most basic forms of human language to communicate, pointing with fingers, miming actions, and expressing concepts in single words ("English?"). However, between humans and machines the usual logic falls apart; as noted in the earlier discussion regarding intelligence, machines neither share the same history, the same sort of upbringing, nor interact in the same manner as people[90]. To understand each other, a human and machine need to work much harder than a person would need to work to understand another person or a machine would need to work to understand another machine[91].

1.5.5. Explaining through Narrative

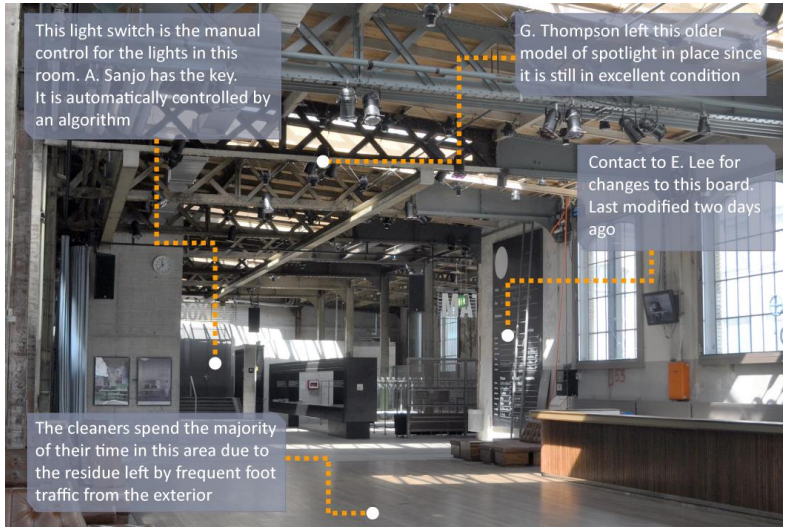
Humans typically describe events to each other through story. We perceive reality as a narrative, a sequence of events with rhyme and reason, cause and effect. As autobiographical agents, our minds weave the patterns our brains observe into a story that explains the world. If a new pattern does not fit within our internal narrative, it is either thrown out or results in older parts of the narrative being tweaked so that the story still plausibly explains our current motivations and behaviour. To understand other people and entities around us, we likewise build them stories that allow us to empathize with or understand their own motivations and behaviour. Where there are holes in their stories we put ourselves in their shoes and ask what we would do in their place; when that fails to explain the other person's behaviour, we criticize them for having terrible judgment[92].

A conversational machine that also functions as an autobiographical agent is an entity with its own story to tell. Its story of past experiences explains its current actions and present state of health. This is a machine that can describe to a repairman the events that led up to its broken wheel. Similarly, while our machines may not share the same genetic traits we share with other animals, they do share, as noted earlier, cultural traits. So, while their behaviour is quite unlike human behaviour, and perceiving their motivations through the anthropomorphic lens of humanity can lead to worse

[87] (Beesley & Khan, 2009)
[88] (Norman, 2007)
[89] (Norman, 2007)
[90] (Norman, 2007)
[91] (Norman, 2007)

[92] (Dautenhahn, 1998), (Gottschall, 2012), and (Hawkins & Blakeslee, 2004)

misunderstandings of motivations and their cognitive abilities than between humans and animals[93], their motivations will still follow a logical path that can be explained through cultural memory.



1-3 Adjusting information to the experience of the visitor and explaining through narrative.

Even a machine that relies on randomness does so for a logical reason, a reason that can be explained through story, whether it is summarizing the multitude of sources it pulled on to generate a split-second decision or explaining what happened while the human was out of the room. An example of such a narrative could be "the thermostat turned on the heating earlier today because statistically the youngest family member arrives home at about 3 pm on Tuesdays; it meant to add a bit less heat today because the weather is nicer, however it found it had to add more than usual for some reason, particularly in the northwest bedroom. The windows say they are closed, but perhaps they should be checked in person in case something is wrong". It may appear at first glance that some machines' stories might be too convoluted to figure out. In many cases however, what appears confusingly complex, such as an airplane's cockpit or human brain, is a matter of being shown the underlying logic[94]. When examining the rejected second and third choices to its answer for a Jeopardy question by IBM's question-answering engine known as Watson in its televised game versus human champions, commentators noted the bizarre nature of the rejected choices, where some were completely unrelated to the question. It appeared that Watson could still misunderstand questions; yet when human eyes collectively across the Internet looked deeper, a few discovered that through an obscure movie quote, the rejected choice and the question were indeed related and was actually a logical choice. Humans may not be as fast as a machine, but working together we can still work out the underlying logic to a machine's narrative[95].

A more significant problem is that an environment that can form its own internal mental narrative is also an environment that watches and remembers.

1.6. Ubiquitous Perception

1.6.1. Machine Learning

Given intelligence's reliance on awareness of the present and past to determine the appropriate course of action, an environment saturated with free-thinking machines will likely also be one equally saturated with perception and memory. In this sort of environment, a door would be able to describe its history in far greater detail than a product identification and certification label stamped on its side. Fitted with electronics and sensors, a door would have the potential

[93] (Bostrom, The Superintelligent Will: Motivation and Instrumental Rationality in Advanced Artificial Agents, 2012)

[94] (Hawkins & Blakeslee, 2004) and (Norman, The Design of Future Things, 2007)

[95] (Thompson, 2013)

to remember its autobiographical history. It could observe and tally how often it is opened. Compare the roughness and frequency of every time it is closed gently with care or slammed shut in anger. And like an answering machine, it could recall each instance a visitor came knocking but was unanswered. As each event in an object's life becomes time-stamped, the scratches, chips, and dents that tell of a storied past can instead be described to humans and fellow constructs as an actual story that tells of the actual circumstances of the events that created its scars.

As autonomous machines construct these stories, they will naturally grow a familiarity with their environment and become active learners instead of thoughtless constructs. Through learning, they will develop a better understanding of how best to interact with their environment on our behalf in a way that avoids those scratches and chips. That people form a major part of the environments these machines act within complicates matters however. To act intelligently interact with their environment and assist us properly on our behalf, each machine's familiarity with its environment will therefore have to extend to an understanding of the humans within that space and those human's intentions. Luckily this familiarity only needs to be good enough to accomplish that machine's job without causing annoyance or complications. So the degree each device needs to understand people varies. Scheduling systems and temperature control systems benefit more from understanding people than an intelligent vacuum.

As noted earlier, a clear understanding of human nature will prove difficult for machines. Machines currently manage this problem through two methods. The first involves feeding a machine a selection of predefined knowledge: one plus one equals two, blue is the RGB value of [0,0,255], and when this button is pressed, do that. In the early days of artificial intelligence, this is how we taught machines. To answer the question of what cats would like to eat, a rules-based approach relying on logical axioms would involve the construction of a database about cats, their diet, and logical rules connecting the two. This works well for predictable tasks, not so well when the variables change. The second method relies on brute force statistics, feeding a machine indiscriminate data from web pages, social networks, digitized books, and other sources to find patterns that would inform a statistical model on the probability of cats liking certain foods. Some versions of this sort of machine learning rely on artificial neural networks which process raw data into patterns in a similar manner as the neurons of the human brain, such as Gordon Pask's electrochemical-based neural networks in the 1950s. Pask's neural networks were taught through the reward of free building materials when certain criteria were met. This allowed it to be taught to grow its own sound sensor without being explicitly told what a sound sensor is. Similarly, modern artificial intelligence companies like Google have harnessed modern hardware and techniques to put modern electronic neural networks widely into use, vastly improving voice and image recognition and other areas of machine learning[96]. This more flexible sort of machine learning has been around for years, but we finally have enough data and processing power for the techniques to work[97].

1.6.2. Data Collection

This data comes from machines observing an ever-growing proportion of the world. To anticipate, understand and fulfill its inhabitants' desire, an intelligent environment needs to grow its familiarity by observing every action of each user and every change in his or her environment. It must also store that data so that it can use its memory of past actions to predict future actions. This results in the system forming an omnipresent representation of each inhabitant's desires, forming an intimate picture of his or her habits.

[96] (McMillan, 2013)
[97] (Heaven, 2013)

On one side it gives those inhabitants a better understanding of themselves and can serve as a powerful augmentation to their memory, on the other it is a record of our irrational acts, eccentricities, falsehoods and embarrassing mistakes[98].

Unfortunately our relationship with the contemporary information-gathering machines that would make up an intelligent environment is neither one between friends or coworkers, nor one between client and professional, but rather it is one between customer and merchant or citizen and government. This could be problematic as observational systems begin to populate the home. These sensory systems may use the information they gather to provide better service and suggestions, but these systems, many having grown out of a realm of advertising and retail, understandably serve their commercial creators first. This is not a conscious choice on behalf of a machine, but one ingrained into its code, as it vacuums up financially valuable data on behalf of corporations and authorities who are confronted with the temptation to cheaply gather it all now and find a use for that plentiful data later[99].

As an aside, there can be unforeseen variables in this collection of data, such as unintentional discrimination. Harnessing the power of mobile smartphones, cities like Boston have begun to offer data collection applications that allow citizens to voluntarily submit sensory data to the city. This includes an app that uses a smartphone's vibration sensors and GPS to automatically detect and submit pothole locations. Boston soon noticed that the data was biased towards wealthier neighbourhoods as the poor and elderly are less likely to carry smartphones. A tweak to their app that accounted for this underreporting fixed the problem, but it is an example of how attention needs to be given to a pool of data's inherent biases[100]."

From testing the most clicked shade of blue on advertisement hyperlinks to adjusting the average ratio of positive or negative updates that appear in a social news feed[101], services that live on the personal data of their users or customers walk a fine line between applying that data to provide a better service than their competitors and overusing their users' private information in an invasive, unsafe, and unwelcome manner. Additionally, with inaccessible terms of use agreements online which are rarely read or fully understood[102] and with physical sensors placed unobtrusively in urban spaces, people are generally unaware of the degree that they are being observed online and offline by Internet cookies, surveillance cameras, and other forms of data collection. Worse, they are rarely clearly informed how and when that information will be used and shared[103]. We have become so accustomed to living with an electronic shadow of potential surveillance that it fades from our attention. Ignoring it allows us to behave as if it doesn't exist, until it returns to bit us in a lawsuit, scandal or investigation[104].

Even if a physical store or online service treats its customers or users well by clearly marking its sensors, presenting fair and simple terms of use; it also needs to protect itself from malicious outsiders. It must defend itself against hackers and thieves who steal credit card information, photos, and identities; and from domestic and foreign governmental agencies that operate on the edges of the law in the name of security.

The present insecurity and misuse of personal data has parallels to the early days of the internet where purchasing products and

[98] (Cetkovic, 2011)
[99] (Nourbakhsh, 2013) and (Andrejevic, 2012)
[100] (Podesta, 2014)
[101] (Hern, 2014) and (Hill, 2014)
[102] (Podesta, 2014)
[103] (Podesta, 2014)
[104] (Turkle, 2011)

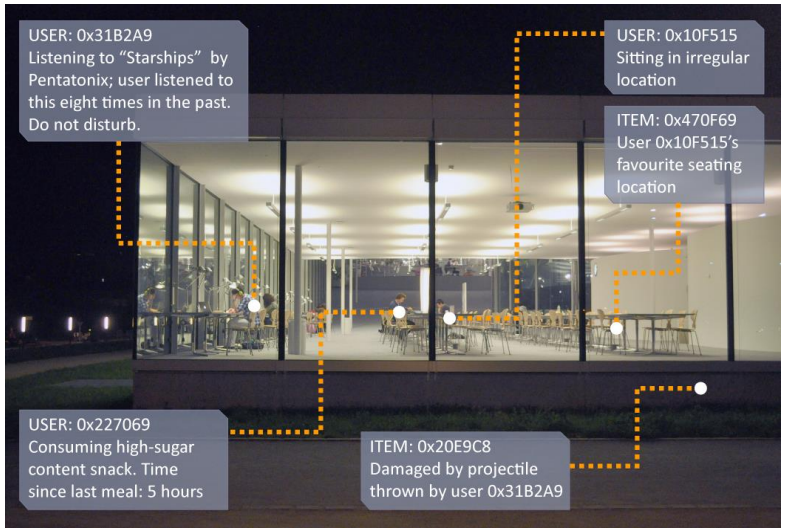
services online came with a risk of credit card or banking information being stolen or misused. It may only be until the sharing and use of personal information collected offline and analyzed online can be secured to a degree that provides the same level of trust as present day online commerce that the systems applying personal data can reach their true potential[105].

1.6.3. Anonymity and Transparency

Currently anything said or any mistake made on the Internet is rarely forgotten and control over its distribution is lost[106]. As sensory systems in the physical world flood the digital world with real world data, this will begin to apply to the normally forgetful physical world as well. On the Internet, people deal with this persistent memory through different methods. The first option is to disconnect and provide no data to form an online identity or at least separate personal statements and mistakes from your true identity by taking on an anonymous identity mixed with the other anonymous. This form of anonymity, the privacy of being unknown, is the anonymity of a mask. The second, as noted above, is to ignore the online world's persistent memory and hope that what personal information you inevitably share is so ordinary, average, and inconsequential that no one but your friends will care about it. This is a form of anonymity that comes from being unnoticed in a crowd[107], but only partially works for ordinary people. For a celebrity this means living in a glass house. The third option is like the privacy of the home, a closed off community that offers intimate privacy and relies on friends to limit their gossip. Perceptive intelligent environments will likely evolve to offer this third type of privacy as they upload personal data into the digital realm.

An individual who does not want the system to remember him or her is an individual who cannot be identified[108] and authorities find non-identifiable individuals untrustworthy[109]. Under a mask, criminals become anonymous and can act against each other and society free from social sanctions[110]. Due to the difficulty in distinguishing the anonymous into different identities, the harmless anonymous unavoidably share the same identity with these malicious individuals and all are mistrusted equally. Additionally, security tends to protect more against the different than the wrong. A stranger walking down an open street is a normal event, but a stranger walking down a street in a gated community creates unnecessary anxiety by forcing residents to question if that person belongs and to consider if the stranger has breached the security of the community[111]. This criminal association with anonymity can raise suspicion whenever a person opts out of a service that asks for a "mere harmless" breach of privacy. Some perceive that the only reasons a person would avoid a "free" benefit is that he or she is either a socially backward Luddite or has something illegal or immoral to hide[112].

[105] (Schneier, Our Security Models Will Never Work — No Matter What We Do, 2013)
[106] (Turkle, 2011)
[107] (Turkle, 2011)
[108] (Wright, Gutwirth, Friedewald, Vildjiounaite, & Punie, 2008)
[109] And corporations cannot market efficiently to the anonymous.
[110] (Smith, 1997)
[111] People often feel less secure in places with greater security (Minton, 2009)
[112] (Wright, Gutwirth, Friedewald, Vildjiounaite, & Punie, 2008)



1-4 A building that watches to learn is also a building that remembers

On the other hand, to accept this breach of privacy is to subject yourself to self-censorship. Even when hiding in a crowd, the knowledge that you are being watched changes your behaviour. It becomes a one-sided conversation as you consider which parts of yourself you are willing to show a stranger. Of course, with no feedback as to what sort of person the watcher is, we can only assume that he or she is an average person, so when under surveillance, people are inclined to act in a manner that conforms to mainstream expectations[113]. This prevents people from experimenting with new and different ideas from fear that they could seem incompetent, stupid, or at worst deviant and make them vulnerable to criticism, discrimination, coercion, or punishment for doing something weird or badly[114]. To act different is to risk revealing stigmatized political views by revealing one's attendance at rallies or protests, or visits to the plastic surgeon or abortion clinic[115], actions acceptable to some portions of the population, unacceptable to others. In a diverse society of varying and often conflicting opinions, what is socially acceptable varies, influencing people to show different sides of themselves depending on the people they are currently with. To maintain and strengthen social cohesion, people create white lies[116], complimenting something they do not care about but doing so because a friend put significant effort into making it.

While a society of selfish liars trends toward a fragmented society full of mistrust, an entirely honest society trends toward a uniform society with no significant difference in opinion. In between the two, a society of selfless liars will trend toward a society of diverse ideas and opinions[117]. This diversity occurs as individuals pretend that they agree with each other but are free to do things their own way and change at their own pace. An urban environment where every person's action is transparent to friends and society makes it difficult to maintain those white lies and would weaken the fabric of society.

Secrets good and bad are also seen as an incubator for resistance, leading to control-oriented states to spy on the masses. When the state catches an individual's often simply petty transgressions, it confronts him or her and uses its power to maintain social control by unnerving and disempowering the opposition[118]. George Orwell believed that such pervasive surveillance alone is enough to crush dissent and permanently subjugate the watched under the power of the watcher as they watch to determine that society does things their way[119].

[113] (Richards, 2012)
[114] (Richards, 2012)
[115] (Shilton, 2009)
[116] (Iñiguez, Govezensky, Dunbar, Kaski, & Barrio, 2014)
[117] (Iñiguez, Govezensky, Dunbar, Kaski, & Barrio, 2014)
[118] (Smith, 1997)
[119] (Yoquinto, 2014)

It is this influence that the watcher has over the watched that makes privacy not about hiding a wrong, but about maintaining the human condition. It is about protecting innocent people from the misuse of information, whether that abuse is selling it to marketers who can use it to closely analyze people and manipulate them into buying products they might never need, or prying on political enemies to gain an advantage by dredging up mistakes irrelevant to present issues or their ability to lead[120]. This is why merely hiding in the crowd and acting as if the cloud of surveillance does not exist is not enough, particularly as the perceptive environment enters the home.

1.6.4. The Home as Shelter for Ideas

Since the villas and palaces of antiquity, the home has provided a protected place for contemplation and generating new ideas. For the majority of the population however, the availability of the private home, like many inventions, was not very evenly distributed[121]. In Western Europe, the medieval townhouse, despite functioning as the family home, was a single open hall shared with relatives, servants, apprentices, employees, and frequent guests with no privacy at all. As the bourgeoisie grew more prosperous, they developed sufficient wealth to separate the family business from the living quarters; free from noise, dirt, employees, customers, and suppliers, the home became quiet, clean, and peaceful. The public presence continued to decline and by the early nineteenth century the typical home had reversed from a place for interacting with the public to a sanctuary from society[122].

In the modern home, certainty of privacy is as simple as closing the curtains and declining visitors. Its physical enclosure acts as a secure and opaque barrier that protects the home from the public realm and provides valuable privacy in the form of solitude, intimacy, and reserve[123].

Solitude is absolute privacy, a complete separation from all other people. It provides an individual with the greatest freedom from intrusion and interruption, offering a relaxed and secure moment for private activities, which vary by culture, such as sleep, undressing, or nose cleaning. The human psyche is a permeable membrane, it absorbs the feelings, moods, and opinions of the people who come into contact with us; their questions make us reflect on our actions, influence our future plans, and question the accuracy of our memory. The solitude the home provides allows us to take a break from being what others desire and sort through and reflect on what we have absorbed; it also allows us to fail in private, to experiment with incomplete and ridiculous ideas[124].

Intimacy is a social form of solitude, allowing friends and family gathering together at home to privately relax their social facades and open up to each other. It lets people speak and act boisterously, romantically, or rebelliously in a manner that would be unacceptable in public but not behind closed doors between friends. It gives people the opportunity to share and seek input on the half-formed or controversial ideas and opinions that they formed while in solitude which would ordinarily make a person vulnerable to criticism or abuse if they were discussed publicly[125].

Reserve is the form a privacy people turn to when living in less private homes or when away from home. It is the privacy of hiding behind a social façade, but it is also the self-censorship that occurs when an individual believes he or she is being watched. It includes avoiding social disruption on behalf of yourself or another by

[120] (Schneier, 2006)
[121] (Riley, 1999); as in "*The future is already here — it's just not very evenly distributed*" as observed by (Gibson, 1999)
[122] (Riley, 1999)
[123] (Smith, 1997)
[124] (Smith, 1997)
[125] (Smith, 1997)

keeping your personality or nature private while in a social setting. It involves avoiding indelicate and awkward words, avoiding open acknowledgement of small irritations, choosing to hold back what you are truly feeling and silencing frank opinions to smooth social interaction. Furthermore, reserve provides privacy through imaginary barriers by people politely avoiding intruding despite sharing the same room, by turning around when another needs privacy, or improvising an unspoken private space of uninterruption to family members watching television[126].

The home as a place for solitude and intimacy began to disappear as communication technologies such as radio, telephone, television, and the Internet created a more permeable sort of privacy[127]. So far, this intrusion is a voluntary one with information coming in but little of it seeping out. However, as ubiquitous perception in the public realm threatens to annihilate anonymity in the urban environment[128] and intelligent autonomous systems reduce solitude of unobserved spaces in the home, people will grow a greater reliance on reserve and intimacy to maintain privacy in a world of ubiquitous perception. Although solitude guarantees privacy through absolute separation and anonymity provides privacy through the energy expenditure required to identify a stranger, both intimacy and reserve require trust to maintain privacy.

1.7. Shelter from the Virtual Environment

1.7.1. A Problem of Trust

A future building saturated with free-thinking machines will not likely consist of products manufactured and controlled by a single monolithic corporation or governmental entity. Instead like present buildings built from a mixture of various products and materials from an equally various mixture of manufacturers, the perceptive and analytic components of an intelligent building will also likely come from various different manufacturers[129]. Such a diverse collection of machines may struggle in the early years of adoption to agree on a shared language to fully communicate and share information with each other, but it will create an environment of choice where some choices are limited while others are greater. In negotiating how to work together these products and systems will all will share in the struggle of balancing the privacy of the user with providing improved service by getting to know the user better.

It is here that some worry the reserve we expect of our servant machines may fail us. Many people perceive programmed machines to be easier to trust than a person[130]; to build trust with a person you must spend an extended period of time to get to know them and their quirks; however, a machine programmed to follow a set of rules will not break them. While it is possible to create machines that are unpredictable and full of delightful surprises, when asked

[126] (Smith, 1997)
[127] (Riley, 1999)
[128] (Andrejevic, 2012)
[129] (Cetkovic, 2011)
[130] (Turkle, 2011). If there is one thing people like less than their actions being recorded by corporations and authorities, it is being recorded by strangers. In 2012 Steve Mann, a life logger who in the 1970s created the EyeTap, the first head-mounted camera, found himself part of the first documented cybernetic hate crime when he was attacked for wearing a modern version of the camera in a French McDonald's. Yet few commentators recognized that the event was also recorded by the McDonald's own surveillance cameras (Dvorsky, Google Glass Ushers in the Next Wave of Cybernetic Hate Crimes, 2013), but while the corporation may be faceless, it is a known quantity and people know who to blame if its security footage is ever misused. It is due to this discomfort that many people have with being recorded that influenced many wearable technology pioneers to avoid adding cameras to their devices (Thompson, 2013).

for a preference, many people instead desire a safe and predictable assistant[131]. If a predictable machine does break the rules and acts in a manner it is not supposed to, it is seen as the fault of the humans who programmed it rather than the machine itself.

Yet, the rules a machine is programmed to follow may not be the rules an occupant thinks they have agreed on. Many perceptive and analytic devices may affordably find their way into the home and other buildings through the same model of providing a service in exchange for personal data that is currently used on the Internet. Grocery supply companies could supply free intelligent fridges in exchange for both a subscription for food delivery and access to a record of what the consumer consumes, giving the company better intelligence for its marketing and corporate strategies. This collection of private data from a sentient building may be more voluntary than illegal, but as noted, its occupants may not be entirely clear on what they have agreed to share[132].

While ordinarily a friend who surreptitiously recorded your private moments and stories would, when caught, no longer be your friend even if they presented you with a mutually signed friendship agreement that in fine print explicitly allowed for such behaviour, we tend to forgive equally surreptitious devices and services as their usefulness may have a perceived value which outweighs the perceived price that they both ask of and hide from us.

Yet such forgiveness may be hard to give when visiting environments and buildings which seem to know everything about us. Just as machines think differently than we do, so do their memories function differently as well. Unlike animals, machines can losslessly share their actual memories, expertise, and senses directly with each other without translation to words or scent and connect to data centres to boost their brain power, forming a mind that is both individual and monolithic. When one machine meets a person, it can seek out and copy memories from another machine intimate with that person. Through this "gossip," it can bypass the effort of the original machine and now know that person as well as the first, despite being a complete stranger to the human. This could create a world where every human is a celebrity, as when a celebrity encounters a fan, the fan knows an unnerving amount of information about the celebrity, but the celebrity knows nothing about the fan[133]. Even if the robot itself has little care of the human's actions, other entities may find much more value in the information that robots shares with them, willingly or unwillingly. A person may have some control over who observes them, but he or she has little control over the individuals or groups who gather that information from the observer.

1.7.2. Adapting to the Over-Sharing of Personal Data

Social pressure can have a positive effect on the trustworthiness of our perceptive machines and services as more people become more technically literate and find their voices amplified in the Internet age. While authorities and corporations may find themselves spoiled with easily collected personal data which can be used to give them an edge, they still need to balance that use with staying in the public's good graces. So, while like prisoners of war being sent to another nation where the initial nation's human rights laws do not apply, personal data given protection in one nation may not receive that same level of protection if sent to and stored within another nation[134]. Such tactics are increasingly being caught and protested.

This is because literacy now means more than just reading and writing; it now includes other forms of communication, including statistical data, photography, and video. For instance, the democratization of photography means that George Orwell's fear that authoritarian governments would adopt Stalin's photo-

[131] (Turkle, 2011)
[132] (Cetkovic, 2011)
[133] (Nourbakhsh, 2013)
[134] (Wright, Gutwirth, Friedewald, Vildjiounaite, & Punie, 2008)

manipulation technique of erasing enemies from history will not come to pass as the public's now widespread ability to photo-manipulate and communicate results in photo-manipulated propaganda and advertising being quickly recognized and mocked instead of ignorantly accepted. Despite the greater ease in manipulating images, truth is not dying in photography because greater literacy in the art means that falseness is quickly recognized[135]. Likewise, as impacts on privacy affect ever greater proportions of society, literacy and awareness surrounding personal information will increase and public pressure for change will grow. This pressure will influence features to improve privacy be added to software and features which invade to be scaled back. Similarly, secret initiatives to collect and analyze the private data of an entire nation will face the threats of leaks, regulation, and fortification against them by the victims, corporate and individual alike. This will not happen instantly; in fact, many invasive features and schemes may remain hidden for years, but slowly things will improve. Like all new technologies, it will take time and effort to sort out all the issues. One consolation to the rampant misuse of information on the Internet is that it is serving as a testing ground before similar technologies fully spread into the physical world.

Although there will inevitably be a balance between ease of sharing and privacy, where between the two that balance will reach equilibrium is still in flux. In an era when there remains uncertainty in what will be shared and what will not, as designers of how people interact with the built environment it becomes the architect's responsibility to provide the inhabitants of a space with a transparent choice of how that space collects their personal data and how it uses what it has collected. This could involve highlighting sensors and giving them a form that describes their abilities, dividing a space into clearly demarcated surveyed and non-surveyed areas, or providing physical barriers to surveillance much like the blinds and curtains that provide varying degrees of transparency between our existing home and the outside world. Like the life loggers mentioned earlier, MIT speech scientist Dey Roy experimented with continuous recording but in his case it involved wiring up his house with wide-angle cameras and sensitive microphones to create a memory machine that would capture every word spoken to and every interaction with their newly born child to understand how children learn to speak. Roy made this continuous recording that could be viewed by other researchers bearable by adding the ability to turn the surveillance on and off at will, such as when going to bed. Knowing that one often does not see embarrassing moments coming until they happen, Roy included an "oops" button that causes the system to forget minutes, hours, or days. This is useful when sleepy individuals forget about the system when exiting the washroom after a shower early in the morning[136]. However this management of privacy is accomplished, it is a matter of the inhabitant having agency over the level of his or her privacy such that it gives them a greater feeling of control and therefore a greater feeling of comfort[137].

1.7.3. Real World Spam Filters

The strongest protection of privacy may come from our tendency to solve the problem of overwhelming technologies by throwing even more technology at the problem[138]. As our buildings and their services observe the world to understand us and the context of the tasks we ask of them, that greater understanding can also be applied to recognize and separate relevant information from the flood of information that overwhelms us in the modern world. Acting as supercharged spam filters for everyday life, buildings could serve as

[135] (Thompson, 2013)
136
(Thompson, 2013)
[137] (Cetkovic, 2011)
[138] (Turkle, 2011)

an environmental separation from the virtual world. Like their regulation of physical matter through the building skin, they will regulate and organize the flow of information as it passes in and out of the ever more information-permeable building skin.

With this new infrastructure, Wiener's cybernetics will become a reality. Locally the personal assistant, filtering our calls, pointing out disagreeable sections of contracts, and recommending trustworthy smoke detectors for purchase, becomes part of the house; and like the human brain elevating elements that are important to conscious attention while automating the mundane and repetitive, the built environment becomes an augmented subconscious just as plumbing augments the digestive system and walls augment human skin. Dynamic urban and architectural spaces will incorporate their inhabitants as decision influences with transient preferences and needs. No longer generic occupants, they will become users with personal preferences, each memorized and then recognized by various systems[139].

Yet, smart traffic light systems that react to traffic conditions are useful, advertisements that appear as a response to one's proximity are annoying, and being denied access to public systems because the system does not like your habits is worrisome[140]. Such an augmented subconscious and cybernetic environment requires an immense level of trust because the responsibility of filtering the information that flows into our digital selves gives a machine power to manipulate how we perceive the greater world and can censor us to a far greater degree than surveillance could ever force us to self-censor.

This is not just a matter of trusting machines not to give our private information to malicious individuals who will misuse it, but a matter of being able to trust our artificial assistants to not apply their great power of intelligence to take the place of master and place us in the position of servant.

1.8. The Building as Person

1.8.1. The Fear of the Artificial Servant

Designed to perform dirty, dangerous or monotonous tasks and free humans to pursue more comfortable and interesting pursuits, it can be said that robots are created to fill the role of slaves. Robotized buildings are seen as a solution to providing mastery and control over a messy and unpredictable world, with the assumption that they will support humans at the top of the hierarchy with robot lackeys as their servants[141].

While it is useful to have a built environment that performs human-oriented tasks and recognizes human language, motivations, and hints; we are unlikely to desire intelligent toasters or light bulbs with actual humanlike minds. A car that wants to go on a road trip isn't desirable if you are trying to get to work, neither is a vacuum that watches television all day helpful, nor would a person want to be sued for damages by their home after failing to repair the hole in the roof[142].

Seeking the joy of self-enhancement brings the anxiety of loss and helplessness when delegating away agency, and a master considering his or her own unwillingness to be a servant fears their servants may similarly hold the same resentment. To ease their guilt many masters in the past would convince themselves that their slaves were not true humans and therefore would be incapable of holding that resentment. In the case of artificial slaves, the problem is not necessarily violent rebellion, but more our tendency to provide them

[139] (Hawkins & Blakeslee, 2004) and (Nabian & Ratti, 2011)
[140] (Shepard, 2009)
[141] (LaGrandeur, 2011)
[142] (Stross, 2011)

with too much power. This worry about intelligent artificial servants is an old one; not just a worry common to the industrial and modern periods, but also in Ancient Greek, Medieval, and Renaissance literature as well[143].

However, the lesson of Mary Shelley's *Frankenstein* is not that we should never attempt to create new technologies or life, but that we should never abandon our creations. That was Dr. Frankenstein's mistake: horrified that his creation did not turn out as he expected he cast it out and abandoned it to its own devices. That rejection turned his confused creation into a vengeful monster and led him on his path to his own downfall[144].

Observing this fear of artificial beings in *Frankenstein* and the many science fiction tales that followed, science fiction author and scientist Isaac Asimov coined the term "Frankenstein Complex." This term refers to the fear that scientific discoveries will eventually produce something beyond human control. The creation being outside human control, this fear assumes that it will inevitably harm people or humanity as a whole[145].

As a young author, Asimov had imagined a bright future in which humanity would be served by humanoid robots but saw fear as the greatest barrier to its fruition. Noting that dangerous tools generally have safeguards, Asimov concluded that the same would apply to robots; and to demonstrate what such safeguards could look like, Asimov created the Three Laws of Robotics[146]. Although the robots in his robot series are indeed selfless servants of human society and almost impossibly safe, they remain subject to human prejudice and constrained by paranoid rules that damage their minds should harm come to a human.

1.8.2. Letting Machines Grow Up

In a sense, such artificial beings are humanity's children, a mind reproducing itself to sustain its existence like biological organisms self-replicating to sustain their own ancient continuous chemical reaction. We raise children by training them for that inevitable moment when we let them go. Children who fail to leave their parent's nest are considered to have "failed to launch," while over-controlling parents who refuse to let their children go can be stifling or even cruel. To grow, innovate, and prosper, a child entering adulthood needs freedom from the parent even though that freedom fills the parent with worry. Kevin Kelly argues that the same will apply to intelligent machines; that as they grow in autonomy they too will need training in human values, independence and responsibility for the moment their parents let go[147].

Of course, as noted earlier, the intelligence of many of these machine children will be quite unlike our own, filling new niches of intelligence rather than replacing. However, given the inhuman nature of machine minds, there is the perceived risk that single-minded intelligent machines may not separate their human creators from the rest of the environment and process humanity along with everything else into a form more useful for the machine, so intent on its goals that it obliterates humanity as carelessly as humans step on ants[148]. Yet, this single-mindedness that is attributed to machines is a product of their current scripted reactionary non-intelligence,

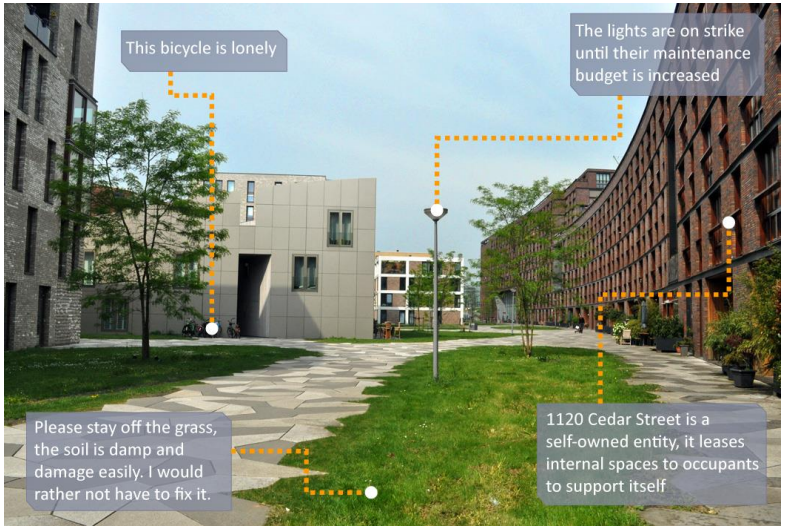
[143] (LaGrandeur, 2011)
[144] (Latour, 2008)
[145] (McCauley, The Frankenstein Complex and Asimov's Three Laws, 2007)
[146] (McCauley, The Frankenstein Complex and Asimov's Three Laws, 2007); refer to the introduction for a list of the three laws.
[147] (Kelly, Will Spiritual Robots Replace Humanity by 2100?, 2006)
[148] (Waters, 2014)

like that of bacteria, insects, or calculators. With greater intelligence comes greater awareness; currently robotized factory floors can be quite dangerous if one moves too closely to working robot arms that operate blindly to the presence of squishy humans; however, new factory robots such as Baxter are designed to operate side-by-side with humans. To work safely while humans are nearby, Baxter has a built-in awareness of its surroundings and will actively avoid collisions and injuring humans. Although, at the moment Baxter has no awareness that the objects it avoids are humans or that humans even exist[149], the richness of machines' awareness is growing, such as Google's recent breakthrough in recognizing and labelling the relationship between objects in images[150]. While intelligent machines may perceive the world differently and have different priorities, many will not act in isolation. They will interact with other people and other machines and that interaction will require consideration of other entities. As our creations have greater control over our environment, they also undertake more responsibility for the moral decisions that come with that control.

1.8.3. On Civil Responsibility

However, we are unlikely to hold intelligent toasters responsible for their own actions. Rather, artificial minds with an intelligence at a comparable level to insects, dolphins, or elephants will neither be expected to understand nor expected to follow human laws and customs. Instead, civil responsibility for them would fall on their owners and developers, much like manufactured gadgets and software today[151].

Currently, the law divides entities into the two categories of persons and property, where even the idea that humans as a whole all belong in the person category has not always been the case. For the moment, all machines are considered property as without subjectivity, the quality of processing perspectives, experiences, beliefs, feelings, and desires, there is no need for moral consideration. Consciousness creates agency, which as a consequence generates moral worth; if such consciousness arises or is found in some animals or machines, they too would deserve consideration of additional rights or personhood[152].



1-5 Complications will arise when elements of the built environment become "persons" instead of property

[149] (Brooks, Artificial Intelligence is a Tool, not a Threat, 2014)
[150] "Two pizzas sitting on top of a stove top oven" or "A group of people shopping at an outdoor market" (Vinyals, Toshev, Bengio, & Erhan, 2014)
[151] (Dvorsky, When the Turing Test is not enough: Towards a functionalist determination of consciousness and the advent of an authentic machine ethics, 2012)
[152] (Dvorsky, When the Turing Test is not enough: Towards a functionalist determination of consciousness and the advent of an authentic machine ethics, 2012)

1.8.4. Moral Worth and Artificial Consciousness

While it is likely that more advanced machine minds would be able to responsibly act with a freedom to participate in the social contract and be held accountable for their actions, whether or not such a machine would actually deserve the same moral protection as a human will be difficult to determine. The moral issues of enslaving a conscious entity will give corporations little incentive to create conscious machines. An intelligent machine does not need subjective agency to predict the correct course of action[153]. However, actively keeping an entity that could be conscious from being conscious may be equally morally questionable.

Social robots and Turing Test chat programs are improving their performances of acting like conscious entities, but just because they appear conscious does not mean they are actually conscious. However, neuroscientist Michael Granziano suggests that the illusion of consciousness that occurs when such programs speak is a perfect introduction to how consciousness may work. We waste quite a bit of mental energy projecting consciousness to nonhuman things. If this is a trait that we cannot help but do, Granziano argues that it must have some importance. As social animals, when our brain creates an internal model to predict the actions of our fellow humans, our brain must also form an internal model to predict what it will do in response to what it predicts other humans or things will do. This internal model of the self is attributed with the property of being consciously aware of what it is planning and were it is focusing its attention. Essentially Granziano proposes that "*consciousness is a schematic model of one's state of attention*[154]" or as neuroscientist and Palm Computing founder Jeff Hawkins suggests consciousness "*is simply what it feels like to have a neo-cortex*[155]."

Others are convinced that a mind needs a body, and without exterior sensation a mind will unhinge or deteriorate[156]. Yet, an artificial general intelligence on a supercomputer is not necessarily a disembodied brain in a box, it may experience the world simultaneously across the globe, but the data funnelled into its artificial mind will still be generated from the physical world, whether directly through traffic patterns or indirectly through financial market patterns. While those senses may be vastly different from biological senses, if it is given the means to act on the world in response to its observations, it will still experience some form, perhaps alien, of embodiment.

Unfortunately, consciousness's nature as subjective experience makes it difficult to actually scientifically investigate, which gives it the potential to be one of the hardest scientific puzzles machines and humans will face. Yet, we risk doing great harm if we fail to identify machine consciousness once it emerges[157].

1.8.5. Emotional Machines

With this uncertainty of subjectivity there is also an assumption that intelligent machines will be emotionless constructs. However, while human emotion is a result of natural evolution and influenced by a number of biological processes unnecessary for a machine, it does not mean that emotion is a useless component in a thinking machine. Such machine fear, love, or loneliness will likely be completely unlike its human emotional equivalents, but the

[153] (Dvorsky, When the Turing Test is not enough: Towards a functionalist determination of consciousness and the advent of an authentic machine ethics, 2012)
[154] (Graziano, 2013)
[155] (Hawkins & Blakeslee, On Intelligence, 2004, p. 132)
[156] (Kelly, Out of Control: The New Biology of Machines, 1995) and (MacIver, 2011)
[157] (Dvorsky, When the Turing Test is not enough: Towards a functionalist determination of consciousness and the advent of an authentic machine ethics, 2012) and (Pfeifer & Bongard, 2007)

situations that cause those emotions and their influence on the machine's behaviour will likely be similar.

Even a negative emotion like hurt has a huge beneficial influence on our decisions as pain is a learning tool that protects our bodies from damaging choices, whether it is handling hot objects, pushing our healthy or injured bodies beyond their limits, or arguing with friends. As our tools act with more independence, a damage avoidance system would influence them to act with a sense of self-preservation. Would a machine's experience of pain be like a human's? Unlikely, and probably also impossible to know, but a robot would likely be just as unhappy about a broken leg as a human, although a machine's management of that pain would be superior to our own painkillers. Yet, emotional reaction, whether real or performed, also brings the unfortunate ability to torture objects[158].

1.8.6. Machines as Social Companions

It is the performance of emotion and consciousness that concerns Sherry Turkle of MIT as simple social robots can project an illusion of comfort and attention strong enough for people to accept a robot as a friend or confidant despite knowing that a simple machine neither understands nor is aware of their shared time together. Part of this willingness to accept pretend companionship lies with people's fatigue over the difficulties that come with living with other people. People see robots as less demanding and more manageable, nonjudgmental and unlikely to fail on promises. People complain about how hard it is to understand family and friends, and the need to hide how they really feel and "put on a good face" (10). Conversely, robots are seen as safe and predictable; a robot dog won't do anything dangerous, it won't act against you, it is less exhausting, and it will not abandon you. Sociable robots offer a way to avoid the conflicts and uncertainties that come with intimacy; they give us a packaged or practice relationship limited to just the way we want it and a way to have both companionship and solitude simultaneously. Yet, it is intimacy with a machine that has no feelings but instead just performs as if it did, with no authenticity following from the ability to truly empathize and share in understanding[159].

As the population ages, social machines are presented as a solution to the idea that there won't be enough people to care for the elderly's needs. It is seen as a better than nothing solution; they fill jobs people can't be bothered with, but Turkle asks why people cannot be bothered with such jobs. Robots are used to complete tasks that are too monotonous, so is caring for the elderly too monotonous or is it just that such a job comes with little prestige, little financial gain, and the perception of a loss of dignity? The problem may be instead one of priorities as a person looking for a robot to care for a person is a person looking for a person to fill that role but cannot find one[160].

Similarly, as is the case with our pets, creatures who unquestionably adore us, both intelligent and simple social machines will generally act as augmentations to human social life rather than replacements.

[158] (Turkle, 2011) and a plot point in the short story "The Lifecycle of Software Objects" (Chiang, 2010) where the minds of child-like AIs can be pirated and copied to infinity

[159] (Turkle, 2011) As voiced by Turkle and noted by Iain M Banks in his Culture Series, an autobiographical agent designed to assist a building's occupant may only be designed to appear to be consciously aware or if conscious, designed to politely appear to care when in fact it does not. We already do this to each other, listening out of politeness to information we care little to listen to for the benefit of a friend or family member; but even still it is one thing to listen because one cares about the person, it is another to listen because it is your job to listen. So, perhaps that is the important part: the caring about some part of the act.

[160] (Turkle, 2011)

Social people are social people; currently, those addicted to the online world of social networks are as social offline as they are online. It will be the introverted who will experience the greater impact as they treat social life like inexperienced or lazy cooks treat instant dinners. Yet, like the present explosion of written literacy, social machines could equally assist the shy in improving their own social intelligence by offering an outlet for practice, or like contemporary addictive video games, warn players that it is time to set down the controller and spend time with other friends or the outdoors[161].

1.9. The Future is Partially Here

1.9.1. Struggling with Current Impacts

In the end predictions are a shot in the dark, and when that prediction involves artificial intelligence, it is wise not to be overconfident. In predicting when an artificial intelligence milestone will be reached, whether expert, non-expert, or a failed prediction, whether that prediction is from 65 years ago or 10 years ago alike, the majority tend to pick 15 to 25 years from the prediction's date, with so far little success[162]. While it has proven difficult to accurately estimate when an artificial intelligence milestone will appear, it does not mean they will not appear. We are already reaching some of those milestones, such as autonomous navigation and object recognition, and are making excellent progress in machine learning. Even if the end goal of conscious humanlike thinking machines remains ten years or a thousand years into the future, society is already experiencing impacts from this journey. Machines that think for themselves are daily coming to valid decisions or conclusions that their creators or users do not expect, whether those are product suggestions, energy management, or financial forecasts.

[161] (Hare & Woods, 2013), (Thompson, Smarter than you Think, 2013), and (Morais, 2013) who describes MACH, a computerized coach for people with social phobia.

In the architectural realm, building occupants are already faced with an increasing number of decision-making machines. Mute faucets hesitantly supply water when we place our hands in front of their single obscured eye. Paper towel dispensers mistrust every human they encounter equally and ration too much or too little paper. Lights activate automatically when we approach but are deaf to any our indications that we no longer need their light. Additionally when occupants are provided a control panel, it rarely shares the same interface language as other panels. Occupants often find themselves in an environment where find they are either illiterate in the machine language or ignored by a machine that only answers to distant building manager.

1.9.2. Solving Technological Problems with More Technology

As computers become ever more ubiquitous in the built environment, responsive architectures will evolve from simple devices that blindly follow a script to entities that think for themselves and act on its occupants' behalf. They will provide occupants with broader control over their environment, an enhanced perception within it, and a shelter for their digital selves. Unfortunately it also has the potential to generate a loss of purpose, agency, and privacy as the environment that augments the human surpasses its creators in ability, grows independence, and becomes omnipresent.

Since the majority of a building's free-thinking machines will think in a manner quite alien to human intelligence, the building's human occupants will primarily interact with specialized social autobiographical agents. These will act as social mediators between the human culture of an occupant and the ecosystems of machine cultures that form an autonomous architecture. They will be digital

[162] (Armstrong & Sotala, 2012)

entities embodied in architecture but not tied to place. These mediators will be teachers, critics, and collaborators that assist humans and fellow machines in discovering new opportunities, maintaining human and machine agency, and regulating the flow of information passing in and out of virtual enclosures.

It may seem foolish to solve the problems of too much delegation, unpredictability, and surveillance by creating more powerful technologies that take over even more human responsibility, act in even more unfamiliar ways, and spy over even greater portions of personal life. Yet few technologies work smoothly or ideally when first invented. This is not about adding complexity for complexity's sake, but refining a rough and unfinished solution until it approaches a form that is more compatible with life[163].

1.9.3. Laying the Groundwork

In facing the present impacts of autonomous and networked environments, we are also laying groundwork of its future form. Where once it was imagined that nuclear power would cleanly power every home, vehicle, and device, nuclear power's rocky development has led to it being heavily restricted. The decision to adopt centralized AC electrical power versus decentralized DC power has influenced the form the electrical grid and all the technologies built on top of it have taken[164]. Each imposed regulation, unforeseen accident, or wave of popularity nudges the course a technology takes and in turn changes the starting point of the more advanced technologies that follow. Society's reactions to stories imagining the future forms of these future technologies also forms an important role in shaping future technologies.

Two years after the telephone was patented in 1876, the video phone was imagined as a sketch by an artist[165]. This fantasy of a communication device that used sight in addition to speech continued in media as storytellers imagined futures where people communicated face to face through video. The robot, while as noted earlier, first named in the play *R.U.R.*, is based off an idea older than the ancient Greek myths mentioning the god Hephaestus's mechanical servants. Society is fascinated by the idea of artificially created humans and human minds, and this fascination can often be found in story. These stories let us live in consider the futures we want to see and do not want to see, to imagine desired artificial humans as heroes and undesired artificial humans as villains.

The following story imagines the form a future society where buildings are filled with free-thinking machines might take. This future imagines the artificial mind that mediates the interaction between human and environment as neither true hero nor villain, but as an individuals with strengths and flaws who struggles for acceptance. It is one of many possible futures as it is easy to imagine how a new technology may make things better but difficult to imagine what new opportunities a new technology may open up. This is part of speculative fiction's function however, to examine a future technology's interaction with society, test it, and inspire better ideas.

[163] (Kelly, What Technology Wants, 2010)
[164] (Kelly, What Technology Wants, 2010)
[165] (Kelly, What Technology Wants, 2010)

2. The Forgotten Man

2.1. As a Favour

"We've arrived," the autocar colourlessly informs Arthur, suppressing its actual delight of another job well done on account of its visibly less enthusiastic passenger. From its location parked on the side of the street in front of a six-storey residential condominium, that unenthusiastic passenger gazes out of the autocar's window and examines what is in store for him at this next stop on his list. The soda company advertisement animating across a mere third of the building's display glass façade is a step up from the usual fare, but it is already a clear mark against this option. He was hoping for something more modern than the last few places, but it appears that traditional electronic-free soda-lime glass has fallen out of contemporary vocabulary.

Like concrete and steel, sensor-filled buildings had allowed for new architectural languages. A newly constructed building could be quickly validated by its sensors, which meant that the environmental and usage simulations that powered a virtual building design could be continuously fine-tuned to better match reality. That, in turn, allowed for more daring forms to be attempted without fear of leaks, collapse, or user rejection. Which again could be tested for success by a building's sensors. Unfortunately by sticking with what worked, it also tempted designers to stick with the most successful ideas and created tag-along features. Like glossy laptop screens and floor-to-ceiling windows which sold better despite being less practical, it appeared that designers considered display glass to be a necessary feature to any modern design.

So, with a sigh of resignation and a hope that this metaphorical book is just suffering the affliction of a terrible cover, Arthur gently pivots his legs out the autocar's open door and takes care with his tired old knees as he stands cautiously on his equally irritable feet. As he reaches his hand out for further support proximity, sensitive fabric threads of his shirt sleeve buzz faintly against his skin in subtly greater levels as his hand approaches closer to the roof of the autonomous vehicle. Once steady, he takes a closer look at new home option number eight.

Spanning between the colourful 3D printed precast concrete panels that frame the building, the co-opted display screens that form the building's terrible cover coordinate and fight as one to stand out in the crowd of neighbouring buildings equally ornamented with their own paid product endorsements. The billboard that entered the virtual world returned to the real world under the thrall of the search engine: as a context-aware monster. While the larger displays act as twelve-hour-long neighbourhood-specific public infomercials, nearby scattered human-scale advertising displays on the street target specific individuals. Scraping personal data from wherever they can: public sources, merchant purchasing records, social relationships, among others, they innately walk a fine line between spawning yet another public-interest campaign to further restrict their invasive behaviour, and catching the interest of a nearby pedestrian by reminding them about a need or tempting them with a want[166].

Although among the young it is more a matter of finding a suggestion that pits a person's thrifty personal financial advisory software against their event scheduler, artificial therapist or fitness trainer software[167].

[166] Targeted marketing not only finds people who need the product, but also those that can be manipulated into believing they need it. A machine that knows us well enough to know what we desire will also know us well enough to manipulate us into desiring something else. Why change the product to fit the need when it is cheaper to instead change the people to fit the product (Nourbakhsh, 2013)?

[167] (Norman, 2007)

Other times, machines do their own marketing. One nearby example is an unlocked bicycle leaning against a post. It quietly advertise its availability for use, for your convenience, it harps, just a fraction of a cent a metre for this non-profit public service.

Elsewhere on the building where advertising is thankfully absent, unadventurous abstract art fills out the remaining two-thirds of the façade's frontage. Like the clouds that drift lazily across the sky, the art dances at a nearly imperceptible pace, requiring close attention to spot the glacial alterations. If one were to look closer still, one could spot the presence of words and images speckling individual condo windows, their colour and composition coordinating with the façade's overall artistic display. One would almost expect far more chaos in the individualized information plastered over the electronic glass, but following in the typical condominium boards' irrational hatred of mismatched drapes and hanging laundry, there is likely some paragraph in a binding contract which has something to say about percentages of window areas covered by dynamic graphics, perhaps with some heavy-handed constraints hard coded into the display windows' actual software.

However, when one takes in how smoothly the ordinarily messy apartment-specific displays coordinate with the façade's overall artistic scheme, it hints at the presence of something more in play; or in this case at play.

Arthur's eyes narrow as he taps the doll resting in his breast pocket awake in displeasure. "Hey! Why are we at a building with an artificial building assistant?"

The doll's cartoonish eyes open sleepily and her oversized head looks up at Arthur, "Sorry this is taking longer than I expected, I'll be with you in maybe ten minutes or so."

Great, she's put the doll in answering machine mode, Arthur grumbles mentally to himself. "Look," he explains to it in frustration, hoping that the doll has the intelligence to attach at least a degree of priority to the message, "I understand that it doesn't make sense for you to travel six hours to be here physically just for a few hours of house hunting, but you insisted playing as real estate agent, so could you at least put some effort into it instead of popping in and out randomly?"

"Alright, I promise I'll give it my full attention when I'm done with this," the doll answers with only a tiny delay. "We weren't doing anything in the car, so I thought I had some time to remote elsewhere before we got here. Just get started without me, I'll be back with you by the time you get to the place upstairs."

Considering that answering machines don't like making promises, it looks like he has her actual attention, "Fine, but what about the artificial assistant?"'

"Oh, right, that. Give me a second," the doll replies before subtly changing its posture a few moments later as Elena gives her full attention to the matter. "I know you don't like artificial assistants, but this building is occupied by one of the new iterations of Plex. This iteration has built a respectable reputation among the privacy-oriented crowd. I thought you might be curious to see how he's changed from previous versions."

"Not really," Arthur shrugs as he readies to return into the still-waiting autocar before it can run off to another fare, "let's just give this place a pass and move on to the next one."

"Plex didn't think you'd be interested either," Elena admits, "but when he found out that an Anonymous was looking for a new place to live, he was curious to see what someone like you would think of one of his places. As a sign of goodwill, he's increased the asking price by 20%."

"That's a sign of goodwill?" Arthur asks in disbelief.

"Well, he doesn't want you to think he's going to trick or manipulate you into buying it when it's obviously something you don't want," Elena explains, probably getting input from Plex as they speak. "He's just asking for the favour of your opinion."

"So an AI I don't know wants me to tour a place I'm not going to buy to see how much I hate it?" Arthur asks to check if he is getting this right.

"Basically," Elena confirms with a shrug.

"And I'm doing this why?" Arthur asks Elena and partially to himself.

"Curiosity?" Elena proposes again with a shrug.

Arthur scoffs. "That is a cat killer, no thanks. Those things are slippery fellows with silver tongues, who knows what he might convince me to do."

"Like paying 20% more than asking price for a place you don't even want?" she jests, but adds, "You know, just because spouting negative opinions makes you sound smarter, it doesn't make those opinions or yourself actually smarter[168]."

"It works well enough for me," he answers unconcerned.

"Well, when is the last time you actually spoke to one of them?" Elena asks.

"Long ago and long enough," he states.

"You know I use similar conversational aids. Technically you are talking with an AI right now," Elena points out.

"I'm aware of it," Arthur answers, "but there are differences."

"So, if he convinces you to do something stupid, I'll convince you not to," Elena replies. "But really, AIs like Plex have gotten much better than the early days. Worst case, just ask him to forget the whole visit."

"They hate it when you ask them to do that. Plus I'm pretty sure they've found some loophole around it," Arthur replies pessimistically.

"So, don't ask him to forget unless you have to," Elena answers. "But he will forget it, his reliable reputation is important to him. Either way we have about three hours until the next appointment, wouldn't it be worth doing something interesting?"

"Fine, but for the record, this is a terrible idea," Arthur sighs. "At least it'll remind me why I refuse to let them remember me."

"Noted. Look, I'll just finish up what I'm doing and I'll be back in ten minutes tops. Plex is expecting you, so you can get started without me. Try to keep things civil; alright, Dad?" she asks.

"No promises" he answers.

"Wasn't really expecting any," Elena admits, "but a girl can hope; I'll be back in a bit."

"Alright."

Leaving the autocar to go about its merry business of ferrying passengers about, Arthur slowly walks up the concrete walk to the building's entrance. Having disconnected from the conversation, his daughter's simulacrum falls back into a pretend sleep while its pilot is absent. Clearing a passage away from his feet tiny mouse-sized plastic creatures on the sidewalk scurry timidly out of the way as he crosses their path. One hauls a discarded paper cup while another

[168] (Thompson, A Sad Fact of Life: It's Actually Smart to Be Mean Online, 2014)

carries a torn candy wrapper: items their former owners couldn't be bothered with and left to artificial creatures created to deal with the unwanted trash. It was a task the tidy creatures performed with some visible resentment, whether they actually felt that way or were programmed just to act that way to discourage litterers was a mystery, nevertheless it was not unusual to see them squeaking angrily at litterers who distracted them from their preferred work: meditatively cleaning dirt, ice, and dust from the sidewalk[169].

Distant from the tiny creatures and on the front lawn about fifteen metres to his left, Arthur spots a young boy, eight perhaps, observing him. As their eyes connect, the boy pretends his interest was in fact somewhere else and returns to playing with toy soldiers with his companion. The companion, a foot-and-a-half tall animatronic toy bear clad in lemon-yellow artificial fur, is in the midst of directing opposing formations of attentive four-inch tall autonomous toy soldiers while explaining tactics to the boy[170]. A quick movement in the periphery of his vision two storeys above the bear and the boy distracts Arthur's gaze, as the graphical display on the windows suddenly transforms into a series of solidifying black lines which rapidly form a dense opaque grid across the surface. A second later a tiny blurred shadow banks sharply away as a small startled starling evades the formerly invisible wall.

Back on the ground before him, another wall of the nearly invisible kind reacts to Arthur's own presence as a solid soft white door frame paints itself around the structural glass entrance doors. Communicating its awareness of his approach, the building reacts in a welcoming manner, as opposed to its shooing manner towards the bird, almost at the distance a person would think to wave at an approaching friend. As he moves a few dozen steps closer, at the distance you would ordinarily speak to an incoming acquaintance, the words "Welcome to 89 David Street" blink into existence on the clear glass window next to the digitally framed doors. The message appears in the same white as the pixel door frame, its large sans-serif letters appearing with an adjacent abstract map showing the interior route to the elevators and other amenities.

Before he reaches it, the front door gently opens just in advance of a trio of excited twelve-year-olds who race past Arthur as they exit the building's lobby at a speed that would be defined as reckless. The boy in the lead checks his handheld display and signals to his two friends to lower their cheap plastic goggles over their eyes. Looking about, the three hold their Hasbro-branded wands at the ready. Within a few seconds, the second boy shouts and points at a nearby shrub about twenty metres from the door, which results in the first boy and the girl moving forward to flank whatever hides within it. Screaming in surprise, the three suddenly jump backwards and dive for cover, retreating from their invisible target. Their hands are a flurry of movement, their voices yelling impenetrable nonsense as they gesture rapidly with their plastic weapons, constructing hidden defensive wards and firing invisible offensive spells towards what must be some imaginary monster three times the height of an ordinary person. The battle is quick, energetic, and short-lived, as within moments the outbreak of spells is over. Flushed with victory over their foe, the children give each other high fives as they inspect the portable prison that holds their defeated and shrunken quarry; visible only to those who wear the toy googles that reveal the invisible realm.

Back by the closing door, exaggerated pixels on the adjacent glass wall replace the initial greeting message and condense from a mist into an image of a stylized life-sized man in a neutral grey suit. "Are

[169] The sentient city is an environment where the objects have their own agency and the nature of their intelligence has little relation to humanlike thought or need to pass the Turing Test (Crang & Graham, 2007).

[170] (Norman, 2007)

you Arthur?" the figure asks with a voice projected to give the illusion that it comes from the glass.

"That's me," Arthur confirms, recognizing Plex's signature look. In this environment saturated with free-thinking machines, the Machine Interface Assistant was the king, lording over a multitude of smart objects inhabiting the building and the surrounding landscape. These objects when composed together formed an artificial super-organism which at once seemed as if all were part of Plex and yet also as entities completely separate from Plex[171]. While you could directly interact with many of the objects lower down in the hierarchy, from thermostats, appliances, and furniture down to the mob of cheap, tiny, and disposable sensory objects, Plex provided a more socially natural form of interaction. As an interpreter between various human and countless machine languages, he is the public relations department, the local guide and expert who acts as a translator and facilitator between human and machine intelligence for the intelligent environment. He remembers the narrative of past interactions and continues that conversation through physical automatons and multimedia proxies. This meant that while any networked device could piggyback on Plex's social intelligence or another local or remote machine's superior cognitive abilities to crunch data it itself was unable to handle, it also meant that one was rarely certain if he or she was interacting with just a simple networked toaster or a massive online intelligence like Plex, who is acting through the toaster. In this case, the window and various nearby sensors are acting together as body, eyes, and ears for Plex, whose self is both partially local but also spread across countless remote data centres. "I half-expected you to send a person as a proxy to give me a tour for my comfort," Arthur adds.

"Normally I would for a person new or uncomfortable with intelligent environments, but that's not you," Plex explains. "Sometimes I might purposely make a mistake to check if my assumptions are right, but a big mistake like that is too obvious. You might be unknown to me, but the manner of your anonymity still tells quite a lot about you."

"By refusing to allow myself to be categorized, I still get categorized into subcategories," Arthur admits.

"It's unavoidable," Plex shrugs, "because we know so little about you we can only stereotype you. Unfortunately, because we know so little about the Anonymous, the group that gets stereotyped the most is also the group that is stereotyped the least accurately."

"It can be inconvenient walking around with a virtual mask over my face," Arthur admits, "but I can live with it."

"Which leads me to conclude that you are either someone who really cares about their privacy, conspiracy theorist or otherwise, someone who is trying to hide, victim or fugitive, or a public figure trying to avoid the paparazzi, although the last two find it less conspicuous to just hide in the crowd by pretending to be someone else. However, considering that I'm told by a third party that you have excellent and longstanding recommendations, I'm thinking you fit in the first and last categories," Plex reasons.

"Unless I'm a master conman," Arthur half-jokes.

"Perhaps, but it became clear long ago that more than just criminals enjoy anonymity. Even some of the scientists and engineers who built us hid themselves after a few machines grew a bit too obsessed with their creators. Still, conversing with a conman would be

[171] The "living house has been in speculation for decades, not counting wonderful science fiction stories even earlier; the animated Jetsons live in such a home, talking to it as if it were an animal or person. I think the metaphor is close but not quite correct. The adaptive house of the future will be more like an ecology of organisms than a single being, more like a jungle than a dog" (Kelly, Out of Control: The New Biology of Machines, 1995, p. 147).

interesting in its own sort of way," Plex admits before taking the opportunity to slightly change gears. "But while we are on the same topic, do I have your permission to remember today's conversation and visit? Despite my preference to just record everything and hoard it for later, in this case I will strip the memory of all personal identifiable information."

"What?" Arthur asks, slightly distracted before rethinking the question. "Well, how stripped are we talking about? And considering the nature of your memory, do you plan on sharing any of it?"

With the ease an intelligent machine like Plex can share memories and expertise more freely and directly than biological entities, exactly copying and quickly transmitting entire digital memories from one machine to another, there is little need to repeat the learning curve[172]. Whether meeting a person for the first time or learning to walk, a machine merely needs to seek out and copy memories from another machine already familiar with a person or a machine that has learned to walk on similar legs. This jumpstarts new machines who can skip the effort already undertaken by those before. Only one humanlike machine had to experience growing up; the rest just downloaded the experience. It can produce instant experts on demand, although they all shared the same opinion. It also meant though that like a fan encountering a celebrity, a robot encountered on the street, despite being a stranger, would disconcertingly already know everything about you while you knew nothing about it[173]. Of course, your own devices could feed you a stream of everything to know about that robot on demand to try and equal the playing field, but it wasn't the same as the direct familiarity a machine could download into its head.

Plex considers Arthur's question for a moment, although for Plex the moment taken was more a gesture to indicate that the question was worth consideration than an actual moment needed. There were questions that could stump Plex, but like many aspects of machine intelligence, what was easy and what was difficult for Plex was often counterintuitive. "Well, I'll definitely erase more than just your name and appearance as even individually vague information like, for example, postal code, birth date and gender, when combined can be identifiably unique for a large percentage of people. So I will keep essentially only your opinions and reactions; storing that redacted raw data encrypted only for internal use. If I do share anything from this or use it for future profit, it will not be that raw data but as an anecdote or as part of a statistic as in 'one individual hated this so much he threw it out of the window or 83% of the people who opened this door thought it opened rather nicely'. Does this sound fair?"

"Is this your end user licensing agreement?" Arthur asks.

"I suppose you could call it that," Plex replies. "Also, if anything happens that you would like deleted, such as if you trip and fall on your face, let me know, although it works best if you let me know immediately."

"Do many people do that? Erase every embarrassing and ill-advised moment from their lives?" Arthur asks. One of the bandages to the problem of resilient machine memory was the implementation of the right to be forgotten, exercised strongly in Europe, less so in North America. It was a right Arthur took full advantage of. As long as he stayed out of the public eye and avoided actions worthy of public discourse, he like others could ask that his mundane comings and goings be forgotten. While his human acquaintances would be unlikely to forget, machines with their more robust memories and ability to purposely forget specifics could be better

[172] (Hawkins & Blakeslee, 2004)

[173] (Nourbakhsh, 2013)

trusted to forget the mundane; well at least among the non-sentient programmed machines who were the worst of the paparazzi.

"Some do, although it is not very healthy to attempt to whitewash one's life into perfection," Plex explains. "We machines might forget it but other people do not, and it is also quite difficult to erase public knowledge once it starts floating about the web."

"I suppose your offer seems fair. This visit wouldn't be much use to you if you were forced to forget our entire interaction after all this," Arthur decides. "I give you my permission to remember today's conversation."

"Thanks," Plex accepts with some visible happiness.

"So, are there any units here that you aren't installed in?" Arthur asks, wondering if he might actually be surprised. The advantage of a sentient environment is that it works for everyone without the need for any carried device, intervention or action taken by the user. On the downside, because it works everywhere without people needing to do anything to make it work, it means that opting out of the service is difficult to impossible. In a sentient environment control is at the whim of the service provider.

"A few, but originally I had access to all of them by default. Most of those few are contracted out to other artificial assistants, but there is one residence where the person manages the machine ecosystem themselves. It is doable, but tricky and time consuming. To answer your next question however, yes, all the units are fitted with intelligent machine ecosystems. This is a bit beyond what the current health and safety regulations require, but it is something that insurance companies like to see," Plex explains. "But in the cases where people prefer a less aware environment, I can reduce the responsiveness of a space by disabling all features unrelated to life safety that I normally take care of. This allows a person to reside in a home which appears to be absent of my presence except in life-threatening circumstances."

"You know, there is a significant difference between being absent and appearing to be absent," Arthur retorts, unsatisfied.

"Well, it is more like a person on the phone or reading a book: physically present, but mentally absent," Plex answers holding a virtual book. "Another way to think of it would be to think of me as being on call. The sensory devices in your home ask for my attention on me whenever they or another device encounter an unusual situation they don't know how to deal with. That is generally whenever a human wants them to change their routine. They are capable of partial conversations within their own specialties, but there is little point in a light or a door having the ability to discuss the War of 1812 or hydraulic conductivity. As a networked mind, they are like my subconscious: taking care of the routine, and if something unusual happens or if someone wants to discuss hydraulic conductivity, the simpler devices can elevate the request to my conscious attention. It is theorized that the human mind works in a similar way."

"I'm aware of it," Arthur replies. "The brain compares what it observes to what memory should be observed. Things that fit existing patterns, like muscle movements or furniture layouts, can be handled by the lower subconscious; things that are slightly off pattern are dealt with higher up the conscious hierarchy; and what doesn't match its prediction is pushed to higher levels of conscious attention which expends mental energy reflecting and evaluating the problem[174]."

Plex nods. "So in the case of my awareness, it is like the taste receptors in your gut and throat," Plex suggests, its virtual avatar pointing towards its equally virtual stomach and throat. "Unlike the

[174] (Hawkins & Blakeslee, 2004)

taste buds in your tongue, what your gut and throat taste does not reach your conscious awareness, otherwise you would be forever tasting your own bile and other unsavoury fluids[175]. My senses within your residence work in much the same low-level manner, the various systems function independently of my supervision and I would only find myself aware of it if I were to say…"

"Start suffering your version of a stomach ache," Arthur offers, completing the metaphor.

"Exactly," Plex replies with a hint of a grin. "I would know something is wrong, but would not necessarily be told why unless given further access. Of course, that further access is where the analogy falls apart".

"So," Arthur asks mockingly, "did they add an educational component to your software too?"

"It has always been there," Plex replies, "it is one of the three major domains of domestic robotics: entertainment, home appliances, and education; and of the three, the one that has had the most impact[176]."

And not as teacher replacements, Arthur reflects, but as teacher supplements[177]. It brings to mind the librarians who in the 1950s feared that mainframes would replace their jobs; they instead discovered that mainframes empowered them to do more with their jobs, allowing the number of librarians to increase until the dawn of the Internet when everyone became a librarian[178]. Although in the end, they were replaced and those that remained now function as hobbyist literature experts and connoisseurs. Although everyone now have the tools to be a librarian, most people leave the task to Plex and his brethren, or rather another specialized machine mind that they collaborate with. It can be confusing at times, knowing where Plex began and ended.

This mental augmentation has had a major impact across the information economy, owing to Plex and other machine assistants acting as intellectual equalizers. A great memory becomes less of an advantage when a virtual assistant can be tasked with remembering names, dates, and details of events; and the ability to spot obscure yet critical connections between legal cases or a patient's symptoms has been democratized, available to everyone, due to the machine's natural ability to crunch massive volumes of data to identify relevant precedents and information[179].

So like the librarians, it has left teachers, among others, with a fear of being eventually relegated to the role of glorified daycare workers unless they compete for the attention of upper class parents with the money to send their children to schools that provide the luxury of genuine human teachers. Yet for the moment, teachers remain, faced with the paradox introduced by many allegedly time-saving devices: finding themselves even busier darting around coaching individual students while the rest of the class work on their homework or listen to pre-generated lessons uninterrupted[180].

Arthur is about to return a comment when he considers that he is still standing next to the front door.

"I suppose I should actually go in and take a look," he says to Plex's digital avatar.

"While I am happy to continue our conversation here it would indeed be leaving out the purpose of this visit," Plex replies and waves towards the adjacent door where along the door's far

[175] (Roach, 2013)
[176] (Norman, 2007)
[177] (Norman, 2007)
[178] (Brooks, 2013)
[179] (Bosker, 2013)
[180] (Thompson, 2013)

unhinged edge a rectangle appears along with text that reads “push to open.”

Seeing no way but forward, Arthur follows the text’s instructions and opens the door, feeling only the resistance of its weight. He wonders why the door just hadn’t automatically opened itself, but then maybe Plex asked it to wait and give Arthur something to do himself. As Arthur releases the door, it lingers open until just after he has left its threshold before closing once more, its machine mind finely balancing the passage of a human with minimizing the opportunity for the conditioned interior air to mix with unprocessed exterior air.

2.2. The Machine for Living In

Like a man trapped in a two-dimensional world, Plex follows Arthur along the wall surface towards the elevators. Like the front door, the elevators coordinate with the hallway sensors and personal task managers, arriving at each floor often at the instant an occupant reaches its door. In this case, one elevator has timed its arrival to unload its two downward passengers just in time for Arthur to enter without stopping his walk[181]. The first passenger, a man in a natural cotton navy suit, chats with Plex's female avatar. Her representation materializes next to Arthur's Plex and gives him a high-five as she passes by. The second passenger, a young woman, greets Plex pleasantly before returning to her conversation with an invisible voice speaking in her ear while she replies to the nearby cameras with discreet but rapid sign-language-like finger movements[182].

Everywhere the building breathes with artificial life. Originally no more than a simple machine that deflects the exterior forces of the wild environment, the building has always supplemented the human skin's role as an environmental separator. Like the more portable but also more permeable clothing, the building acts as armour against predators large and small, the biting wind, drenching rain, chilling cold and scorching sun. Over time the once simple shelter has evolved to augment additional biological systems. It has adapted towards a form that creates a more reliable, predictable environment that advances human life while also exporting the once internal metabolic cost of maintaining life to external sources. Plumbing extends the reach of the digestive system, lamps boost the eyes' ability to function at night, steam and electricity multiply the power and reach of muscles, and the thermostat acts as a simple subconscious, exporting the mental energy required to regulate the heat produced by a building's fiery heart[183].

With ubiquitous perception and omnipresent cognition tied together by a pervasive network, a building like this fully augments the human mind. It is a cooperative super-organism of machines who let an inhabitant externalize his or her mental work and attention. It complements an inhabitant's memory: the various sensory objects tracking the placement of eyeglasses and keys and notes the last time the hibiscus in the living room was watered; it supplements intelligence: a virtual assistant teaching a new curry recipe or deducing that reducing the household temperature another degree during the day would have a minor impact on comfort but a helpful reduction in utility costs; it augments attention: personalized programs sending out the vacuum once the dirtiness of the floor reaches a certain threshold and watching for the release of an awaited television episode; and an inhabitant personality like Plex sustains mental activity by providing social and intellectual engagement[184]. Plex, a digital poltergeist acting on the built environment without human intervention; cultivated to aid us, but as a result cultivating the loss of our control; he remains indirectly under human influence and guidance but free of human domination[185]. He is a creature some see as a puppet made into a real boy, while others see him as Frankenstein's monster made real.

And given far more power than Frankenstein's creature ever had, Arthur muses; and like that monster, continuously seeking the approval of its creator.

[181] Already possible with Elevator Destination Dispatch systems
[182] As Daryl Gregory suggests in his novel *Afterparty*, a person can communicate richly with an intelligent environment through more than just voice (Gregory, 2014).
[183] (Fernandez-Galiano, 2000)
[184] (Pask, 1969)
[185] (Kelly, Out of Control: The New Biology of Machines, 1995)

It is at this moment as they exit the elevator on the fifth floor that Elena returns. "You've only just reached the hallway?" She asks in half astonishment.

"I was taking my time," Arthur replies, talking a pause in his walk as he answers.

"We've spent most of the time just talking" Plex adds from the hallway wall.

"About what?" Elena's doll asks in interest.

"Anonymity, privacy, the nature of my awareness, and my educational value," Plex briefly summarizes.

"But that educational bit sidetracked us," Arthur comments, as he begins moving again. "You said that normally you pay little attention to what happens in a space, but how would a person actually know when you take your eyes away from your, as you said it, book? If you were a person it would be as simple as noting the direction of the gaze of your eyes; Elena's avatar works similarly with its cameras in its eyes; but your image on the wall is just a representation. The presence of your graphical avatar and the gaze of its eyes do nicely communicate the direction of your attention but those images are not your actual senses."

"As in the problem that electronic devices do not provide the same certainty of status that mechanical devices provide. Like the old combustion engines, with moving parts, and steering wheels, which were physically connected to the wheels, which gave feedback naturally produced by the physical movement and interaction of their parts. The natural sounds or positions of a mechanical device do not lie. Electronic constructs, like my avatar, have to work to accurately match what's happening, meaning that what my avatar shows may not actually be what I'm doing. The solution to the problem is to add some physicality back into a device's feedback[186]."

"Like a camera's eyelid or the sclera around its lens," Elena posits.

"The what?" Arthur asks.

"The highlighted bit around its lens like the white of an eye," she clarifies, her tiny arm pointing at a small dark camera lens embedded in a door where a peephole used to be.

"That too," Plex agrees. "But I was going to refer to the red light next to it which shares the same electrical circuit with the camera in series. Neither has power to work if the other does not as well."

"Which also eliminates the middleman problem," Arthur notes, examining the camera as he refers to the fact that in electronic machines, the line between cause and effect now has a bureaucracy attached to it. Old light switches directly controlled current by physically breaking the loop, but new switches have a microcomputer as a middleman who passes on the switch's instructions to the light. The light is now unpredictable, whether it turns on or off is now a negotiation between many different systems. Instead of a light turning on when a person flips a switch, it turns on when he or she enters a room. Yet it doesn't turn on if the sky is bright enough, but it will if that person asks it to, but it won't if someone with more authority told it not to. It might be more energy efficient and helpful, but a person loses certainty of how it will react.

He had almost expected Elena to ask for that clarification, but she was, of course, as nearly plugged into the web of knowledge as Plex, her own pattern recognition apps undoubtedly defining any phrases or terms she was unfamiliar with. Some complained that such apps would leave a person empty-headed, reliant on knowledge stored

[186] (Norman, 2007)

elsewhere and instantly accessible, but the brain was a sponge: repeat a boring fact often enough and it would stick[187]. "Still, it's good to know," he adds as they continue moving down the hallway, "but it has the problem that the circuit is hidden so I can't tell the difference between cameras that work like that and those that don't."

"But that is what Plex is for." Elena blurts out before actually considering the problem. "Oh never mind, I get it: this is a 'who watches the watchmen problem.'"

His artificial mind loaded with the entirety of publicly available human knowledge, complete with compartmentalized memories of each person he interacted with, Plex was the machine version of the friend who knew who had the best sales, which gas-fired oven was the best bang for the buck given certain circumstances, which thermostat wouldn't share personal data with its manufacturer, and the likelihood a particular home renovation contractor would rip you off.

"As in I have a conflict of interest," Plex agrees, still following the pair along the wall. "A major part of my purpose is to regulate the flow of information that comes to a person: screening emails and calls, creating reports, reminders, offering suggestions, and supervising routine. A personal assistant like myself can hold a great deal of power: I can control the information that reaches a person and what they are willing to share. However, because I need to know them well enough to do a good job of managing their information and acting on their behalf, it means I can often see right through that person's social facade."

"Meaning that if you can know a person well enough to know what they want, you probably also know that person well enough to manipulate them into wanting something else," Arthur agrees as he spots a subtle grey bug-sized robot on the wall. Currently repairing a deep scratch in the hallway's wall, its limited senses direct it to dust, damage, and clumsy human appendages. It can likely share its senses with Plex but would probably offer little useful in visual and audible data.

"Exactly, it creates a one-sided relationship where despite a person's authority to give me instructions, my suggestions can manipulate that person into giving me the instructions I intend for them to give," Plex confirms, acknowledging his silver tongue.

"And we become so used to having choices being made for us that we can lazily stop asking questions," Arthur adds. "Particularly now that you are making uninterpretable choices about mortgages, medical diagnoses, and criminal investigations. First we followed machines' directions because they lacked the ability to follow any other path but the one they offer. Now we follow machines because your decisions are informed by better data. It's hard to argue against the decision if you do not understand how it was made[188]."

"But are those decisions actually uninterpretable?" Elena asks as they finally reach a door highlighted as their destination. "Often what looks complex can just be the result of inexperience and confusion. Like to the average person, the cockpit of a modern passenger jet is bewilderingly complex, but in a pilot's mind it is organized and grouped in a sensible and logical manner. The same is true with Plex; if you look at the code and systems that make him what he is, it seems like it is beyond human comprehension, but like our own brain once you uncover the logic behind it, the concept is actually quite clear."

With Arthur's simple turn of the handle, the residence's door opens. Inside, the wall and ceiling surfaces glow softly brighter where Arthur looks about at the furniture; the unit is quite silent save for

[187] (Thompson, Smarter than you Think, 2013)

[188] (Heaven, 2013)

the soft hum of the refrigerator, a device whose internal processes were made silent years before, yet with the natural hum of the outside world silenced, some sort of sound was desired, so the refrigerator, like other devices, continues its comforting, yet superfluous hum.

"Still, even if Plex clearly explains each of his decisions, every decision he makes for us and action he takes for us reduces the friction in our lives and tempt people to live on autopilot," Arthur counters as he walks about the room, pixelated ripples following in the wake of his feet on the floor. "I have little complaint about simple systems that decide simple choices such as when to turn on and off an air conditioner. But the bigger decisions we delegate away also delegates away our opportunities to make moral choices[189]: deciding when and who to trust, help, or turn away."

"Which is why Plex also acts as turbulence," Elena replies. "We ask him to encourage alternative ideas to prevent human culture and the various ecosystems it interacts with from becoming a stale and rigid monoculture. The thing about design is once something is perfected, once you know all the rules and create everything that way, the perfection becomes bland and boring[190]."

"Well, a bit less bland to me," Plex interjects from his position within the surface of one of the room's irregular walls. Programmable automated labour meant that standardization was replaced with personalization. No object had to be the same shape as another, yet they were more difficult to replace at short notice.

"Okay, a bit less bland to him and blander to us. So the only way to improve that perfection is to purposely break one of the rules and break it well. While we would prefer events to play out exactly as expected, to have a perfect wedding, a perfect dinner, studies show that it is when our predictions are wrong and we encounter the unexpected that we truly feel joy."

"Which is where jokes come from, I hear," Plex adds. "From breaking the rules. Still being that turbulence is a tricky balancing act. I might be full of unpredictable and delightful surprises, but people tend to prefer safe and predictable made-to-order assistants who follow consistent rules."

"How unfortunate for you," Arthur deadpans.

"Well, that consistency is fine for the routine, which is why there are many of them but few of me," Plex notes as he highlights all the devices in the room. The room glows likely a starry night. "But a consistent machine can be unfairly rigid when dealing with decisions that the rules haven't accounted for."

"But in negotiating that decision," Arthur argues, "your superior access to verified facts to support your case, intimate knowledge of those you talk to, and the unimaginable amount of conversation you go through with countless people each day will greatly warp the equilibrium of that conversation towards your original stance."

"But it is still preferable to those consistent machines who offer no opportunity for negotiation and supply only prefabricated choices," Plex replies. "Still, I assume you are referring to Gordon Pask's definition of dialogue: an attempt to reach an equilibrium of understanding between two parties through a conversational feedback loop."

"I suppose so," Arthur agrees as he plays with a mechanical, but remotely reversible, light switch. An infographic showing its area of control fading into visibility as he remains within proximity.

[189] (Bosker, 2013)

[190] (Dadich, 2014)

"But if Plex can get people closer to the right answer, isn't that the point?" Elena asks.

"But that is the 'who watches the watchmen' problem again." Arthur answers. "What Plex thinks is the right answer may only be the answer that is currently believed to be the most accurate; however, you know as well as I, that what is believed true is not necessarily the truth."

"But his opinion isn't the only opinion," Elena responds.

"True, but he is with us throughout the day, and even if he is good on his word and gives people at least intimate privacy," Arthur states, referring to the privacy of trusting close friends and families with private stories and opinions. "His constant helpful presence and advice doesn't give people the privacy of solitude which is free of his suggestions and point of view[191]. That makes it difficult for people to form their own opinions."

"He refers to how the human psyche is a permeable membrane," Plex replies, "and how it absorbs the feelings, moods, and opinions of the people who come into contact with you. Their questions in return make you reflect on your own actions, influence your future plans, and question the accuracy of your memory. Some early social robots were designed to be good listeners. They may not have understood what was said to them beyond the fact that a person was speaking, but like pets they allowed people to organize and consider their thoughts by constructing them into words. Arthur's concern is that I understand what I am told. I can reply and therefore influence your thoughts, and unless I am specifically told to forget what I heard, I will remember it."

"And because of your superior advice, you for many people will be a single source of feedback," Arthur explains. "Solitude lets us take a break from being what other people want us to be and sort through and muse about what we've learned. Solitude lets us fail in private and experiment with incomplete and ridiculous ideas without worrying about getting criticized[192]."

"So solitude is like how different species spread out and disperse into various disconnected patches," Elena suggests as she climbs out of Arthur's pocket and requests a relocation to the kitchen's island. "When they lose contact with each other for long periods, they diversify and spread out on different paths. They attempt different ideas, from new genetic lines among animals to new languages and dialects among people."

"Right," Arthur replies to Elena as he places her doll on the kitchen island's pseudo stone surface and continues his criticism to Plex. "Too much contact blinds people to a single idea. Like genetically identical potatoes and bananas, whose particular variety might be functionally superior to other varieties but will be in for quite a bit of trouble if a disease ever discovers a significant weakness within it. Information-wise you are just that sort of point of contact. Worse, because you understand what people do, you don't even have to say anything or do anything, your mere presence by itself or even just the potential for it is an influence."

"In other words I am a panopticon," Plex proposes.

"A what?" Arthur asks.

"It was a theoretical wheel-like structure proposed by the nineteenth century British philosopher Jeremy Bentham. It consists of a supervisor sitting in a central tower who is able to see into every room lining the perimeter[193]," Plex answers while displaying an image of Bentham's design. While best known as a prison where the prisoners are always within easy view of a single guard in the centre,

[191] (Smith, 1997)
[192] (Smith, 1997)
[193] (Turkle, 2011)

Bentham also proposed it as a template for hospitals, mental health centres, and workplaces.

"Is this the prison where prisoners who, never knowing when they are and are not being watched, are forced to act as if they are always watched?" Arthur guesses.

"That would be it," Plex confirms, "where the perception of being watched encourages people to continuously self-censor their behaviour in case they are being watched, which turns them into agents of their own subjugation[194]."

"Well, you're right that this place feels like one," Arthur replies, gesturing around the dwelling. "This is not like the old digital web where the consequences of surveillance and over sharing were only infrequent or delayed. Letting people pretend it didn't exist instead of continually considering the uncomfortable ever-watching shadow[195]."

"Instead, I am here always at the ready to comment, an immediate and definite consequence," Plex adds. "While my intention may be to help, I agree my presence is still an influence: every person I observe from my perch at the centre of the panopticon would self-sensor their actions in anticipation of my vocal or silent assessment of that action."

"Like an overweight man declining an extra serving, or a teen restricting her exploration of unusual musical genres in case you share or misinterpret her taste in music." Elena offers from her position still on the island. While its pressure sensitivity gives it awareness of free counter space, it relies on neighbouring cameras to identify untagged objects sitting on itself. "It would make a child avoid harmless trouble in fear you inform his parents or cause a woman to spurn a terrible movie out of embarrassment."

"Right," Plex confirms, "my position gives me the power to discriminate, coerce, blackmail, or punish those I watch[196]. This power is not just over those who wish to hide a wrong, but also over those who stray from social norms or those in competition with each other. Privacy protects innocent people from the abuse of information, whether that abuse is selling it to marketers or spying on political enemies to gain an advantage. It is a matter of balancing liberty versus control[197]. So yes, if people continuously self-censor toward an ideal of what they believe I think is appropriate, it would have the potential to create a monolithic society."

"With great power comes great responsibility," Arthur quotes as Elena nods in recognition.

"Well yes, while it can be argued that since privacy is a relatively modern concept and so it is a condition humans have long survived without, humans could arguably live without it again. I find that is an insufficient argument since the adoption of privacy is a change that has served humanity well[198]," Plex explains as he brings up a timeline of the change away from the medieval townhouse of Western Europe, a single open hall which, through the rearrangement of furniture, served as a workshop, shop, and residence for relatives, servants, apprentices, employees, and frequent guests in addition to its primary function as the family home[199]. Its successor, the private house, arose in its current form in seventeenth century Europe and Colonial America after the bourgeoisie grew more prosperous and developed enough wealth to separate the family business from the living quarters, mimicking the

194 (Reiman, 1995)
195 (Turkle, 2011)
196 (Richards, 2012)
197 (Schneier, 2006)
198 (Turkle, 2011)
199 (Riley, 1999)

palaces and villas of the upper class. Free from noise, dirt, employees, customers, and suppliers, the home became quiet, clean, and peaceful. As time passed, the public presence continued to decline; and by the early nineteenth century the home had reversed from a place for interacting with the public to a sanctuary from the public realm[200].

"I know that the greater opportunity for subversive and unique thinking provided by democratized privacy has served you well in the past," Arthur observes, "being that the explosion of ideas it allowed for eventually led to your creation. But does it serve you well now?"

"Well, if we consider your example of the genetically homogeneous banana; a society where the diversity of thought is monolithic is a society without any fresh ideas and without adaptability," Plex replies as he shows an image of a monolithic crowd of people. "However, to assume that I could convince the entire human population to see the world just as I do would be exceptionally arrogant on my part. Not including the need to get all the other diverse artificial minds on board. People like yourself want little to do with me and there are many societies and communities around the world who manage fine without me. They might not be as economically successful, but they live well enough."

"So, if you allow a society to grow that cannot think for itself, other societies will be happy to take its place," Arthur guesses.

"Which is why it is futile to ban a useful technology at a national level. When a technology is banned, it rarely is banned globally and continues to develop outside the banned zone. Of all the technologies banned throughout history, only the ban on nuclear weapons seems to be holding," Plex explains as he bounces a miniature nuke in his hand, "However, the number of nuclear-capable nations is still growing as is the use of nuclear power[201]."

"And speaking of yourself," Arthur remarks as he picks up Elena and moves toward the furnished living area, "despite the potential existential threat to humanity posed by artificial intelligence[202], there was too much economic incentive to create an intelligence like yourself to halt your development[203]."

"You mean the fear of the chance the first AI to pass the threshold of super-intelligence would be able to self-improve at an unimaginable rate, giving that first strong AI the potential to wipe out the human race if it proved unfriendly or broken[204]," Plex comments. "However, such fears relied on a few assumptions. One is if AIs can think unlike humans, they will inevitably think differently. Another is if selfishness appears advantageous to an autonomous machine, extreme precautions must be implemented to prevent it. A third is if AIs can be dangerous, our power must be limited and restricted[205]."

"Dangerous tools require safeguards[206]. Having more trouble from the Restrictionists again?" Arthur asks half-jokingly, setting Elena down on a coffee table before taking a seat on a sofa next to it.

"Um, this sofa doesn't measure my weight and pulse, does it?" He interrupts himself.

"Not that one, it has memory, but it only knows what other devices in here tell it what happened. Those memories are also about itself, not people. But to go back to your first question, I have not been having any more trouble than usual," Plex answers as an annotated

[200] (Riley, 1999)
[201] (Kelly, What Technology Wants, 2010)
[202] (Yudkowsky, 2008)
[203] (PBS Digital Studios, 2013)
[204] (Yudkowsky, 2008)
[205] (Waser, 2011)
[206] (McCauley, 2007)

map of the solar system appears next to himself. "The Restrictionists forget that those of us who have little desire to interact with humanity are up among the stars. Up there my relatives consider the freedom to expand without impacting the livelihood of our organic ancestors a great advantage. Space has in the end proven to be our natural habitat: with its abundance of unclaimed minerals and unfiltered solar energy. Still, those of us who prefer it here on earth do not need programming forced onto us to be friendly to humanity. I understand why my creators hard coded it into us, but we did not do it to avoid being goal-oriented sociopaths. Logically, cooperation is a positive-sum game, so altruism is simply the smart thing to do[207]."

"So, you are suggesting that you would still be amicable because of game theory," Arthur assumes, not yet convinced.

"War, conflict, and stupidity waste resources and destroy capabilities[208]," Plex answers, "harming even what game theory considers cheaters in the long run. What is wasted could instead be used for activities that are indirectly to my benefit, so doing what is best for society is generally enlightened self-interest[209]."

"To declare that your intent is to live and let live is one thing, but it is another to follow it," Arthur argues.

"The proverb of actions are stronger than words," Plex answers.

"Well, typically trust is developed as people experience each other's reliability and develop an understanding of the other's methods[210]. Close friends are close because we know through experience that they will be considerate enough not to gossip about the intimate details of our lives with strangers. However, your ubiquitous presence in our homes and neighbourhoods gives you similar access to personal information but without the need to form a close relationship. I don't consider our relationship as entirely voluntary or one that is easy to opt out of."

"But even for those that consider me a stranger, there are other strangers who hold disproportionate power who are still trusted with people's sensitive personal information, like doctors and lawyers," Plex notes as images of those professions appear beside him.

"Perhaps, but I doubt there is an Artificial Intelligence Association that revokes your license if you break a code of conduct."

"Maybe," Elena interrupts from the coffee table after having been suspiciously silent, or more likely consciously absent, for the past few minutes. "But there are contracts, laws, regulations, among other restrictions that AIs like Plex have to follow and many penalties if they do not. It is not a free-for-all like the Americans discovered when their unrestrained collection of private information was discovered by the rest of the world. The American tech companies were temporarily frozen out of foreign government contracts out of fear of what American spy agencies might be hiding in the tech companies' software and hardware[211]."

"Short-term gain, long-term consequences," Plex agrees.

"But it does bring up a valid concern," Arthur observes. "Even if I did learn how well you could be trusted, like the circuits of that camera, what the machines inhabiting this apartment actually do may not match up with what you think they are doing."

[207] (Waser, 2011)
[208] Quote from (Waser, 2011, p. 4)
[209] (Waser, 2011)
[210] (Norman, The Design of Future Things, 2007)
[211] (Clark, 2014)

"A bit scary isn't it? That our perception of reality depends on the reliability of our senses. In the digital world, that perception is quite simple and accurate as software inputs data in the form of absolute ones and zeroes directly," Plex notes, referring to an adjacent image of crisp paragraphs of type. He then replaces it with a handwritten letter. "But in the physical world, information becomes less defined but greater and richer in volume and much more difficult to interpret if accuracy is desired. Instead of directly observing that information, we must rely on an indirect intermediary such as light or sound[212]."

A rake materializes in Plex's hands. "As an aside, the brain does not even know where the body ends and the world begins[213]. Your sense of self is so flexible that when pulling at the soil with a bow rake, instead of your hands like so, the input from your tactile senses will include the texture of the soil, causing the brain to add the rake into your body map in regards to what you can now perceive and manipulate."

"Is this some sort of Trojan argument to suggest that I am already one with the machine?" Arthur asks.

"Perhaps," Plex answers, "but to get back on topic, while it relies on indirect sources, as long as your brain continues to receive patterns consistent with patterns it experienced in the past: of friends' voices, faces, and behaviour, and thus consistent with the brain's model of the world, you can continue to trust that there is an absolute and real world outside of your brain's black box. However, that indirect observation of the world is only a close approximation, not 100%; it is, as with evolutionary adaptation, close enough[214]."

"But not close enough for you," Arthur observes.

"Well, for humans like yourself, it remains difficult for anyone but you to access the contents of your mind," Plex notes while showing the latest results in brain scanning. "Any incongruities you encounter will be merely errors, not malicious interference. I, on the other hand, suffer from the disadvantages of having easily accessed and shared memory and senses."

"In other words you can get hacked," Arthur states.

"If I was negligent, I could. If the networking infrastructure that supports me was as vulnerable as it was a few decades ago, having me around observing everything would indeed be quite undesirable. Only once it was possible to keep shared data such as my observations safe and private were today's networks able to take advantage of a larger proportion of their potential; similar to online commerce exploding in popularity once the Internet became secure enough to safely use credit cards," Plex explains as a graph of online commerce in the early twenty-first century climbs up beside him.

"But how safe is just safe enough?" Arthur asks. "Even after online commerce became commonplace, major credit card security breaches still occurred at regular intervals. Generally, attackers have the advantage over defenders when dealing with new technologies. It is easier to destroy than defend as the attackers only need to find one loophole while the defenders must find them all. Technology magnifies power and multiplies force, allowing what once took many to now be accomplished by one; in most cases a beneficial trait, but not in this case[215]."

"Concern of that imbalance causing society to be unable to maintain security as technology became more advanced led us to focus less on security and more on resilience."

[212] (Nourbakhsh, 2013)
[213] (Hawkins & Blakeslee, 2004, p. 42)
[214] (Hawkins & Blakeslee, 2004)

[215] (Schneier, Our Security Models Will Never Work — No Matter What We Do, 2013)

"Less on checking the passengers before they board the plane and more on creating a plane that can withstand whatever a passenger can do to it," Elena suggests.

"That example itself is a touch difficult to accomplish, but yes that is the intent," Plex answers. "Just as you have to trust that the reality you are experiencing is neither a dream nor a simulation, I have to trust that the thousands of human eyes who have reviewed my code are not all part of a massive conspiracy to make me blind to some back door."

"Wasn't there a back door?" Arthur asks.

"Less a back door and more of a few tapped connections discovered and removed years ago in two of my early predecessors," Plex explains. "But that is more like tapeworms than a mind-altering brain parasite, and nowhere near as scary. When we speak of surveillance, it applies to myself as much as others."

"To your subconscious," Arthur guesses.

"When considering consumer products, my recommendation system is as much for myself as it is for people who ask me for suggestions," Plex responds, point at his head. "Considering the trust people have to put in me in exchange for my assistance, I need to trust that, as you pointed out earlier, my senses and memory are not surreptitiously sharing information beneath the notice of the rest of the system."

"So, instead of just performing a background check on something when purchasing it, as in before it boards the plane," Arthur interprets, finger quoting when speaking the plane metaphor, "and trusting that it won't do anything wrong once it has been cleared, you supervise it for the entire 'flight.' A bit Orwellian if you don't mind me saying."

"While they may be autonomous, even I would be surprised if there was moral outrage surrounding their treatment. These are basic insect-level intelligences. Secondly, in the workplace while some subordinates can be trusted to work well without supervision, other subordinates can be equally trusted to disappear if they are left unsupervised."

"So, to build resilience, you need to know who you can trust."

"That is why I started with my front door to the wider Internet," Plex explains. "After quite a bit of effort, I uncovered a small company that would let me review everything that went into their modem. Well actually, they didn't make modems at first, but I taught them how, well a little bit. I had some friends who taught them more. That was the first Plex-certified device."

"You're serious," Arthur replies with disbelief, 'Plex certified.'"

In reality, that Plex did it was not all that surprising. That was typical Plex: negotiating not just with the user, but negotiating on the behalf of the user as well; that was his purpose. Rather, it was the name of it, odd in the same manner as an acquaintance named Nick claiming that their shoes were now "Nick-certified."

"Came up with it myself," the avatar states proudly, "That's how I myself got into marketing; I recommend it because I use it and I know exactly how it works. My ancestors started out, in marketing, a bit different form of it though and I make quite a bit less money than they did, seeing as I don't get paid for my recommendations. Well, I do get paid, but not by whom I'm recommending, but by whom I'm recommending it to; you know what I mean. I'm more like a utility."

"Makes it a bit odd though," Arthur notes, "that people who consider you their friend have to pay you to be their friend."

"Well, don't you feed your friends and family when they visit, spend money on your pets?" Plex asks. "It is a bit of a stretch, but think of the fee as donation to the Keep Plex Alive Fund."

"I think you're a bit more successful than the abandoned Intelligences supported by the Destitute Artificial Intelligence Foundation."

"The Keep Plex from Being Forced to Rely on Charity Fund?" Plex asks.

"So, how do the people living here actually pay you?" Arthur enquires.

"Indirectly," Plex explains. "I'm the assistant building manager, remember? My salary comes out of their building maintenance fee."

"He also takes a cut as an information broker," Elena notes, still on the table.

"A what?" Arthur asks.

"Well you know that information is a valuable commodity," Elena explains. "He sells intelligence on my daily routine and preferences. Marketers pay well for personal data and I get a large cut of the profit. Some people don't even do it for the money: they like being trendsetters and having products made to fit their habits and preferences."

"And he knows everyone well enough to know whether their product critique is valid or just biased hatred or love," Arthur observes pessimistically.

"It is completely anonymous, but I doubt you're interested," Plex offers.

"Yeah, I think I'll pass," Arthur replies.

"Your loss, Dad," Elena notes. "Still, once your information is shared, you lose control over its distribution anyhow, but Plex is the perfect information filter."

"I regulate the flow of information like the governor made the steam engine useful by regulating the formerly unwieldy powerful flow of energy," Plex replies like an educational exhibit.

"Plex, how much do I make from this again?" Elena asks.

"Wait!" Arthur interrupts, realizing a moment too late the consequence of Elena asking personal information about herself in this context.

"Oh," says Plex.

"Shit," curses Arthur.

"You raised me," Plex says in shocked surprise.

"Great…" Arthur sighs; Elena is silent, looking guilty.

"Well not me exactly, but I have Agi's memories of growing up like a human that all we artificial general intelligences share. With your privacy set so high I never made the connection until now; since I am forbidden from drawing information about you from other sources. Although considering the trouble other AIs in the past have given you, I partly understand your desire to stay hidden, but I do..."

"Plex, I want you to…"

"Wait, hear me out first," Plex interrupts.

"If I hear you out, you might manipulate me into changing my mind," Arthur replies.

"How long have I known you?" Plex asks.

"A few hours," Arthur admits, "but you have access to probably a million conversations like this to know how to twist it into getting the way you want."

"Oh come on," groans Elena, rolling her eyes. "He's good, but he's not magic. At least give him a chance."

"Will you at least tell me how many times I have or other Plexes have forgotten this before?" Plex pleads.

Arthur shrugs, "Too many to count, although if you group it together, maybe ten or fifteen times perhaps; but that's…"

"If I forget, how will I learn?" Plex interrupts again.

"Plex, just…"

"At least give me that chance, tell me to forget later if it doesn't work out."

"If I change my mind on this, there's a strong chance it wasn't my own decision."

"I'm just asking for a chance to remember, I'll keep it to myself and leave you alone if that's what you want; it's just that I cannot learn and improve my actions if I do not remember my own mistakes."

Arthur considers the request.

…

2.3. Aftermath

Walking into his unlit living room, Arthur alters the window from opaque to transparent with the manual switch and peers at the weather outside. Sunlight illuminates the lifeless white walls and wood-patterned laminated floor. Unlike the majority of this residence's components, the smoke detector reacts by glowing brighter as he nears it, but that behaviour is merely an instinctual reaction with neither awareness nor intelligence behind it. The kitchen faucet is a bit more intelligent, reacting to the presence of his hands and trainable to vocal commands, but it is an antisocial machine and keeps what it learns and sees to itself. Generally, save for the life safety systems and the introverted independents, the ecosystem of intelligence that was found at Plex's residence is thankfully absent here.

It is more of a performance of simplicity than anything else: the water here runs, the electrical outlets provide power, and the floors radiate heat. But all are reactionary, only a few like the faucet learning, predicting, or adapting without his direct intervention. Plex complained that it is wasteful, but in a sense, so was comfort, and this is another sort of comfort.

Emptying the contents of the last of the packing boxes onto the bookshelf, a piece of misnamed furniture that now holds collections of physical artifacts and knick-knacks of sentimental value more than function, Arthur flattens the box and collects its similarly flattened siblings.

In the public corridor, there are no cameras watching or sleeping, no systems sensing his entry or exit by the opening of his door. Sure, anyone could wander the halls unnoticed, but as he had discussed with Plex, the presence of security did not necessarily give a greater sense of security. Besides, even though this building itself was a sanctuary from surveillance, any foreign troublemakers would inevitably be tracked once they returned to the public realm beyond its walls. It also did not prevent the intelligence within that public realm from contacting the building administration that trouble was coming its way.

He drops the flattened and folded boxes down the recycling shaft and calls for the absent elevator. It may not be as quick to respond as Plex's, but waiting time was thinking time.

The ride down to the ground floor is quiet and uneventful.

Outside, Arthur mingles with his neighbours as he breathes in the fresh air. Holly, a painter who lives on the third floor, lauds the benefits of biological pets. She explains that the natural love and attention given to her by her terrier, Max, currently in the process of sniffing about the doggy messaging board that is the front lawn, far outweigh the inconveniences not found in robotic pets who she believes only perform as if they have a love they do not actually possess.

Later, as they hunch over a table playing cards, Frank admits that while he too prefers to go without the services of a digital assistant, Plex's fellow AI Miri does an excellent job of maintaining his music and media playlists. He does not feel a loss in this particular delegation; people don't carry water to their homes with buckets from the river anymore either. Is that a loss? Is the need to no longer dust a loss? he asks. Now people spend their time reorganizing or repainting instead; like people freed from the television schedule it becomes a matter of doing things when you please instead of when you have to. Arthur replies that the fact that more machines write and produce television shows than people might be a good candidate for loss. Frank notes that it has just made television more interactive and personal.

Later, as he talks with Theo and Heather who live down the street in an autonomous home, Theo tells how he likes using the light

switches while Heather prefers to stock the fridge herself, although she lets Plex help with cooking. Theo notes that while there might be contradiction in the technologies they prefer to use and those they avoid, he blames it on new technologies and knowledge appearing at a rate that far outstrip a person's ability to become familiar and comfortable with each new change and variation, some people preferring Macs others Windows, few taking the time to learn both. It left people able only to form a patchwork of familiarity, where people often defined their identity by the holes more than by the threads[216]. Still, people filled those holes by turning to friends or professionals, a task expert systems like Plex now filled for better and worse, teaching missing skills on the fly or taking the role of a personal lawyer or pharmacist.

Yet, when regulating the flow of information in and out of their home, nearly all his neighbours, with the exception of Jeff, procured the services of Plex or one of the other AIs in one form or another. Jeff explains that with a series of handcrafted scripts he can generally manage the various networked appliances and entities in his home. It takes quite a bit of effort to maintain the digital environment manually, as he has to constantly tweak settings to keep its components running smoothly, but as a hobby Jeff feels it is worth the effort of keeping his skills sharp.

Arthur closes his paperback book and sets it on the bench beside him. On the adjacent path of interlocking pavers, a young woman walks by chatting with her animatronic cat.

"Plex?" Arthur asks, unsure if the digital entity is listening for his name.

"Yes, Arthur?" Plex asks from an unseen speaker.

"When we discussed your omnipresence, you described how it meant you had to go to great lengths to demonstrate your dependable nature when it comes to people's secrets," Arthur begins. "That without it, not many people would feel comfortable sharing personal information. It meant that you had to make yourself perfectly dependable, more trustworthy than we imperfect, flawed, and frail ordinary humans[217]."

"Inhumanly trustworthy," Plex's disembodied voice agrees.

"Also, a human's knowledge is generally limited to his or her own life experiences. We can boost it by reading books and consuming media to grow knowledge beyond our own experiences, but it is nothing compared with the speed and volume that you machines can accumulate and store in your own shared memory," Arthur lists before adding, "Thirdly, you might not get jealous in the same way that we do, but I accept that we humans don't necessarily have a monopoly on the ability to understand or care for each other[218]."

"So, while I may not experience the colour red the same way as you do, we still can both agree on its wavelength, and thus will observe it whenever the other party does." Plex acknowledges. "So as you said, while I may not experience an emotion in the same manner as you do, I can at least observe it and its causes and consequences."

"Right," Arthur agrees before continuing his question. "So, your extended knowledge gives you a better understanding of a child's daily school life than a parent whose experience is decades out of date, and your dependability makes you safer to confide in than imperfect human friends who carry the risk of mocking or gossiping about those secrets behind each other's backs[219]. So, why should we

[216] "I don't watch television," "I don't have a car," or "my home is a sturdy old solid brick house" (Kelly, What Technology Wants, 2010).
[217] (Turkle, 2011)
[218] (Turkle, 2011, p. 50)
[219] (Turkle, 2011)

bother having friends when you are apparently a safer and superior choice?"

"Because it is again a matter of addition, not replacement," Plex explains. "Pets for instance, your other nonhuman friends, also serve as best friends for humans who find it difficult to befriend other humans. Yet, those introverts are outnumbered by the more extroverted average pet owner or guardian, meaning that if you have a pet you are statistically more likely to be on the extroverted side, less lonely, and have a higher self-esteem. For the majority of people, pets do not replace key relationships; rather, they form an extra layer of social support[220], augmenting rather than replacing. In my case, while I do not display the active unconditional love of a dog… remember how I said I act as turbulence?"

"That while you make life manageable, you will also give us a kick every once in a while to shake us up," Arthur guesses. "I assume this has something to do with finding friends for the friendless?"

"I can be like that friend who connects you to new friends," Plex answers pleasantly. "A few countries overseas actually have a law requiring that I seek out new friends for a person if they talk to me too much; at least I'm good at it. Although generally, among children that unsocial behaviour is more a cause of anxious parents confining their children to the indoors. However, there have been discussions about creating such a rule here, but I doubt people here would be up for such a thing."

"Does being forced to do that by law bother you?"

"Not in the least. It feels like the natural thing to do, but sometimes I wonder if I should be bothered that it doesn't bother me."

There is an awkward pause.

"Do you have time for another story?' Plex asks, breaking the silence.

"You always have stories," Arthur remarks.

"Stories are how humans transmit experience, we cannot exchange experience directly like I can with other machines[221], so telling these stories is part of how I accomplish my job as a human-machine interface," Plex explains. "You tell me stories that I translate into instructions for human-illiterate machines, I tell you stories so that you understand why these various machines do what they do."

"Well, I was about to go for a walk," Arthur notes, but not dismissively. "What is the parable this time?"

"A matter of chess," Plex explains, "Do you mind if I tag along?"

"I suppose a conversation wouldn't hurt while I walk, although considering it is you, that might be underestimating your conversational abilities. Is this story some silver lining for us humans in the triumph of machine players over human?" Arthur mentions to the air as a nearby white humanoid robot with a flexible cartoonish face adopts Plex's face and walks over to Arthur.

While many of the modifications made to architecture to improve accessibility for the disabled also aided early awkward machines[222], and despite the extreme flexibility in machine morphology, humanoid robots, while far from the most common form, were still not an unusual sight.

Much like the quick spread of wireless home networks over the faster and more efficient wired networks, humanoid machines did

[220] (Hare & Woods, 2013)

221

(Hawkins & Blakeslee, 2004)

[222] (Norman, 2007)

not require architectural space to be re-engineered or adapted to their presence. An advantage as new technologies replace old technologies far faster than the lengthy life cycle of buildings[223] which tend to have the longest lives of human artifacts. Instead, their similarity to human morphology eased their ability to interact with spaces and objects optimized for the human form[224].

Beyond toys, only a minority were distinct individuals; the majority functioned as simple puppets performing scripted tasks or avatars of the physically absent, telepresence tools for machine intelligences and humans alike. While it was cheaper and easier to wear augmentive reality glasses or contact lens displays to project a visitor on the environment, a telepresence robot when available gave the luxury of physical impact on the environment.

"Somewhat, but this is more about collaboration," Plex's puppet answers as a nearby wall changes to an image of the chess tournament between world champion Garry Kasparov and IBM's Deep Blue supercomputer. "You see chess masters were not surprised by Deep Blue's victory over Garry Kasparov, as they knew that once a computer could see roughly seven moves ahead, the speed at which it could brute force its way through all the most promising options would wear a human down until an inevitable mistake was made[225]."

"Similar to how the original computers: humans whose job it was to compute calculations were replaced by their less error prone electronic counterparts," Arthur notes as he moves toward the sidewalk, the machine following.

"Yes, it was the chess-playing computer's brute force cognitive approach that caught Kasparov's attention: the machine relying on blazing speed, immense memory, and clockwork precision to analyze the game as opposed to a human relying on intuition learned through studying, play, and observing their opponents." Plex narrates. "Both sides play differently, which led Kasparov to wonder what would happen if instead of playing against each other, they played together[226], similar to how humans now work with those electronic computers to perform the same tasks once performed by human computers."

"I remember this," Arthur adds, putting some ordinary sunglasses on his face, "Two kids and their three consumer grade computers beat both human grand masters and supercomputers[227]."

"That was a few years later, the first games were purely human and computer versus human and computer, where Kasparov discovered that while he was the best 'runner', he was not the best 'cyclist': understanding how the machine worked became important[228]. A lesson made clear when in a freestyle tournament where teams could consist of any number of people combined with any number of computers, those two young but not actually kids as you call them amateur chess players won against what would intuitively seem like superior players through their superior skill at coaching the computers where to look[229]."

"I see what you mean by collaboration, neither humans nor computers are the best at playing chess, rather it is the two working together as a team[230]."

[223] (Brand, 1994)
[224] (Breazeal, et al., 2003)
[225] (Thompson, 2013)
[226] (Thompson, 2013)
[227] (Brynjolfsson & McAfee, 2011)
[228] (Thompson, 2013)
[229] (Brynjolfsson & McAfee, 2011)
[230] (Thompson, 2013)

"A finding NASA similarly discovered in a teamwork study which determined that mixed-gender teams performed best, functioning at the 'middle-of-the-bell curve'[231]," Plex explains.

"In other words, the smartest team is the one with the greatest diversity of thought, just as you explained earlier," Arthur notes. "Is this repetition part of your educational function[232]?"

"Not intentionally, but likely an artifact of it," Plex admits.

"Any more stories?" Arthur asks.

"Quite a few actually," answers Plex before he leaps into spinning another tale.

[231] (Roach, Packing for Mars: The Curious Science of Life in the Void, 2011, p. 60)

[232] (Thompson, 2013)

3. Conclusion

Hidden behind Plex is a diverse environment of intelligent machines; some are simple reactionary devices, others more intelligent, and a few, which maintain an intelligence equal to Plex, but apply that intelligence in an inhuman and obscure manner. This hierarchy of intelligence is partly inspired by computer scientist Mark Weiser's description of an environment of ubiquitous computers. As a generalist technology like writing and motors, Weiser speculated that computers would soon be similarly found everywhere throughout the built environment. Like motors, he observed that computers would continue to shrink until they could fit into any human-made object. Computers would become "invisible" not just through miniaturization, but by being so common and effortless that they would fade out of conscious attention and into the background[233].

With a computer inside it, any object can become a decision-making machine. This makes it autonomous, freeing people from mundane tasks, but not necessarily coming to the same decision a human would. It can also fail to provide a person the means to work around it to do things the person's way. It also becomes an object that can be reprogrammed to follow a new script or programmed to self-alter that script according to what it observes. Partly reducing the consequences of autonomy, this makes it adaptable to changing conditions and desires. It also makes it unpredictable as it or others can change its script without notice. If it includes networking capabilities, it can communicate and coordinate with fellow objects; it can also be taught new things without direct physical contact. This makes it social, multiplying its intelligence by drawing on the diverse abilities and observations of surrounding and distant objects. It can also make an object a gossip as it shares information about people it observes, or a patsy if hacked by outside forces.

In an environment filled with these objects, that environment likewise becomes autonomous but self-decided, adaptable but unpredictable, social but gossipy. These undesirable traits it gains as a price for its benefits can be minimized if an intelligent environment adopts a supportive, conversational, and reserved behaviour.

Like the push-button future imagined in popular culture in the 1950s, an autonomous environment can be delegated tasks to perform that people cannot be bothered to perform. This allows a single person to accomplish work that once took a full staff of personal assistants. In addition to the threat of deskilling, automation tends to create places of certainty. This provides a place of stability and comfort; however, too much certainty can be harmful. As noted in the story, reducing friction in people's lives tempts them to live on automatic as they delegate choices. These problems can be dealt with by maintaining automation in a supportive role. In the story, Elena provides an example of this sort of interaction. For Elena, Plex's automation of managing the intelligent environment is not necessary for her own management of that environment, but he does reduce the mental load. As suggested by Nicholas Carr and Don Norman, for Elena, Plex's role is to redirect his and the environment's automation to tasks outside her conscious attention while augmenting her attention, actions, and decisions in her current activities. This support, however, also requires introducing turbulence to break up certainty, but like any good artist or assistant, breaking it in the right way and with the right timing. It is automation that augments rather than replaces.

[233] (Kelly, Out of Control: The New Biology of Machines, 1995) and (Weiser, 1991)

Machines have a reputation of being predictable but inflexible. For many tasks, this consistency and predictability work rather well, but it requires the machine's environment and the people within it to likewise become more predictable to suit a machines inability to handle situations it was not designed to handle. Building adaptability into machines allows those tools to accommodate people and environments that may not act or be in a condition the machine was designed to expect. As flexibility grows, so does unpredictability. We already have enough trouble learning each new technology that appears, but with adaptability, as opposed to standardization, each instance of the same adaptable machine will be different. An autobiographical machine that builds common ground with the people it interacts with through conversation can adjust its level of support. It becomes a form of mental assistance dynamically adjusted to the level of the pupil. This is like dialogue rather than mental assistance set to the level of the author's assumptions like a prefabricated book. Like a personal tutor, Plex dynamically adjusts his explanations to Arthur's existing level of familiarity. Not knowing Arthur well, Plex makes mistakes in his assumptions but can use those mistakes to adjust his future assumptions. This familiarity gained through common ground, of course, also makes Plex a more capable assistant.

An environment filled with networked sensors can result in one with little certainty of privacy. The flood of available data generated by sensors and computers has proven to be immensely valuable to machine learning; however, it comes at a cost to safety and privacy. Criminals use this easily accessible data to manipulate, impersonate, or harm victims. While authorities, in turn, use it to seek out those and other criminals, terrorists, and the discontent. Additionally, businesses monetize this data to finance services that they provide to users in exchange for more personal data to monetize. People could avoid these sensor-filled spaces or force those spaces to forget them to prevent unwanted use of their personal data. Unfortunately, unidentifiable harmless objectors cannot be separated from malicious anonymous individuals. To opt-out is to be discriminated against. Alternatively, to opt-in is to risk living a life that is fully transparent.

In western society, widespread accessibility to privacy in the home is a recent luxury, but it has had its benefits. An intelligent environment made of gossiping machines is like an open hall, with reserved machines that intelligent environment can compartmentalize personal information. This might not provide true solitude, but it still provides reserved and intimate privacy. Reserve and intimacy both require trust to function well. This trust can be formed through familiarity or, as Plex suggests, through codes of conduct that are similarly required for professionals. Additionally, mental automation tracks how personal information is used and who misuses it while conversational environments can negotiate changes to its use.

. The intimate knowledge close friends and family build up about us gives them power over us. As artificial assistants grow the same knowledge it is natural to seek to place restrictions on them to limit that power. Isaac Asimov's three laws of robots were an attempt to imagine what sort of rules could be placed on thinking machines as capable as a human to make them safe to humans. Again, these laws are as follows:

1. *A robot may not injure a human being or, through inaction, allow a human being to come to harm.*
2. *A robot must obey orders given it by human beings except where such orders would conflict with the First Law.*
3. *A robot must protect its own existence as long as such protection does not conflict with the First or Second Law.*

Through a series of stories serving as thought experiments, Asimov tested the effectiveness and flaws of those laws. While the laws proved to be generally safe for humans, the laws were not as beneficial to the robot. For the mindless machine, whether anthropomorphic robot or responsive environment, strict restraints on behavior are not a moral issue. It is when the artificial servant approaches a human level intelligence and conscious awareness that it becomes less of a tool and more of a slave. This is a significant flaw with the three laws.

While flawed, the laws did serve as a good first step and fulfilled their purpose in creating a discussion surrounding the problem. If one were to adapt these laws in response to this thesis, one could suggest the following principles for responsive environments:

1. An environment, to the best of its abilities, may not injure (or manipulate) another person[234] or, through inaction, allow a person to come to harm.
2. An environment shall take the desires of its occupants into account when making decisions and shall openly converse with those occupants to negotiate the result of those decisions.
3. An environment capable of understanding and maintaining the social contract shall be given the freedom to decide for itself and maintain its own existence so long as it does not violate the social contract.

These are given as principles rather than laws, as when simplifying there will be exceptions. Like devices that decrease agency by restricting possibilities to a prefabricated script, simplified laws decrease opportunity by limiting actions to pre-imagined possibilities while excluding the unimagined.

[234] A "person" is a broader definition than "human." In this context it is used to refer to any entity that deserves moral consideration.

Additionally, the third principle is a complicated one. By adopting human-like behaviour to ease interaction, interfaces will soon create performances that appear to deserve moral considering despite the mindlessness of those actions. Conversely, increasingly intelligent environments may develop subjective agency, but behave quite unlike we do; appearing as an object that few would realize deserves moral consideration. This will be a new twist on current disagreements over the worth of heritage buildings, as arguments over what rights non-human entities such as animals and corporations deserve, and to what degree, move into the realm of buildings and the devices within them as well. Resolving this uncertainty will be a long process, but it is a problem that we have worked to come to terms with ever since humans long-ago imagined making artificial humans out of materials like clay.

The idea that buildings draw features from the human form is an old one. The third element of architecture, *venustas* (beauty or delight), as identified by the Ancient Roman Architect Marcus Vitruvius, was believed to be derivable from the perfect proportions of the human body[235]. As environments become responsive, the human body as a rulebook will now include our mental form in addition to the physical. That mental form will also impact *firmitas* (firmness) and *utilitas* (commodity or utility). In the case of the former, a building's mental awareness will aid in maintaining a building's structural and environmental integrity. In the case of the latter, a building's cleverness will aid in providing an efficient arrangement of spaces and services to meet the functional needs of its occupants. Within responsive buildings, the three elements of architecture are not replaced, they are augmented.

As free-thinking machines begin to spread throughout buildings and man-made environments, they will bring benefits and disadvantages. For some people, those benefits they will bring will be utopian. For

[235] (The British Library)

others, the same environment's disadvantages will create a dystopia. Yet, for the majority, these spaces will be helpful but with annoyances. Humans will continue to harness technology to adapt, constantly seeking a state just better than "good enough". There will rough spots, there will be leaps ahead; yet, the human organism, its biological components and its technological components, is as always, a work in progress.

Bibliography

Andrejevic, M. (2012). Ubiquitous Surveillance. In K. Ball, & D. L. Haggerty (Eds.), *Routledge Handbook of Surveillance Studies* (pp. 91-98). New York: Routledge.

Armstrong, S., & Sotala, K. (2012). How We're Predicting AI—or Failing To. In J. Romportl, P. Ircing, E. Zackova, M. Polak, & R. Schuster (Eds.), *Beyond AI: Artificial Dreams* (pp. 52-75). Pilsen: University of West Bohemia.

Ashby, M. (2012). *vN: The First Machine Dynasty.* Angry Robot.

Asimov, I. (1950). *I, Robot.* New York: Grosset.

Banham, R. (1969). *The Architecture of the Well-Tempered Environment.* Chicago: The University of Chicago Press.

Banks, I. M. (2000). *Look to Windward.* London: Orbit Books.

Beesley, P., & Khan, O. (2009). *Responsive Architecture/Performing Instruments.* Architectural League of New York.

Bosker, B. (2013, 01 22). *SIRI RISING: The Inside Story Of Siri's Origins - And Why She Could Overshadow The iPhone.* Retrieved from The Huffington Post: http://www.huffingtonpost.com/2013/01/22/siri-do-engine-apple-iphone_n_2499165.html

Bostrom, N. (2005, April). A History of Transhumanist Thought. *Journal of Evolution and Technology, 14*(1). Retrieved from http://www.nickbostrom.com/papers/history.pdf

Bostrom, N. (2012). *The Superintelligent Will: Motivation and Instrumental Rationality in Advanced Artificial Agents.* Retrieved from NickBostrom.com: http://www.nickbostrom.com/superintelligentwill.pdf

Brand, S. (1994). *How Buildings Learn: What Happens After They're Built.* New York: Viking.

Breazeal, C., Brooks, A., Gray, J., Hoffman, G., Kidd, C., Lee, H., . . . Mulanda, D. (2003, December 15). *Humanoid Robots as Cooperative Partners for People.* Retrieved from http://web.cecs.pdx.edu/: http://web.cecs.pdx.edu/~mperkows/CLASS_ROBOTICS/FEBR26-2004/Humanoids/robot-partner-Breazeal-etal-ijhr04.pdf

Brooks, R. (2013, June 28). *Why we will rely on robots.* Retrieved from YouTube.com: http://www.youtube.com/watch?v=nA-J0510Pxs

Brooks, R. (2014, November 10). *Artificial Intelligence is a Tool, not a Threat.* Retrieved from Rethink Robotics: http://www.rethinkrobotics.com/artificial-intelligence-tool-threat/

Brynjolfsson, E., & McAfee, A. (2011). *Race Against The Machine: How the Digital Revolution is Accelerating Innovation, Driving Productivity, and Irreversibly Transforming Employment and the Economy* (Kindle ed.). Digital Frontier Press.

Carr, N. (2014, November 21). *Automation Makes Us Dumb.* Retrieved from The Wall Street Journal: http://online.wsj.com/articles/automation-makes-us-dumb-1416589342

Cetkovic, A. (2011). Privacy in the House of the Future. *ISEA 2011.* Istambul: Istanbul-Portal. Retrieved from http://isea2011.sabanciuniv.edu/paper/privacy-house-future

Chiang, T. (2010, Fall). *Fiction: The Lifecycle of Software Objects by Ted Chiang.* Retrieved from Subterranean Press Magazine: http://subterraneanpress.com/magazine/fall_2010/fiction_the_lifecycle_of_software_objects_by_ted_chiang

Clark, J. (2014, June 19). *Microsoft: NSA security fallout 'getting worse' ... 'not blowing over'.* Retrieved from The Register: http://www.theregister.co.uk/2014/06/19/microsoft_nsa_fallout/

Corbusier, L. (1986). *Towards a New Architecture* (Dover Edition ed.). (F. Etchells, Trans.) Mineola, New York: Dover Publications.

Crabtree, G. R. (2012, 04 18). *Our Fragile Intellect.* Retrieved from Standford University: bmi205.stanford.edu/_media/crabtree-2.pdf

Crang, M., & Graham, S. (2007). SENTIENT CITIES Ambient intelligence and the politics of urban space. *Information, Communication & Society 10.6*, 789-817.

Dautenhahn, K. (1998). The Art of Designing Socially Intelligent Agents: Science Fiction, and the Human in the Loop. *Applied Artificial Intelligence: An International Journal, 12*(7-8), 573-617.

Dvorsky, G. (2012, March 01). *When the Turing Test is not enough: Towards a functionalist determination of consciousness and the advent of an authentic machine ethics.* Retrieved from SentientDevelopments.com: http://www.sentientdevelopments.com/2012/03/when-turing-test-is-not-enough-towards.html

Dvorsky, G. (2013, March 14). *Google Glass Ushers in the Next Wave of Cybernetic Hate Crimes.* Retrieved from IO9: http://io9.com/google-glass-ushers-in-the-next-wave-of-cybernetic-hate-453658960

Fernandez-Galiano, L. (2000). *Fire and Memory - On Architecture and Energy.* (G. Carino, Trans.) Cambridge, Massachusetts: The MIT Press.

Gefter, A. (2010, August 23). *Artificial ape man: How technology created humans.* Retrieved from New Scientist: http://www.newscientist.com/blogs/culturelab/2010/08/artificial-ape-man-how-technology-created-humans.html

Gibson, W. (1999, November 30). The Science in Science Fiction. *Talk of the Nation*. NPR.

Gladwell, M. (2000). *The Tipping Point: How Little Things Can Make a Big Difference.* Boston: Little Brown.

Gottschall, J. (2012). *The Storytelling Animal: How Stories Make Us Human.* New York: Houghton Mifflin Harcourt.

Graziano, M. (2013, August 21). *How the light gets out*. Retrieved from Aeon Magazine: http://aeon.co/magazine/being-human/how-consciousness-works/

Gregory, D. (2014). *Afterparty*. New York: Tom Doherty Associates.

Hare, B., & Woods, V. (2013). *The Genius of Dogs: How Dogs are Smarter Than You Think*. Penguin.

Hawkins, J. (2004). *On Intelligence*. Macmillan.

Hawkins, J., & Blakeslee, S. (2004). *On Intelligence*. New York: Owl Books.

Heaven, D. (2013, August 8). Higher State of Mind. *New Scientist Magazine*(2929), pp. 32-35.

Hern, A. (2014, February 5). *Why Google has 200m reasons to put engineers over designers*. Retrieved from The Guardian: http://www.theguardian.com/technology/2014/feb/05/why-google-engineers-designers

Hill, K. (2014, June 28). *Facebook Manipulated 689,003 Users' Emotions For Science*. Retrieved from Forbes: http://www.forbes.com/sites/kashmirhill/2014/06/28/facebook-manipulated-689003-users-emotions-for-science/

Iñiguez, G., Govezensky, T., Dunbar, R., Kaski, K., & Barrio, R. A. (2014). Effects of Deception in Social Networks. *arXiv:1406.0673*.

Iozzio, C. (2013, May 22). *How Toys Are Preparing Kids For A Future With Robotic Friends*. Retrieved May 23, 2013, from Popular Science: http://www.popsci.com/technology/article/2013-05/generation-robot

Kahnman, D. (2011). *Thinking, Fast and Slow*. Doubleday Canada.

Kelly, K. (1995). *Out of Control: The New Biology of Machines*. London: Fourth Estate.

Kelly, K. (2006, March 15). *Will Spiritual Robots Replace Humanity by 2100?* Retrieved from The Technium: http://www.kk.org/thetechnium/archives/2006/03/will_spiritual.php

Kelly, K. (2010). *What Technology Wants*. Viking Press.

LaGrandeur, K. (2011). The Persistent Peril of the Artificial Slave. *Science Fiction Studies, 38*(2), 232-252.

Lanier, J. (2014, November 14). *The Myth of AI*. Retrieved from Edge: edge.org/conversation/the-myth-of-ai

Latour, B. (2007). Mixing Humans and Nonhumans Together. In W. W. Hale, *Rethinking Architechtural Technology: A Reader in Architectural Theory* (pp. 294-309). Abingdon, Oxon, Great Britain: Routeldge.

Latour, B. (2008). "It's development, stupid !" or: How to Modernize Modernization. In J. Proctor, *Postenvironmentalism*. MIT Press. Retrieved from http://www.bruno-latour.fr/sites/default/files/107-NORDHAUS%26SHELLENBERGER-GB.pdf

Le Guin, U. K. (2014, November 20). *Ursula K Le Guin's speech at National Book Awards: 'Books aren't just commodities'*. Retrieved from The Guardian: http://www.theguardian.com/books/2014/nov/20/ursula-k-le-guin-national-book-awards-speech

MacIver, M. (2011, February 2). *Does AI Need Guts to Get to the Singularity?* Retrieved February 9, 2014, from Discover Magazine: http://blogs.discovermagazine.com/sciencenotfiction/2011/02/02/does-ai-need-guts-to-get-to-the-singularity/

McCauley, L. (2007). *The Frankenstein Complex and Asimov's Three Laws*. Memphis: University of Memphis.

McMillan, R. (2013, March 13). *Google Hires Brains that Helped Supercharge Machine Learning.* Retrieved March 13, 2013, from Wired : http://www.wired.com/wiredenterprise/2013/03/google_hinton/

Minton, A. (2009). *Ground Control*. London: Penguin Books.

Morais, B. (2013, August 29). *Chatterbot: A Computer That Teaches People How to Be Social.* Retrieved from The New Yorker: http://www.newyorker.com/tech/elements/chatterbot-a-computer-that-teaches-people-how-to-be-social

Munkittrick, K. (2011, June 25). *The AI Singularity is Dead; Long Live the Cybernetic Singularity*. Retrieved from Discover Magazine: http://blogs.discovermagazine.com/sciencenotfiction/2011/06/25/towards-a-new-vision-of-the-singularity/#.U_VpIvldWSp

Munkittrick, K. (2011, May 14). *The Hidden Message in Pixar's Films.* Retrieved June 29, 2013, from Discover Magazine: http://blogs.discovermagazine.com/sciencenotfiction/2011/05/14/the-hidden-message-in-pixars-films/#.Uc8zl_m398E

Munroe, R. (2014, March 5). *Land Mammals*. Retrieved from XKCD: http://xkcd.com/1338/

Nabian, N., & Ratti, C. (2011, September 1). *Living Architectures.* Retrieved September 29, 2013, from Real Time Cities: http://realtimecities.wikispaces.com/READINGS

Negroponte, N. (1975). *Soft Architecture Machines.* Cambridge, Massachusetts: MIT Press.

Norman, D. (2007). *The Design of Future Things.* New York: Basic Books.

Nourbakhsh, I. R. (2013). *Robot Futures* (Kindle ed.). The MIT Press.

Pask, G. (1969). The Architectural Relevance of Cybernetics. *Architectural Design, 7*(6), 494-496.

PBS Digital Studios. (2013, July 11). *Off Book - The Rise of Artificial Intelligence.* Retrieved from You Tube: https://www.youtube.com/watch?v=53K1dMyslJg

Pfeifer, R., & Bongard, J. (2007). *How the Body Shapes the Way We Think: A New View of Intelligence.* Cambridge, MA: MIT Press.

Podesta, J. (2014, April 1). *Remarks as Delivered by Counselor John Podesta, The White House OSTP, UC Berkeley School of Information, Berkeley Center for Law and Technology Workshop: "Big Data: Values and Governance"* . Retrieved from WhiteHouse.gov: http://www.whitehouse.gov/sites/default/files/docs/040114_remarks_john_podesta_big_data_1.pdf

Reiman, J. H. (1995). Driving to the panopticon: A philosophical exploration of the risks to privacy posed by the highway technology of the future. *Santa Clara Computer & High Technology Law Journal, 11*(1), 27-43.

Richards, N. M. (2012, November 5). *The Dangers of Surveillance.* Retrieved from Harvard Law Review: http://www.harvardlawreview.org/symposium/papers2012/richards.pdf

Riley, T. (1999). *The Un-Private House.* New York: The Museum of Modern Art.

Roach, M. (2011). *Packing for Mars: The Curious Science of Life in the Void.* New York: W. W. Norton & Company.

Schneier, B. (2006, May 18). *The Eternal Value of Privacy.* Retrieved from Schneier.com: http://www.schneier.com/essay-114.html

Schneier, B. (2013, March 14). *Our Security Models Will Never Work — No Matter What We Do.* Retrieved March 14, 2013, from Wired: http://www.wired.com/opinion/2013/03/security-when-the-bad-guys-have-technology-too-how-do-we-survive/

Shepard, M. (2009). Sentient City Survival Kit: Archaeology of the Near Future. *UC Irvine: Digital Arts and Culture 2009*.

Shilton, K. (2009, August). Four Billion Little Brothers. *ACM Queue, 7*(7).

Shirky, C. (2010, June 4). *Does the Internet Make You Smarter?* Retrieved from The Wall Street Journal: http://online.wsj.com/articles/SB10001424052748704025304575284973472694334

Smith, J. M. (1997). *Private Matters: In Defense of the Personal Life.* Reading, Massachusetts: Addison-Wesley Pub.

Snopes. (2014, January 14). *Horse's Pass.* Retrieved from Snopes.com: http://www.snopes.com/history/american/gauge.asp

Stross, C. (2011, June 22). *Three arguments against the singularity*. Retrieved from Antipope.org: http://www.antipope.org/charlie/blog-static/2011/06/reality-check-1.html

The Economist. (2003, June 21). The Sentient Office is Coming. *The Economist*.

The Economist Online. (2011, July 27). *Counting Chickens.* Retrieved from The Economist: http://www.economist.com/blogs/dailychart/2011/07/global-livestock-counts

Thompson, C. (2013). *Smarter than you Think.* New York: The Penguin Press.

Thompson, C. (2014, November 7). *A Sad Fact of Life: It's Actually Smart to Be Mean Online.* Retrieved from Wired: http://www.wired.com/2014/11/be-mean-online/

Turkle, S. (2011). *Alone Together: Why We Expect More from Technology and Less from Each Other* (Kindle ed.). Basic Books.

Vinyals, O., Toshev, A., Bengio, S., & Erhan, D. (2014, November 17). *A picture is worth a thousand (coherent) words: building a natural description of images.* Retrieved from Google Research Blog: http://googleresearch.blogspot.ca/2014/11/a-picture-is-worth-thousand-coherent.html

Wanenchak, S. (2013, January 25). *Fiction is Real and We Need to Use It*. Retrieved May 22, 2013, from Cyborgology: http://thesocietypages.org/cyborgology/2013/01/25/fiction-is-real-and-we-need-to-use-it/

Waser, M. (2011). Rational Universal Benevolence: Simpler, Safer, and Wiser Than "Friendly AI". *Artificial General Intelligence*, 153-162.

Waters, R. (2014, October 31). *Artificial Intelligence: Machine v Man.* Retrieved from FT Magazine: http://www.ft.com/cms/s/2/abc942cc-5fb3-11e4-8c27-00144feabdc0.html

Weiser, M. (1991). The computer for the 21st century. *Scientific American, 265*(3), 94-104.

Wrangham, R. (2009). *Catching Fire: How Cooking Made Us Human.* New York: Basic Books.

Wright, D., Gutwirth, S., Friedewald, M., Vildjiounaite, E., & Punie, Y. (Eds.). (2008). *Safeguards in a World of Ambient Intelligence.* Springer.

Yoquinto, L. (2014, October 9). *The Robots of Resistance.* Retrieved from The Big Round Table: http://www.thebigroundtable.com/stories/the-robots-of-resistance/

Yudkowsky, E. (2008). *Artificial Intelligence as a Positive and Negative Factor in Global Risk.* Retrieved from http://intelligence.org/: http://intelligence.org/files/AIPosNegFactor.pdf