

# Chapter 2

## Three Case Studies

Donald R. Searing and Elizabeth A.M. Searing

### 2.1 Introduction

Case studies can help one get a better understanding of an unfamiliar topic and the extent of its related ethical and conceptual issues and concerns. The area of emerging technologies known as pervasive information and communication technologies (PICT) includes a number of seemingly disparate issues and technologies. Analyzing a sample of real-world examples of these technologies, how they behave, and where they are headed can bring substantial clarity to this broad and novel field.

To properly grasp the effect of the emerging PICT systems in their full breadth, it is first important to define the types of technologies that fall under this broad umbrella. The key word in the acronym, in our opinion, is *pervasive*. *Pervasive* indicates that we are looking at a phenomenon that completely suffuses a world and all aspects of the lives of the inhabitants in that world. In this book, the pervasive phenomenon being examined is that of information and communications technology; as you read this, pervasive technologies are creeping into every aspect of our lives, such as mobile telephones, computers, the Internet, robotics, and automation. These technologies have had a significant effect on the way we live our lives already, and trends indicate that we are moving in the direction of even greater penetration and integration of these technologies into our everyday experience. The constant cost reduction and capacity growth in the basic units of computation represented by Moore's Law means that more and more capable technology is finding its way into ever more inexpensive devices, essentially putting a computer in everything with which we interact and in every environment we

---

D.R. Searing, Ph.D. (✉)  
Syncere Systems, Lawrenceville, GA, USA  
e-mail: [desearing@synceresystems.com](mailto:desearing@synceresystems.com)

E.A.M. Searing  
Andrew Young School of Policy Studies, Georgia State University, Atlanta, GA, USA

inhabit. This computing capacity is amplified by the interconnectedness provided by the ever-present network that wirelessly flows through the radio spectrum all around us.

The world is in the fairly early stages of the spread of these technologies, and one could argue that there is a long way to go until the word *pervasive* would truly be appropriate to describe it; however, the exponential nature of the growth being experienced indicates that pervasiveness, in the truest sense of the word, is coming sooner than one might think. The case studies included below focus on recent events that presage the true extent of where the technology is going to take us and give us a glimpse of what issues are being brought to the forefront by the technologies they represent.

## 2.2 Functional Decomposition

It is important, before we look at an area as broad as the PICT technologies, to unpack the words used in its definition and build a structure for the space we are going to represent with the cases being discussed. From an engineering standpoint, we can use the process of functional decomposition to break this area and its technologies into their constituent pieces and examine the effects of those pieces on their contextual socio-technical systems to get a starting point to examine the overall emergent behavior of the system as a whole. The framework we will be using here for this decomposition is that of the basic sensor-effector system model commonly used in agent design architectures.

The sensor-effector system model describes a common pattern found in nature and in most technological designs. Its approach decomposes a system into sets of interconnected sensors, effectors, and the decision modules that connect them. A sensor is a component that perceives an aspect of the world surrounding the system. An effector is a component that affects the world surrounding the system in some manner. The logic in the decision modules converts the signals from the sensor into signals that drive the effectors to some action.

The most commonly used example of a designed sensor-effector system is the heating system in a house. The thermostat mounted on the wall contains a thermometer (sensor) and a control system, while the furnace is the effector. The thermometer *senses* the temperature in the house, and the furnace turns on to *effect* a change. The control logic in the thermostat compares the difference between the desired and the measured temperature and turns on the furnace to change the environment and thus a future measured temperature. When there is no difference between the desired and measured temperature, the logic in the thermostat turns off the furnace, thus removing the effect driving the change in the system. The temperature will then naturally drop and the cycle will repeat.

The sensor-effector model can be used to describe natural systems as well. Living organisms can be expressed as sensor-effector systems, as can larger-scale phenomena such as climate. Even the simplest forms of life exhibit this type of behavior. Bacteria sense chemical concentrations in their environment and use their

various methods of locomotion (effectors) to move to a more advantageous location (e.g., one with a higher concentration of nutrients, or one with a lower concentration of chemicals toxic to the organism). Trees and plants sense light and grow (effector) toward the area of greatest intensity.

Using this model, the concept of PICT systems can be broken into the slightly more compact concepts of pervasive sensors, effectors, and control logic. Pervasive sensors imply a world where the many types of sensors available in our technological repertoire are found throughout all aspects of our lives and environments: physical, virtual, and mental. One can imagine a world where every aspect of our lives is under constant multi-modal scrutiny no matter where we are. The Internet gives us a good starting point for considering the issue of constant surveillance in a more tangible fashion as we examine our experiences and relationships (open and hidden) with companies like Google (which tracks our searched terms, email, and site visits to create detailed marketing profiles for their customers) and governmental agencies like the National Security Agency (NSA) in the United States (which monitors our online communications) (Bamford 2012). Pervasive monitoring is not only a concern in our online world, but in the physical world as well, as evidenced by the proliferation of cameras in the public places of the major cities of the world (e.g., London, Chicago). These two examples hint at another aspect for us to consider in our analysis: the pervasiveness of the technology may affect us differently or advance at different paces in each of those worlds. Google's and the NSA's pervasive monitoring efforts are currently focused on the virtual or online world of the Internet, while the city police's surveillance cameras are focused on the real or physical world. The mental world is another world that is increasingly becoming available to technological sensors, though not quite at the rate of the other worlds. It bears consideration, though, because recent advances in tools such as the fMRI indicate that the mental world of each of us may, one day soon, be as accessible as the other worlds already are to surveillance. The ethical implications and issues surrounding this type of surveillance, both real and virtual, are examined more closely in a number of other chapters later in this book.

The case study related to pervasive sensing in this chapter takes a different tack on the issue of pervasive sensors. It focuses on the changes in the behaviors of the inhabitants (agents) of the system when useful sensors become pervasive. The specific sensors being considered are the global position system (GPS) handset and the mobile communication technologies. The case and its analysis touch primarily on the implications of living in a world where the sensor technology has become pervasive enough to make most users completely dependent on it and subject to severe consequences when the technology fails or does not work as anticipated.

Worlds in which technological systems can affect all of the inhabitants and their behaviors and environments fall under the pervasive effectors umbrella in our conceptual model. Pervasiveness is the key here, as technological systems already provide control in many areas of our lives (e.g., the thermostat discussed earlier). The increasing penetration of networked computational units into our real-world objects increases the risk that currently non-threatening objects can become more dangerous or more open to allowing others to cause us harm. For example,

over the last several decades automobiles have become increasingly computerized and networked. The computerization was undertaken to provide benefits such as additional safety (e.g., stability systems, anti-lock brakes, air bags), improved performance (e.g., fuel injection controllers), and convenience (e.g., OnStar, climate control, tire inflation sensors, entertainment systems). Though most of us do not realize it, our cars are in essence rolling data centers, full of computers, sensors, and network components, but lacking the basic protections that would be expected of an actual online data center (e.g., firewalls, encrypted communications protocols, intrusion detection devices). It has recently been demonstrated that a malicious person could take advantage of this lack of security to take control of someone else's car and cause it to unexpectedly brake, flash its lights, unlock its doors, and display messages to the driver through the instrument cluster (Koscher et al. 2010). The pervasiveness of the computing systems throughout the mechanical systems in the car gives those systems unprecedented intentional and unintentional abilities to cause physical effects and damage. This cross-over of the effects from the virtual to the physical world is a tell-tale characteristic of the systems that fall into the pervasive effector category.

This example of the computerized automobile also gives rise to another aspect for consideration when looking at these pervasive technologies. Besides the sensor-effector considerations and the multiple world considerations, one must also look at the intentionality of the harmful effects caused by the technology. In the example of pervasive effectors demonstrated in the automobile, it should be apparent that none of the systems provided in the car was provided with the intention to cause harm. They were furnished to provide additional safety, performance enhancements, and occupant convenience and comfort. The dangers came from the unforeseen uses of the system by a malefactor to cause harm to the car, its occupants, and others nearby. While it may be fairly easy to see how intentionally building a system to cause harm to its occupants would be subject to moral sanction, there is more room for debate in claiming that a failure to provide adequate safeguards against the possible misuse of the technology by malefactors is also subject to a similar level of moral opprobrium. The example case below focuses on both these aspects of intentionality while examining the deployment of the Stuxnet virus – the first computer virus generally recognized as designed specifically to cause real-world damage outside of a computer system (Brenner 2012) and the first generally regarded as an attempt at cyberwarefare between nations or their proxies (Farwell and Rohozinski 2011; Gross 2011).

The pervasiveness of control logic is a bit more difficult to see as a taxonomical category, due to the fact that the logic is not useful without its associated sensors and effectors. So, to imagine this world of pervasive control logic, one needs to imagine a world of pervasive systems, each using its sensors to fire its control logic and actuate its effectors. These systems can have varying levels of autonomy depending on the amount of oversight and involvement of a user in the control logic and the setting of system goals. Systems with greater levels of autonomy are more likely to be pervasive since they are not limited by a population of users who direct them. Thus, this last category of the taxonomy can be taken

to be focused on the pervasiveness of autonomous systems and agents. Pervasive autonomous systems are also subject to the multiple worlds and intentionality aspects outlined in the other category above, and all of these aspects will be considered in the third case study in this chapter. In May of 2010, the stock markets in the United States experienced an approximately 20-min period, now known as the “Flash Crash,” where trading in a number of derivatives and stocks experienced an unprecedented drop in value followed by a mysteriously quick recovery. In the ensuing investigation, we observed that the calamity was driven by the unanticipated interactions between automated trading systems that were operating properly in pursuit of their programmed goals.

## 2.3 Lost in Death Valley – Sensor Pervasiveness

### 2.3.1 *The Case*<sup>1</sup>

In the summer of 2009, a mother, son, and their dog set out on a camping trip in the mountains and deserts west of Las Vegas. The woman had just recently moved to the Las Vegas area to take a nursing job, and she wanted to spend some quality time with her son while exploring her new environs. She had informed her family that she and her son would be camping for the weekend at a popular free camping site near the south end of the Death Valley National Park in California. They set off for the weekend in their SUV with enough supplies (including water) for the trip and a GPS system to help in the navigation from their home to the camp site.

As they crossed through a remote part of southern Death Valley, the woman’s GPS system instructed her to turn off the paved road onto a road that was little more than a dirt trace barely suitable for any sort of vehicle. Following the instructions on the GPS, she continued down the road despite its condition. The road got progressively worse, with deeper sand and larger rocks, and eventually one of the tires on the SUV went flat. She stopped and changed the tire, and then continued down the road following the directions being provided to her as any traces of an actual road disappeared. Several miles further down the track, the SUV passed over an animal’s den, collapsing it and hopelessly miring the car in sand and dirt up to the axles.

Once her car had become immobilized, the woman attempted to call for help, but her mobile phone was well outside of the range of any towers that could provide it connectivity. The family used their weekend’s provisions to stay alive as they first hiked unsuccessfully to find help or a mobile phone signal and then resigned themselves to wait near the car with the hopes that someone would eventually pass by and find them.

---

<sup>1</sup>Based on information provided in the August 8, 2009 Associated Press story, “[11-Year-Old Boy Dies After Mom Says GPS Left Them Stranded in Death Valley.](#)”

Not hearing from the mother and son for four days, her extended family members called the authorities to report them missing. They provided the search and rescue teams that were assembled with the location where the woman and son were to be camping, and air and ground search and rescue teams went out looking for them. On Thursday, a full five days after the woman had left on her trip, a park ranger noticed tire tracks heading off the main road onto an abandoned mining road, and upon following the tracks for a few miles discovered the evidence of the tire change. He continued following the tracks until he found the family's stranded SUV with the mother frantically signaling him. By this point, the mother and dog were severely dehydrated and the son was dead from the combination of the heat (daytime highs at the time had reached 111 °F) and lack of water. Upon being questioned after her rescue, the mother insisted that the GPS had directed her down that road.

### 2.3.2 *Analysis*

There are two pervasive technologies that are at work in this case: the GPS system and the mobile telephone. These two technologies have become indispensable in many people's lives. The mobile phone gives its users constant connection and availability to communicate with friends, coworkers, and family, not to mention emergency authorities when necessary. The GPS unit gives users the navigation information they need, and in fact many smartphones have GPS systems built into them. Users of these technologies have become habituated to always knowing where they are and where they are going, and having ready contact with everyone in their lives at all times. The technology works wonderfully, until it doesn't, and that transition can be sharp and unexpected. Both systems require continuous contact with their base stations (cell towers for mobile devices, and the constellation of GPS satellites for the GPS unit) and both technologies become useless when they are taken out of the range of their larger system context. Mobile phones lose their connectivity in rural areas and in areas with geographies or structures that block signals, or when they are intentionally turned off (e.g., on commercial aircraft flights). Additionally, as many discovered in New York on September 11th, 2001, the cell networks don't always have the capacity to service everyone if there are high volumes of calls or if critical infrastructure (towers, cables, etc.) has been damaged. GPS systems work poorly in areas where the view of the sky is limited, including inside buildings, parking structures, and tunnels. For the continuously-connected generations, it can be difficult to function normally in areas where these tools lose their reference signals and cease functioning properly.

Both the behavior of the users and the structure of the worlds they inhabit have been changed by the ubiquity of these technologies. The mobile phone has essentially eliminated the public pay phone. Thus, its sheer ubiquity has removed the alternatives that would be the fall-back options when the mobile systems fail or cease functioning in the expected ways. Additionally, people's behavior has changed as a result of the new technology. The idea of never being out of contact

with friends, family, and authorities provides a sense of security. If one's car breaks down, emergency services are a quick call away. Coordination between people, such as in finding a meeting place, is now done in real-time while walking towards the proposed location and each other. These are tasks that in a pre-mobile telephone world had to be planned out in advance and thought through before commencing the activity. Now they take place in an ad hoc fashion while on the move. These types of ad hoc planning sessions fall apart completely when the enabling technology becomes unavailable, and our world has removed many of the fall-backs we once had which would allow for a less "digital" boundary between functioning and failure.

With the changes in the world and the behaviors of its inhabitants, we do not find it out of the ordinary that the woman in this case would head out on a camping trip feeling secure in the presence of her enabling technologies. If trouble arose, the cell phone would be there to call for help, and besides, the GPS unit would keep them out of trouble by leading them to and from their destination. But this case illustrates that these assumptions can be deadly and that the limitations of the technology, while widely known (everyone has had their phone drop out of signal at some point), can be easily forgotten by the casual user, especially when it comes to considering a back-up plan. Experts in the areas of hiking and mountain climbing will often bring back-up gear, and in the most forbidding of locations, emergency beacons that can allow search and rescue teams to locate them. Experts do not assume that their activities will be non-threatening, and they better understand the risks inherent in what they are doing and take suitable precautions. In the past, getting to many remote locations in the wilderness required being an expert or having one as a guide, and it is this that has changed with these enabling technologies. Casual users are finding themselves in locations only experts would have gone in the past. Their device-enabled lack of fear allows them to push out into areas previously only accessible to prepared experts, leading to tragedies such as the one described above. In these areas, remote dirt tracks that would give pause to an expert in a normal SUV now are driven on because the GPS unit directs it and the casual explorer does not second-guess the device or necessarily understand that the trail in front of them is beyond the capabilities of their vehicle.

So, it is understandable that the socio-technical system of the users and their environment has changed, but it begs the question of just who in this case is responsible for the tragedy that befell this family. The GPS unit is a tool containing the necessary relativistic algorithms to locate the unit in space, a set of maps in memory that scroll to the location indicated by the sensors, and route-planning algorithms that provide directions to the user based on the current and desired locations.

## **Designers of the Devices**

GPS is a non-autonomous tool, so the responsibility for its behavior must continue to fall upon its developers. These developers provided the algorithms and provided the data used in the mapping system, which in this case first indicated that there was a road where there was not one, and then directed the user onto it. The inclusion

of roads that no longer exist on these maps is of great concern to the rangers in Death Valley National Park, who have subsequently lobbied the map data providers to remove these old roads from the maps of their park (Clark 2011). Dangerous incidents related to outdated maps and modified user behavior are not just an American phenomenon; there is a documented case in Spain where a GPS system directed a driver down a road into a reservoir that had been in place for over 20 years (Tremlett 2010).<sup>2</sup> The designers of these systems understood that maps do change and most if not all GPS units are capable of receiving map data updates online or through other means. What they failed to do, in the authors' collective opinion, is put in place the safeguards that would (1) prevent the bad data from being entered into the system and (2) ensure that no one could use the system without the updated data. They failed, not surprisingly, to anticipate the types of changes that would occur in the socio-technical system of users and their tools when this new technology was provided to them and did not build in these sorts of safeguards.

Beyond the data provided, the route-planning algorithms aren't required to be designed to prevent routes from being planned down all manner of roads. Most units not only contain information that there is a road but also have access to the properties of the road in terms of speed limit and road type (divided highway, multiple lanes, paved or unpaved, etc.). When navigating between destinations near major roads or interstates, it has been the authors' experience that the GPS systems do take road attributes into account, but it has also been our experience that when away from major roads, the directions can become increasingly focused on getting from one point to another and less focused on the quality of the roads being selected in the routing. The planning algorithms should not include unsuitable roads as connections between destinations for all types of routing or should allow the user to configure the types of roads they are comfortable traversing. Recent patents show that these algorithms are moving in the direction of being more cognizant of the dangers of the routes they are suggesting and allowing users to input the criteria used for their route planning (Herbert 2012). Of course in the changes discussed in the article cited, Microsoft's algorithm, which gives the user the ability to set their GPS unit to avoid "bad" (i.e., crime-ridden) neighborhoods in its directions, trades in our current set of moral concerns for an even more contentious set involving the labeling of neighborhoods as "bad".

## Users of the Devices

Aside from the developers of these systems, who else bears responsibility for the use of these technologies? The users of the technologies should logically be a second set of responsible parties in this analysis. The users of these technologies get into trouble when they cede too much of their judgment to the devices, and that is driven

---

<sup>2</sup>Chapter 8 by Bo Brinkman looks at additional issues that will arise in the approaching world of augmented reality, where the data being blindly relied upon contains more than just simple mapping data and the user's perception of reality itself comes into question.



by a lack of understanding of the devices and their limitations as well as a misplaced balance of trust in the device's judgment over their own. As stated earlier in the analysis, expert users are experts due to their understanding of the underlying risks in the undertaken activities. Casual users do not always understand the dangers that they are putting themselves in, and in a pre-GPS and mobile phone-enabled world would not have been as easily capable of putting themselves into the situations they now find themselves in. That being said, the tragedy outlined above was not only a result of a user of these risk-enabling technologies going outside of their safe zone of travel. It was also a failure in the judgment of the user of the system in following the directions uncritically. The user in this case continued down a path that in all reports was obviously not a passable road and left no sign of trying to turn around even after her tire was damaged.

In the Spanish case, the user was driving late at night and unable to see the reservoir (Tremlett 2010). That death seems to us to be less the responsibility of the driver than of the designers of the system that lacked updated map data.

It is the change in our behavior due to the pervasiveness of these technologies that was not necessarily foreseen by the designers of these systems. The original users of the GPS systems were the military and others who were experienced in using older systems (e.g., LORAN) – in other words, expert users. The system worked so well, though, that additional civilian demand for its services led to its inclusion into more and more devices and thus brought its power into the hands of the casual user. These users saw a useful feature added to their car's electronics or to their mobile phone and they changed their behaviors accordingly. Now they would never be lost, never need directions, and with their phone would always be in contact with others who could help them when they needed it. And they were correct: When used in well-covered cities and metropolitan areas, the devices work quite well. The danger comes, as it always does, on the fringes of these areas, where signals fade out and roads become less maintained.

The resolution to this problem is going to come from both of these responsible parties – the designers and the users. The GPS systems are going to get smarter, the map data is going to get better through the efforts of those responsible for the maps and even through the efforts of those who maintain wild areas such as the rangers in Death Valley. The failures of the GPS navigation systems will serve to ensure that the next generation of devices will contain some of the safeguards outlined above to prevent future incidents. Of equal importance, though, is the continued understanding that these enabling technologies require additional user training as they become easier to use. This seems counter-intuitive, simpler devices requiring more training, but the users of these technologies must be alerted to the fact that these devices are not foolproof and have limitations that may endanger their safety. Additionally, the training should have the primary function of alerting users that these devices may alter their behavior and lead them to discount the risks being taken.

With the lessons learned from these GPS-related tragedies, the GPS designers are building in more safeguards, and through exposure to these sorts of cases, the users are at least more cognizant of the types of issues that might arise. But, this change in behavior is not limited to only this type specific type of technology (Brennan 2010;

Denning et al. 2009). These types of enabling technologies and the perception of increased safety they engender commonly trigger risk compensation and actually drive the increases in risky behavior. The designers of these types of systems should be aware of this phenomenon and commensurately design appropriate safeguards. The users must also bear responsibility for understanding the nature and the limitations of the devices to which they are entrusting their safety.

## 2.4 The Worm Turns – Effector Pervasiveness

### 2.4.1 *The Case*<sup>3</sup>

In April 2010, scientists in Iran were finally beginning to implement their plans to enrich their store of uranium ore to isotopes usable in their plans for future power generation and possibly nuclear weapons research. They had run afoul of the United Nations and the International Atomic Energy Agency (IAEA)<sup>4</sup> in the past and the work was being done in a number of separate underground facilities deep in the heart of the country. These facilities consisted of great halls that contained hundreds of computer-controlled, high-speed centrifuges that were used to separate the more useful heavier isotopes of uranium from the lighter ones. The process requires precise control of the speed of the centrifuges, and Iran had purchased top-of-the-line motors from Finland and controllers from a manufacturer in Germany known for its quality and precision.

As the initial runs of the centrifuges completed, the scientists noticed that the expected output from the cascade of centrifuges was not being achieved. Additionally, the centrifuges were failing at a rate greater than was expected. The problems continued to hamper the effectiveness of the program as investigations began into what was causing the failures.

In June, an anti-virus researcher in Europe found an interesting worm infecting some of his client's computers in Iran that appeared to use a number of novel mechanisms to spread itself and had a large, apparently non-functional payload. It was given the name Stuxnet, based on a sample of text found within the compiled code of the virus (Gross 2011). Further research continued at a number of anti-virus system providers and they discovered that the payload was not inert, but it was a highly targeted attack aimed at an industrial controller used in high-speed motor

---

<sup>3</sup>The case is based on a very detailed narrative of the timeline provided in (Gross 2011) and information provided in a number of IAEA reports such as (IAEA 2009). Both are excellent reading and we would highly recommend the interested reader seek them out for more details regarding the events surrounding this case.

<sup>4</sup>The IAEA is charged with verifying Iran's compliance with a number of UN resolutions related to their nuclear program. See the February 19, 2009 report for details of the resolutions and non-compliance (IAEA 2009).

controls manufactured by a specific German company. Additionally, the trigger for the payload would only execute if certain other conditions were met: The controller must be operating within a certain target range of parameters and must be in a system that contained a combination of frequency converters built by certain Finnish and Iranian manufacturers (Gross 2011). When the payload was triggered, the worm was designed to cause the frequency converters used to control the high-speed motors to cycle unpredictably and remain unable to hold a specific speed within a certain RPM range. A second prong of the payload would then modify the logging systems in the controllers to report that a constant speed actually had been maintained (Gross 2011). All of the speed cycling would cause premature wear in the motor and eventually cause the failure of the system well inside of its normal expected lifespan while leaving no trace in the system logs as to the cause of the wear.

Researchers concluded that the code was likely aimed specifically at the cascades of uranium enrichment centrifuges being deployed in Iran based on the trigger conditions and the effects of the worm. Secretive Iran had not publicly admitted to any issues at this point with their centrifuges, but satellite observations indicated to other countries that something was going wrong at the enrichment facilities as the centrifuges were being replaced at a higher than normal rate.

Speculation began on the source of this highly targeted worm, and based on certain code metrics (size and complexity of the payload, the languages and libraries used, and the constructs and patterns followed), the advanced nature of the exploits (four vulnerabilities in software that had not previously been exploited were used for propagation), the evidence of the use of a stolen code-signing certificate (Gross 2011), and the precise targeting, most researchers came to believe that the worm was created by professional software developers likely in the employ of the intelligence services of the government of the United States or Israel. As of this writing, though, neither state has officially admitted to having played any role in the development or deployment of the worm, though recent leaks provided to the *New York Times* seem to lend additional credence to that conclusion (Sanger 2012).

### 2.4.2 Analysis

The Stuxnet worm is not the first attempt at attacking the computers of an opposing or hostile state, but it is apparent evidence of the escalation both of the severity of such attacks and of the intention of harm inherent in them (i.e., the intention is now to target systems that control physical effectors). Previous cyber-attacks have focused on the virtual world, such as interference (i.e., preventing access to sites through denial of service attacks) and espionage (Hollis 2011). Stuxnet is the first known actual attack that used a virtual approach to cause physical harm. The goal in this instance was physical harm to a set of machines, causing their subsequent failure, but the cross-over from virtual to physical harm does not have to stop there. This escalation and modality change in the harm coupled with the growing pervasiveness of the computational effectors leads to our primary concern.

The more the computer-driven effectors permeate our environment, the greater the potential risk to the public. As stated earlier, our automobiles are a good example of the potential risk we face. The modern automobile is a collection of networked computers, sensors, and effectors. If a malicious person or a worm/virus attacked the braking system, for example, accidents could be instigated and injury and death could result. As has been shown by Koscher et al. (2010), these systems are currently rather insecure and the types of effects posited above come directly from the effects they managed to invoke in their testing. Again, though, this is an area of concern that has already been identified and mitigation efforts have begun. Other areas are awakening to the risks of physically-targeted cyberwarefare attacks, including on personal medical devices (e.g., pacemakers, insulin pumps) (Halperin et al. 2008) and the avionics systems of commercial airliners (Arthur 2012).

The escalation of the risk from these attacks takes on two different meanings for the designers and users of these systems depending on which side of the exploit the designer and user are on: the aggressor or the defender. As the viruses and worms move out of their traditional anarchical headwaters and into the mainstream as professionally developed weapons of cyberwarefare, one must ask about the responsibility of the engineers designing them and the states deploying them as weapons. Like other weapons, these tools can be targeted at the military and their supporting manufacturing centers (e.g., like Stuxnet, purportedly), or they can be used against the civilian population directly or indirectly to cause collateral damage. One could imagine a weaponized worm or virus used to disable a country's computerized power grid or to disable the safety features on a nuclear power station causing massive damage to nearby civilians and property. These systems would need to be handled as weapons, which is what they are, and they should be developed and utilized following the provisions allotted to other weapons development programs, and be subject to all established rules of war, rules of engagement, and any additional conventions placed on weapons that have the capabilities of mass civilian destruction. The developers of these systems would by necessity need to understand the extent and uses of their systems and be reconciled with doing weapons work even if they do not believe that such activity is likely to be pursued. If weapon-like capacity is designed, such a capacity should be assumed in the moral considerations of the design process.

Of equal interest are the decisions and the dilemmas of the developers responsible for defending these effector systems against attack. As stated in the introduction, the moral issues related to building weapons or tools that might harm people intentionally is fairly well developed (e.g., *jus ad bellum* and *jus in bello*<sup>5</sup> discussions), but the moral obligation of a developer of a system to keep it secure from such exploitation (foreseen and unforeseen) is a more difficult case to make. Much is made in engineering ethics of holding the safety of the public paramount, but that is primarily in terms of preventing accidents or intentional damage done by side effects of the engineers' own systems and their usage. Only a few areas

---

<sup>5</sup>These are concepts in just war theory: *jus ad bellum* concerns the right to wage a war and *jus in bello* concerns acceptable behavior in waging war.

in engineering (e.g., military systems design and software engineering) routinely incorporate defenses against malicious use of systems. In these situations, the level of care and responsibility taken by the engineer is usually the level requested by each client or user and verified by third party auditors. But, in order for this approach to work, the clients and the designers/engineers need to be aware of the risks inherent in these sorts of systems. As shown with the automobile example, the systems in an automobile were originally designed in a non-networked world or in a world where there was a tightly controlled network amongst the systems of the automobile without external interface points outside of the diagnostic ports inside the car. But with the advent of wireless technologies used to communicate tire pressures from the tires to the onboard computer and dashboard displays, and systems like OnStar which provide active controls to authorized systems over a wireless network, the internal network of the automobile is more exposed than ever and thus new levels of responsibility for security and safety need to be considered.

This shifting of the priorities related to security and safety was exactly what the designers of Stuxnet exploited. The industrial control systems targeted were traditionally not networked together with external access points, so the necessity of providing a secure barrier around the systems was not present during their initial design and development. This made these systems open for exploitation once they began being networked to personal computers inside their facilities. These are the exact types of control systems that are finding their way into all sorts of devices as “smart” technology is deployed to our vehicles, homes, offices, and critical infrastructure systems. As these systems become more pervasive and carry the possibility of having greater impact on our lives through their interactions with us, it becomes incumbent upon the designers of the systems to build in the necessary safeguards to secure the systems directing these effectors and to limit the amount of damage that can be done with them. The users of these systems play a role here as well: they must educate themselves about the risks inherent in the systems they choose to deploy, and they must deploy them wisely with the proper configurations in place to ensure their safe and secure operation.

## 2.5 The Flash Crash – System Pervasiveness

### 2.5.1 *The Case*<sup>6</sup>

On May 6th, 2010 the American stock markets experienced one of the most unprecedented single-day events in their history. In the early afternoon of what was an unimpressive trading day, one firm’s misconfigured trade execution program

---

<sup>6</sup>The timeline of this case study is based on information provided in the CFTC and SEC report issued in September of 2010 (U.S. Commodity Futures Trading Commission [CFTC] and U.S. Securities and Exchange Commission [SEC] 2010). The events were reconstructed using market records in their attempt to trace the events of that afternoon back to its root causes.

precipitated a crisis which rippled across the futures markets and into the stock markets, at one point sending the Dow Jones Industrial Average (DJIA) down almost 1,000 points (10 % of the index's value) in a matter of minutes. Then, just as quickly, the slide reversed itself and the markets recovered most of their value.

The Security and Exchange Commission (SEC) and Commodity Futures Trading Commission (CFTC) have performed an extensive analysis of the events that occurred on May 6th, 2010 during the so-called "Flash Crash" using all available records of the trades made that day across all of the markets that were affected. The timeline that resulted from this analysis is summarized below.

### **May 6th, 2010 – Morning and Early Afternoon**

May 6th started as a fairly negative day for the market. News from the ongoing European debt crisis was driving the market lower and by 2:30 PM the DJIA was down close to 2.5 %. The volatility measurements had spiked up (i.e., VIX was up 22.5 %), the Euro was declining against the Dollar and the Yen, and a specific high-volume type of futures product known as an "E-mini" had dropped in value significantly during the session (down 55 %).

### **May 6th – 2:32 PM**

At 2:32 PM, a large trading firm began the execution of a sell program into the E-mini market. The firm was selling 75,000 futures contracts worth around \$4.1 billion in order to hedge a position they had already taken in other equities. The traders at the firm decided to use an automated selling program to handle the transaction due to the large number of contracts that would be traded. The program had been configured to sell the contracts at a rate that was solely dependent on the volume of contracts that were currently trading in the market. This trader had executed trades this large before, but always as a combination of manual and automated programs that had their rates controlled by volume, price, and time. Whereas the previous sell programs had executed over a number of hours, the program used in this instance executed the entire sale in 20 min.

The large sales of contracts by the trading firm was initially absorbed by other automated trading programs known as High-Frequency Traders (HFTs), which execute trades in sub-millisecond time frames. These systems tend to trade amongst themselves at high speeds and volumes without substantially changing their overall positions; attempting to profit from arbitraging across the near-instantaneous price disconnects. Some of these systems were trading in just the E-minis, while others were attempting arbitrage between the exchanges. E-mini futures contracts are fundamentally related to the S&P 500 SPDR exchange-traded fund (CFTC & SEC 2010), so there are opportunities to take positions in the futures and in the exchange-traded fund (ETF) and profit from discrepancies in the pricing between the two markets.

**May 6th – 2:41 PM**

Starting at 2:41 PM, a large volume of E-mini contracts began to trade back and forth between the HFTs, a process which had been set in motion by the initial sales from the hedging trading firm and its volume-only-paced trading algorithm. The increase in volume thus drove the trading firm algorithm to increase its sales volume into the E-mini market. This positive-feedback loop worked upon itself for the next four minutes, affecting the E-mini market and spilling over into the related SPDR ETF market due to those that were attempting arbitrage between the markets.

In the four minutes between 2:41 PM and 2:45 PM, over 140,000 E-mini contracts were traded back and forth amongst the HFTs in the market (representing 33 % of the trading volume in that market). The E-mini contract price had dropped 3 % in those four minutes, and due to the arbitrage with the SPDR ETF market, so had the price of the ETF.

**May 6th – 2:45 PM**

In the 14 s after 2:45:13 PM, the HFTs traded an additional 27,000 E-mini contracts (representing 49 % of the market during that time period). This accelerating trading and the commensurate drop in liquidity in the market triggered a 5 s market stop in the E-mini marketplace, halting any further buying in E-mini contracts. When trading resumed after the stop, the E-mini market stabilized and prices on those contracts and the related SPDR ETF began to rise.

**May 6th – 2:45 PM–3:00 PM**

With the crisis over in the E-mini marketplace, it seemed that disaster had been averted. The market had stabilized, and prices began to rise. But during this recovery in the E-mini contract market, a far broader crisis began; one that had been instigated by the problems that had occurred over the last few minutes. As previously stated, the E-mini contract is a futures contract related to the SPDR ETF which in itself is a product that is derived from the values of stocks in the S&P 500. Traders (human and automated programs) in the markets where the S&P 500 equities are traded took the price signal from the SPDR ETF as a portent of some significant unknown event and they began to react. The equities involved and other ETFs related to them then began to see enormous increases in trading volume as the broader markets responded to the price drops seen in the SPDR ETF. By 3:00 PM, two billion related shares (worth around \$56 Billion) would be traded in these equities.

The quickly increasing trading volume caused a number of these equities and ETFs to undergo their own liquidity crises as the available shares evaporated from the market and several of them (e.g., Accenture) began trading on “stub” quotes. “Stub” quotes are quotes that are used as a pulse by trading firms and exchanges to essentially keep quotes on a stock continuously in play in the market. If these

quotes are not actively being utilized for trading, the quotes are given trading values well outside the normal price range (e.g., a \$0.01 bid, or a \$100,000 offer). During this time period, some stocks experienced enough of a liquidity crisis to cause the stub quotes to become the only actionable quotes, which caused prices to drop to irrational levels. For example, quotes on Accenture dropped from around \$30 to \$0.01 in 7 s during the height of the volume increase (CFTC & SEC 2010). These quotes continued to be purchased since many of the automated programs were issuing “Sell at Any Price” orders, which normally execute in the rational range around the market price, but in this instance executed on the stub quotes.

As 3:00 PM approached, the market mechanisms (human and automated) began to react to the stabilization in the original futures and ETF market that had instigated the crisis, and the prices quickly reverted to more normal levels. In the same way that the automated execution and feedback mechanisms had driven the markets down like dominoes, they just as quickly drove each other back-up to a stable level.

The DJIA numbers for the day show the effects of this event at a macro level with a swing of around 700 points down and 600 points up happening in the span of 15 min before 3:00 PM. But on a position-by-position basis in the market many trades ended up broken and many traders and programs adjusted their positions unnecessarily (albeit that was unknown to most of the participants in the markets at the time) and incurred costs associated with that trading. Regulators and the exchanges determined that 20,000 of the trades that had occurred during the 2:45 to 3:00 PM time period had occurred at prices deemed irrational (further than 60 % away from their prices at 2:40 PM), and these were subsequently reversed, thus mitigating some of the damage done. What could not be mitigated, though, was the damage done to the perception of the stability of the markets and the new cast of characters (HFTs, automated trading programs) which now inhabit them.

### 2.5.2 *Analysis*

In this final analysis, we move past the simplistic linear causal relationships that we have been concerned with related to effectors and sensors, and step into the complicated ecology of pervasive interconnected systems. In this scenario, multiple agents involved with the trading algorithm and its environment – designers, builders, implementers, regulators, users, and consumers – are brought into the discussion of assigning responsibility for its behavior and developing approaches for mitigating the risk of this sort of behavior in the future. The sensor-effector relationships are knit into a causal fabric, and we need to examine and understand each piece in order to fully grasp the emerging (and often unanticipated) synergies. To do so, we’ll examine the ethical obligations of the primary players and then integrate them to produce an ethical schematic of the environment surrounding the Flash Crash.



## Financial Professionals

The individuals who design and execute the financial instruments and strategies used in contemporary markets spend their lives estimating, quantifying, and mitigating risk. Transactions are based on there being, at least theoretically, a buyer and a seller: depending on which way the market turns (or whether there's a late frost, or whether someone defaults on a mortgage, etc.), someone is going to be monetarily worse off than someone else. However, the argument is made that, at the moment the transaction occurred, the act made both parties think they were better off or it would not have taken place.

This is an example of Adam Smith's famous invisible hand – through individual pursuit of selfish interests, the net welfare of the group increases (Smith 1991). It is the reliance on this familiar assignment of aggregate benefit to individual actors that originally made systemic problems such as externalities a difficult issue to explain through economic theory. For example, a paper mill in a small town during the 1970s would provide jobs for both factory workers and loggers in exchange for the market price for their good; this price would cover wages, raw materials, operating costs, etc. What would probably not have been reflected in this price is the cost of the environmental damage that would occur from the release into the environment of chlorine and other chemical agents used in the processing of paper. This is not a cost that is billed every month, but it is a cumulative negative impact directly attributable to the paper mill. What should be done to the price and quantity produced to reflect this adverse side effect? And what if the externality was something less tangible than chlorine levels, such as impact on native animal habitat, or if the damage was not directly attributable to a single producer, such as smog?

As with industrial pollution, economists have long urged what are called "Coasian" solutions (after 1991 Nobel Prize winner Ronald Coase) for such situations: assign ownership to the problem, and then the market will take care of itself. For example, if one farmer's cattle keep wandering over to a neighboring farm and eating corn, all that is required is for each farmer to accept ownership of their cattle or crop and then negotiate between them whether the cow owner should preemptively pay for any bovine transgression, the corn owner should subsidize the cow owner's extra cattle hand, or they should jointly finance a fence (Coase 1960). The anonymity that comes from the large-scale aggregation of the market can act as an inhibitor to these types of solutions; Coase himself argued that his approach to valuing externalities did not function as well when the number of agents involved became large. Unfortunately, that anonymity is also a lubricant that helps the market function as efficiently as it does – you maximize profit by being a step ahead of the herd, which only happens when the herd isn't watching you closely. Additionally, there are often numerous endemic conflicts of interest, such as commission systems and the internalization of stock and bond orders, which keep the field lean and efficient, but perhaps not ethical (Battalio and Loughran 2008). These adversarial, efficient, and often mutually-rewarding relationships run throughout the incentive framework and culture of finance.

In order to keep these competitive drives harnessed, market professionals are subject to several layers of ethical guidelines, such as fiduciary or suitability standards and professional codes of ethics. At the base level of ethical compliance is the accountability level assigned by law, which for financial advisors is a “fiduciary standard” and for stock brokers is a “suitability standard” (Angel and McCabe 2010). A suitability standard only requires the agent find a satisfactory or acceptable financial instrument; a fiduciary standard, however, requires disclosure and avoidance of any potential conflicts of interest and placing of the client’s interests above the financial advisor’s (Bell 2010). Currently, there is debate as to whether all client-facing financial professionals should meet the fiduciary requirement, which would strengthen the accountability between professional and client. Stockbrokers argue that not having such exhaustive informational requirements allows them to remain unbiased and at much lower cost; financial advisors claim that such positions are merely sales and are often used to soak up securities that suit the broker’s interests (such as an in-house offering), rather than offer any cost savings on the client’s part (Serchuk 2009). Additionally, the possible arrival and implications of a new standard called “best interest” is now being debated (for a thorough explanation of “best interest,” see Fein 2010). Currently, however, there are no professional standards for the individual behind the computer developing the algorithm for statistical arbitrage (this will be discussed in more detail in the next section), which begs the question if we are legislating a solution to a problem, but not the problem we thought we were solving.

Many corporations, such as Charles Schwab and the Seattle Northwest Securities Corporation, and many professional organizations, such as the National Association of Personal Finance Advisors (NAPFA) and Certified Financial Planner Board of Standards (CFP), have documented codes of ethics. Aside from the legal standards mentioned above, such codes are market signals to potential clients that ethical and professional considerations are openly posted, discussed, and, thus, more likely to be heeded (Adams et al. 2001); this signaling serves as marketing, ironically making even codes of ethics part of the competitive framework. The results of the use of these codes of ethics, including whether they generate more ethical behavior, are mixed, though the relationship falls between positive and neutral (Schwartz 2001). Further it can be argued (and very often is) that the self-regulating nature of the market would punish those considered unethical; however, the necessary assumptions of perfect information and no transaction costs may be too stringent to reflect reality. In an industry designed to push the limits of risk and profit on the margins, how do we ensure that the structure given by a sense of professional ethics is not approached in the same fashion?

## **Software Engineer**

The algorithms designed by the financial professionals are actualized as software systems by the software engineers. The systems in the financial marketplace are increasingly expected to react to changes in prices and conditions that are on the

millisecond scale. At this speed, the decisions being made through the algorithm are not manageable in real-time by a human user. If a misjudgment is being made by the algorithm's implementation, then potentially thousands of them will occur before anyone will notice and be able to correct or terminate the behavior. Thus, due to the speed of their action and the resultant lack of oversight, these algorithm implementations would have to be autonomous by default. They are provided goals by the user, but then they proceed to execute them per their algorithm, whether pre-programmed or a self-learning variant, in which the algorithm operating at any given moment is not known by the managing user.

The possibility of thousands of unseen failures and the severe monetary consequences that would be engendered leads to these algorithms being extensively tested. But most of this testing is in environments that are pale imitations of the entire interconnected financial systems into which they will eventually be deployed. The failure mode that occurred during the Flash Crash was one that emerged only when the full range of connections and systems encountered in the "real world" were present. Even testing within the first ring of systems the algorithm is connected to will not provide an environment in which this failure mode can be found. Connecting to multiple systems in a test environment does not capture the mode either, due to the fact that unless those systems are connected together through other realistic nodes, it becomes impossible to replicate the complex dynamics of actual interconnected set of systems and algorithms.

These types of failures across interconnected systems of systems are especially pernicious. First, they are dangerous because they are not obviously in a failure mode when considering the constituent systems on their own. There is no method to use to test an individual system and find these kinds of failures. Second, the inclusion of so many systems into the eventual failure mode rapidly leads to the 'problem of many hands' and the difficulty in determining who is responsible for fixing the problem once it is found. In the financial markets, the interconnected systems belong to multitudes of organizations from countries around the world. These organizations are in competition with each other, so the propensity to work together is already low. It becomes too easy to point a finger at your competitors and their "secret" algorithms, and all too easy for them to point back at you.

The other players in the market that can create stability are the interconnection points known as exchanges (e.g., NASDAQ, NYSE) that have the duty to keep their systems up and running for their connected clients, but they can only dictate behavior on the systems' actions that are executed on their exchange. As was seen in the Flash Crash, the failure mode involved systems connected to multiple markets and sending buy and sell signals between the markets. These systems that were attached to multiple markets were seeking arbitrage opportunities (i.e., momentary discrepancies in prices between two markets); they became a critical coupling point between the systems and, though subject to the interface contracts and conditions for each exchange, the signals transmitted through them and their actions in the different markets were a primary cause of the breadth of the failure. Additionally, the circuit-breakers present in each market only exacerbated the problems by creating

mixed signals and confusion for those systems that were interconnecting them; this confusion lead to improper orders being executed across all of the systems.

To mitigate the risk exposed by the Flash Crash, the designers of these algorithm implementations and the exchanges they connect to need to be able to test and certify these systems in a large-scale testing environment that can properly simulate the behavior of the full marketplace. This opinion was expressed both by the authors of this paper (Searing 2011) and additionally in a paper published by the government of the United Kingdom (Cliff and Northrop 2011) related to the understanding of the Flash Crash as an emergent condition resulting from the structure of the financial markets as an ultra-large-scale system of systems.

It is this through these new understandings of these large-scale systems of systems and their emergent failure modes that software engineers should be striving to find a way to predict and prevent such catastrophic events, but while this model gives a window into the understanding of the problem and its general cause, it does not provide motivation for the developers of these systems to accept the responsibilities that their autonomous systems may be even partially at fault. What is needed are new moral guidelines, as mentioned earlier and detailed in later chapters of this book, that address the issues related to the ‘problem of many hands’ in these systems and the propensity of people to ascribe blame to their autonomous creations rather than themselves. The rules discussed in a later chapter of this book and described by one of these authors as a Creator’s Ethics (Searing 2012) are the first attempts at laying the groundwork for these new responsibilities for designers of these pervasive and autonomous systems.

## Market Participant

For those of us not comfortable with our ability to solve simultaneous equation models or run Monte Carlo simulations, the world of finance may seem beyond our reach and understanding. Just as we let the auto mechanic work on our car and the plumber fix the leaky faucet, we rely on our retained financial experts to increase our net worth and are happy to delegate this task. However, does this absolve us of any ethical transgressions on the part of our professionals? If we take pride in our ability to pick professionals based on how much money they make on our behalf, should we not be ashamed if we pick ones who violate ethical standards? If your broker made you a lot of money by violating insider trading laws, would you rather he or she not have conducted the violation, or that you simply had not known how they did it?

As the purchaser of financial services, market participants have two dominant obligations. The first is to become and remain an informed consumer. We may never be able to conduct complex mathematics or design an algorithm; however, we are capable of assembling a great deal of information on both the tools and the agents who would use them on our behalf. Objecting to something insidiously called a “dark pool” by casual news coverage is easy, but understanding that it exists in

order to circumvent a 30-ms advantage given to high-frequency traders before the market gains knowledge of the trade may cast additional (and sympathetic) light on the subject. Further, the effectiveness of any regulations put in place to prevent situations such as the Flash Crash from happening again will vary depending on the course of information – you need to know enough about what happened in order to know which sources of information are giving you unbiased analysis. Ignorance may be bliss, but it is not conducive to ethical behavior.

The second obligation is to take the knowledge you have gained and reveal your preference. In the most straightforward sense, this means to spend your money in a way that reflects your judgment. If you are willing to restrict your potential profit-making in order to provide capital for “green” companies, there is a boutique industry which allows you to do so; if you put more emphasis on the behavior of the end recipients of the money, you may pay more for a company whose published code of ethics (and record of service in line with that code) meets your criteria. If market participants fail to signal with their money that such things are important, the industry will adapt accordingly.

An example of this mechanism is dolphin-safe tuna labeling. Few people would call themselves advocates of wanton dolphin suffocation – it was simply a by-product of the most efficient way of netting tuna for commercial sale. It is not unreasonable to assume that most people, given a choice between two brands of tuna totally alike in all ways but this one, would in fact prefer tuna that did not involve killing dolphins. However, the pursuit of a dolphin-safe method of fishing means higher costs for the fisherman. How was this externality of dolphin safety handled?

In this case, it was a combination of industry self-regulation, government intervention, and product labeling so that the informed consumer could choose to pay more per can if they wanted tuna considered “dolphin safe.” Now, all tuna sold at retail in the U.S. needs to meet particular standards of dolphin safety, and consumers have the option of choosing fishing method (which is often proudly displayed on the can; see, e.g., American Tuna) if this is important to them. Since the gathering and posting of information is costly, consumers of both tuna and financial services have an obligation to inform themselves to ensure that the methods of obtaining their services are ethical and then communicate that preference to the market.<sup>7</sup>

## Emergent Behaviors and Systemic Responsibilities

We have now discussed the ethical obligations owed between agents, but what is owed to the market itself? An ecosystem populated with complex autonomous

---

<sup>7</sup>The World Trade Organization ruled against the use of “dolphin-safe” tuna labeling by the U.S., siding with Mexico that such information was limiting Mexico’s access to the U.S. tuna industry (WTO 2011a); a similar ruling was made recently regarding labeling foods in the U.S. with their country of origin (WTO 2011b). The authors encourage the reader to follow these developments in consumers’ rights to information as they unfold.

agents of both natural and manufactured origins begins to show emergent behaviors of its own, such as stability trends, interaction effects, and unforeseen elements. How do you disaggregate in a Coasian fashion an emergent property of the aggregate?

What this case suggests is that, beyond ethical obligations to clients, consumers, and professional standards, participants in the financial industry have an obligation to the market itself to own their share of potential externality. Unlike professional error in the hard sciences, such as an engineer's fallen bridge or a doctor's collapsed patient, the sins of the social scientist or economist can often be quietly aggregated into the very matter of study. This dispersion weakens the feedback loop used to monitor and enforce ethical codes; however, it does not alter the requirement of stability within certain bands to exist for the market (or other subject) to function at all. It is crucial to remember that, despite the recovery of the market in the Flash Crash, this was not an example of market self-regulation and recovery, but rather the delayed and fortunate intervention of both human and human-designed elements which, in their original conception, should have kicked in significantly sooner with significantly greater effect across the markets.

There is legal precedent for this particular approach to market responsibilities. In 2010, two Norwegian traders were found guilty of "market manipulation" crimes for trades made from Oslo in 2007–2008; they discovered that a machine working for an American trading platform would react in a particular way to a particular trade pattern, so they conducted trades that made the machine drive up the share price and give them the profits (Ward 2010). In financial circles, this has caused considerable discussion since the convicted traders were not breaking laws per se by exploiting a discovered weakness in the algorithm. So what was the crime, especially since, in the words of an individual whose firm designs such algorithms: "[i]t's not always easy to guess what an algo[rithm] is doing otherwise a lot of people would be doing it" (Ward and Grant 2010).<sup>8</sup>

The crime, as the Norway court stated, is against the market itself. Regardless of whether any trading regulations were broken or any person endured financial harms, the violation of the system is considered criminal. There is a point of complexity where the causal cloth cannot be separated into strands without losing some of the sensor-effector relationships which have emerged. When this occurs, systemic properties incur systemic ethical obligations. This is not as unfamiliar as you may think: Citizens of the United States both pledge allegiance to and can be executed for treason against the United States as a collective entity (US Constitution, Article 3, Section 3). In addition to protecting the agency of another consumer or the professional standards of your industry group, there is an ethical obligation to the

---

<sup>8</sup>As of this writing, the Norwegian day traders won their appeal (3–2) and had their sentences suspended. The ruling was based on their actions being public and in line with traditional industry behavior ("Norwegian court acquits day traders in algo cracking case." 2012). Unfortunately, industry-specific ethical relativism may not be the path which offers the most stability or long-term benefit for all market participants.

healthy functioning of the ecosystem which contains the financial markets akin to the obligations engineers have to the sustainability of the actual ecosystem. Even if no one suffers financial calamity, the release of an untrained or misunderstood autonomous system undermines the foundations on which the financial markets rest; even though the most damaging of the trades made during the Flash Crash were reversed, this does not eradicate the ethical transgressions made by the agents of various roles active in the market that day. Whether we're eradicating the rats on Guam by introducing snakes or making ourselves millionaires by letting computers execute our algorithms, the ecosystem will endure changes for which those involved will bear responsibility.

## 2.6 Conclusion

The three cases and associated analyses presented above are meant to give the reader a broad understanding of the types of risks and issues that can occur in these new worlds where sensors, effectors, and systems constructed out of communications and computational resources grow more and more pervasive and autonomous throughout our environment.

Pervasive sensors have invaded our virtual and real worlds, opening up new vistas for those who want to track our every move and click. Data-mining algorithms scan these vast streams of data, assembling profiles of our behavior that can be used without our consent to shape the information that we receive on a daily basis (e.g., your search results and the ads you see across the web). Currently, it is advertising that is being targeted at us, but there is nothing preventing news and other information from being similarly filtered and packaged for your specific consumption. The designers and utilizers of these technologies have a responsibility to the public to gather their data with proper consent for the types of information gathered as well as for the ways in which it is utilized. These responsibilities of the builders and fielders of this technology do not absolve we as users from our responsibilities to be aware of the systems our data flows in or the surveillance we are constantly under; we should demand the rightful ownership of our information and insist that its purveyors use it responsibly and guard it zealously against others who may try to steal it or use against us.

Pervasive effectors simplify our lives and have the capability of making our lives safer and more comfortable, but they also magnify the risks we face as we hand over more and more control of the physical world to these systems. A future where computer viruses do more than destroy your data on your computer is already here. Viruses like Stuxnet and its progeny are the first forays into what will likely prove to be a computational arms race between nation states, corporations, and individuals (e.g., the "hacktivist" group Anonymous), where the stakes transcend the virtual world and the harms caused include damage to persons and property. The dangers here lie also in the inadvertent exposure of the workings of these systems to other malefactors as well. The designers and implementers of these systems must not

only worry that their systems can do intentional harm but that even the most benign system may be used by hackers or cyber-terrorists to cause mayhem in unanticipated ways; as these systems become more complex, these potentially harmful paths increase exponentially in number.

Pervasive autonomous systems bring yet another level of risk and ethical concern to the designers, implementers, and users of those systems. Their autonomous nature coupled with their computational speed make these kinds of systems capable of doing enormous damage before human users or managers can even notice something is going wrong. The ethical rules and reasoning in these systems will need to be built into their command modules in order for any moral consideration to be taken before the decisions are executed. Implementing this form of artificial moral agency is not a trivial task (nor a non-controversial one) and is a field that is in its infancy. Without these necessary safeguards and constraints on the designers (or at minimum a strict testing regimen in simulated real-world environments), the risks of a system failure orders of magnitude greater than the Flash Crash are great. When you combine these types of automated systems with effectors that can affect the real-world, the damage will be far more than just a few hundred points on a stock market index.

These new technologies are not just changing our world – they are also changing us. The changes in our behaviors that come about as we accept and use these systems can offer new opportunities for intentional and unintentional harm. The responsibilities for the changes in behavior, from becoming too reliant on the technology to not understanding the limitations of the systems we use, fall upon designer and user alike. It is a brave new world, in which these designers and users need to rethink some fundamental assumptions that have served us well in the use of our technical systems in the past. The worlds we inhabit and we ourselves are changing, and it is not surprising that new rules and new roles will have to be developed and understood for all those involved: designers, builders, and users.

## References

- 2009. 11-year-old boy dies after mom says GPS left them stranded in Death Valley. Associated Press. August 8, 2009. <http://www.foxnews.com/story/0,2933,538323,00.html>. Accessed online 13 Apr 2012.
- Adams, J., A. Tashchian, and T. Shore. 2001. Codes of ethics as signals for ethical behavior. *Journal of Business Ethics* 29(3): 199–211. <http://www.jstor.org/stable/25074455>.
- Angel, J., and D. McCabe. 2010. Ethical standards for stockbrokers: Fiduciary or suitability? September 30, 2010. <http://ssrn.com/abstract=1686756>. Accessed online 13 Oct 2011.
- Arthur, C. 2012. Cyber-attack concerns raised over Boeing 787 chip's 'back door'. *The Guardian*, May 29, 2012. <http://www.guardian.co.uk/technology/2012/may/29/cyber-attack-concerns-boeing-chip>. Accessed online 29 May 2012.
- Bamford, J. 2012. The NSA is building the country's biggest spy center (Watch what you say). *Wired Magazine*. <http://www.wired.com/threatlevel/2012/03/ff-nsadatacenter/all/1>. Accessed online 13 Apr 2012.
- Battaglio, R., and T. Loughran. 2008. Does payment for order flow to your broker help or hurt you? *Journal of Business Ethics* 80(1): 37–44. doi:10.1007/s10551-007-9445-x.



- Bell, A. 2010. FINRA sticks with securities in suitability proposal. *National Underwriter*, August 19, 2010. <http://www.lifeandhealthinsurancenews.com/News/2010/8/Pages/FINRA-Sticks-With-Securities-in-Suitability-Proposal.aspx?page=1>. Accessed online 13 Oct 2011.
- Brennan, T. 2010. Emergency beacons can be beneficial, but can cause problems for rescuers. *KBTX*, September 16, 2010. <http://www.ktvb.com/home/Locating-beacons-provide-benefits-to-users-headaches-for-others-103099444.html>. Accessed online 25 May 2012.
- Brenner, B. 2012. Analyst: Duqu is all a bunch of hype. *CSO*, December 2011/January 2012, 13.
- Clark, K. 2011. *The GPS: A fatally misleading travel companion*, July 26, 2011. <http://www.npr.org/2011/07/26/137646147/the-gps-a-fatally-misleading-travel-companion>. Accessed online 15 Apr 2012.
- Cliff, D. and L. Northrop. 2011. *The global financial markets: An ultra-large-scale systems perspective*, December 23, 2011. <http://www.bis.gov.uk/assets/foresight/docs/computer-trading/11-1223-dr4-global-financial-markets-systems-perspective>. Accessed online on 1 May 2012.
- Coase, R.H. 1960. The problem of social cost. *Journal of Law and Economics* 3: 1–44. <http://www.jstor.org/stable/724810>.
- Denning, T., C. Matuszek, K. Koscher, J. Smith, and T. Kohno. 2009. *A spotlight on security and privacy risks with future household robots: Attacks and lessons*. Ubicomp '09- Proceedings of the 11th international conference on ubiquitous computing, 105–114. Orlando, FL, USA.
- Farwell, J., and R. Rohozinski. 2011. Stuxnet and the future of cyber war. *Survival: Global Politics and Strategy* 53(1): 23–40.
- Fein, M. 2010. *Brokers and investment advisers standards of conduct: Suitability vs. fiduciary duty*, August 31, 2010. <http://ssrn.com/abstract=1682089>
- Gross, M. 2011. A declaration of cyber-war. *Vanity Fair*. <http://www.vanityfair.com/culture/features/2011/04/stuxnet-201104>. Accessed online 13 Apr 2012.
- Halperin, D., T. Kohno, T. Heydt-Benjamin, K. Fu, and W. Maisel. 2008. Security and privacy for implantable medical devices. *IEEE Pervasive Computing* 7: 30–39.
- Herbert, G. 2012. Microsoft 'avoid ghetto' app for smartphones' GPS navigation stirs controversy. *The Post Standard*, January 11, 2012. Syracuse. [http://www.syracuse.com/news/index.ssf/2012/01/microsoft\\_avoid\\_ghetto\\_app\\_windows\\_phone.html](http://www.syracuse.com/news/index.ssf/2012/01/microsoft_avoid_ghetto_app_windows_phone.html). Accessed online 18 Apr 2012.
- Hollis, D. 2011. Cyberwar case study: Georgia 2008. *Small Wars Journal*, January 6, 2011. <http://smallwarsjournal.com/blog/journal/docs-temp/639-hollis.pdf>. Accessed online on 28 May 2012.
- IAEA Board of Governors. 2009. *Implementation of the NPT safeguards agreement and relevant provisions of security council resolutions 1737 (2006), 1747 (2007), 1803 (2008) and 1835 (2008) in the Islamic Republic of Iran*, February 19, 2009. <http://www.iaea.org/Publications/Documents/Board/2009/gov2009-8.pdf>. Accessed online 13 Apr 2012.
- Koscher, K., A. Czeskis, F. Roessner, S. Patel, T. Kohno, S. Checkoway, et al. 2010. *Experimental security analysis of a modern automobile*. Paper presented at the Security and Privacy (SP), 2010 IEEE symposium on security and privacy.
- Koscher, K., A. Czeskis, F. Roessner, S. Patel, and T. Kohno. 2012. *Experimental security analysis of a modern automobile*. Paper presented at the 2010 IEEE symposium on security and privacy, Oakland, 16–19 May 2012.
- Norwegian court acquits day traders in algo cracking case. 2012. *Finextra Research*, May 2, 2012. <http://www.finextra.com/news/fullstory.aspx?newsitemid=23677>. Accessed online 01 June 2012.
- Sanger, D.E. 2012. Obama order sped up wave of cyberattacks against Iran. 2012. *NY Times*, June 1, 2012. <http://www.nytimes.com/2012/06/01/world/middleeast/obama-ordered-wave-of-cyberattacks-against-iran.html>. Accessed online 21 June 2012.
- Schwartz, M. 2001. The nature of the relationship between corporate codes of ethics and behaviour. *Journal of Business Ethics* 32(3): 247–262. doi:10.1023/A:1010787607771.
- Searing, D. 2012. *A creator's ethics- beyond engineering ethics- a first step*. Paper presented at the twenty-first annual meeting of the association for practical and professional ethics, Cincinnati, 1–3 Mar 2012.

- Searing, D. and E. Searing. 2011. *Playing the market good or well?: The ethical implications of the 'flash crash'*. Case study presented at the 12th annual meeting of the association for practical and professional ethics, Cincinnati, 3–15 Mar 2011.
- Serchuk, D. 2009. Suitability: Where brokers fail. *Forbes.com.*, June 24, 2009. <http://www.forbes.com/2009/06/23/suitability-standards-fiduciary-intelligent-investing-brokers.html>. Accessed online 13 Oct 2001.
- Smith, A. 1991. *The wealth of nations*. New York: Alfred A. Knopf.
- Tremlett, G. 2010. GPS directs driver to death in Spain's largest reservoir. *The Guardian*, October 4, 2010. <http://www.guardian.co.uk/world/2010/oct/04/gps-driver-death-spanish-reservoir>. Accessed online 15 Apr 2012.
- U.S. Commodity Futures Trading Commission (CFTC) and U.S. Securities & Exchange Commission (SEC). 2010. *Findings regarding the market events of May 6, 2010: Report of the staffs of the CFTC And SEC to the joint advisory committee on emerging regulatory issues*, 30 Sept 2010. <http://www.sec.gov/news/studies/2010/marketevents-report.pdf>. Accessed online 13 Oct 2010.
- US Const., Art. 3, sec. 3. <http://memory.loc.gov/cgi-bin/query/r?ammem/bdsdcc:@field%28DOCID+@lit%28bdsdccc0801%29%29>. Accessed 14 Apr 2012.
- Ward, A. 2010. Norwegians convicted for outwitting trading system. *Financial Times*, October 13, 2010, 7:17 pm. <http://www.ft.com/intl/cms/s/0/f9d1a74a-d6f3-11df-aaab-00144feabdc0.html#axzz1sAPjPv00>. Accessed 14 Apr 2012.
- Ward, A. and J. Grant. 2010. A tale of man versus algo in Norway. *Financial Times*, October 14, 2010, 10:27 pm. <http://www.ft.com/intl/cms/s/0/9a4aa5b8-d7bd-11df-b478-00144feabdc0.html#axzz1sAPjPv00>. Accessed 14 Apr 2012.
- World Trade Organization. 2011a. United States – Measures Concerning The Importation, Marketing and Sale of Tuna and Tuna Products: Report of the Panel.” Case Number 11-4239. [http://www.wto.org/english/tratop\\_e/dispu\\_e/381r\\_e.pdf](http://www.wto.org/english/tratop_e/dispu_e/381r_e.pdf). Accessed on 14 Apr 2012.
- World Trade Organization. 2011b. United States – Certain Country of Origin Labelling (Cool) Requirements: Reports of the Panel. Case Number 11-5865. [http://www.wto.org/english/tratop\\_e/dispu\\_e/384\\_386r\\_e.pdf](http://www.wto.org/english/tratop_e/dispu_e/384_386r_e.pdf). Accessed on 14 Apr 2012.

Emerging Pervasive Information and Communication  
Technologies (PICT)

Ethical Challenges, Opportunities and Safeguards

Pimple, K.D. (Ed.)

2014, X, 252 p.,

ISBN: 978-94-007-6833-8