12-2012

Moral Responsibility for Computing Artifacts: "The Rules" and Issues of Trust

Frances Grodzinsky
Sacred Heart University

Keith Miller
University of Illinois at Springfield

Marty J. Wolf
Bemidji State University

Recommended Citation

Grodzinsky, F., Miller, K., & Wolf, M. J. (2012). Moral responsibility for computing artifacts. *ACM SIGCAS Computers and Society, 42*(2),15-25. Doi: 10.1145/2422509.2422511

Moral Responsibility for Computing Artifacts: "The Rules" and Issues of Trust

FS Grodzinsky
Sacred Heart University, Fairfield, CT USA
grodzinskyf@sacredheart.edu

K Miller
University of Illinois Springfield, Springfield, IL USA
miller.keith@uis.edu

MJ Wolf
Bemidji State University, Bemidji, MN, USA
mjwolf@bemidjistate.edu

Abstract
"The Rules" are found in a collaborative document (started in March 2010) that states principles for responsibility when a computer artifact is designed, developed and deployed into a sociotechnical system. At this writing, over 50 people from nine countries have signed onto The Rules (Ad Hoc Committee, 2010). Unlike codes of ethics, The Rules are not tied to any organization, and computer users as well as computing professionals are invited to sign onto The Rules. The emphasis in The Rules is that both users and professionals have responsibilities in the *production and use* of computing artifacts. In this paper, we use The Rules to examine issues of trust.

Introduction
"The Rules" are found in a collaborative document (started in March 2010) that states principles for responsibility when a computer artifact is designed, developed and deployed into a sociotechnical system. At this writing, over 50 people (including the authors of this paper) from nine countries have signed onto The Rules (Ad Hoc Committee, 2010). Unlike codes of ethics, The Rules are not tied to any organization, and computer users as well as computing professionals are invited to sign onto The Rules.

Based on the theories of (Floridi and Sanders, 2001) and (Floridi, 2008), we have used levels of abstraction to examine ethical issues created by computing technology (see (Grodzinsky et al., 2008) and (Wolf et al., 2011)). In those analyses, we used three levels of abstraction: LoA1, the users' perspective; LoA2, the developers' perspective; and LoAS, the perspective of society at large. Our analysis of quantum computing and cloud computing, focused on computing professionals at LoA2, delivering functionality to users at LoA1 (Wolf et al., 2011). Our emphasis was on the professionals being worthy of the trust of users in that delivery.

Our analysis of The Rules in this paper differs starkly from the earlier analyses of quantum and cloud computing. The Rules are not a computing paradigm; they are a paradigm for thinking about the impact of computing artifacts. The emphasis in The Rules is different from a technical computing project: both users and professionals are invited to acknowledge their responsibilities in the *production and use* of computing artifacts. Yet there are some aspects of the earlier analyses, especially in the area of trust, that are relevant to The Rules. In quantum computing, the implementers of quantum algorithms will likely not meet most of the users of those algorithms, and although we don't anticipate that communications are likely between individual users and the quantum developers, the trust relationship will be forged through the medium of the quantum algorithms. Inherent in the design of cloud computing is the notion that the people who maintain the computing resources of the cloud are remote from the users of those resources. Humans are clearly crucial in the sociotechnical systems of cloud computing. But most of the relationships will be based on e-trust, not on face-to-face interactions. Trust issues are more complex in these new computing paradigms, and it is our assertion that The Rules can inform a discussion of these issues.

The first part of this paper presents The Rules. The Rules document currently includes five rules that are intended to serve "as a normative guide for people who design, develop, deploy, evaluate or use computing artifacts." Next we briefly examine a model of trust and the relationship between The Rules and society through the lens of trust. In other words, we will examine how computing artifacts and the sociotechnical system of which they are a part, serve as a medium through which trust relationships are played out. Then, we shall examine each rule vis-à-vis the sociotechnical system and trust. The existence and proliferation of computing artifacts and the growing sophistication of sociotechnical systems do not insulate users and developers from the need to trust and the obligation to be trustworthy. Instead, we are convinced that the power and complexity of these systems require us to be more dependent on trust relationships, not less. In the last section of the paper we illustrate this last statement by applying the Rules to the paradigms of quantum and cloud computing.

What Are The Rules?

"The Rules" began at a face-to-face meeting early in 2010, but many of the people who have now signed on to The Rules have not met each other face to face. The Rules has become a project explicitly based on Internet interactions. This aspect of The Rules is something that we often see in analyzing an ethical issue involving sociotechnical systems: there are people who are critical components in any such system, and often some of those people meet face to face. But sometimes, and increasingly often, after a computing artifact is "launched," many (and sometimes most) of the ensuing interactions and relationships are computer-mediated.

The five rules in version 27 are:

Rule 1: The people who design, develop, or deploy a computing artifact are morally responsible for that artifact, and for the foreseeable effects of that artifact. This responsibility is shared with other people who design, develop, deploy or knowingly use the artifact as part of a sociotechnical system.

Rule 2: The shared responsibility of computing artifacts is not a zero-sum game. The responsibility of an individual is not reduced simply because more people become involved in designing, developing, deploying or using the artifact. Instead, a person's responsibility includes being answerable for the behaviors of the artifact and for the artifact's effects after deployment, to the degree to which these effects are reasonably foreseeable by that person.

Rule 3: People who knowingly use a particular computing artifact are morally responsible for that use.

Rule 4: People who knowingly design, develop, deploy, or use a computing artifact can do so responsibly only when they make a reasonable effort to take into account the sociotechnical systems in which the artifact is embedded.

Rule 5: People who design, develop, deploy, promote, or evaluate a computing artifact should not explicitly or implicitly deceive users about the artifact or its foreseeable effects, or about the sociotechnical systems in which the artifact is embedded.

The Rules echo a theme that was developed in our earlier analyses: those at LoA2 owe a special duty of care to users at LoA1 (see (Grodzinsky et al. 2008) and (Wolf et al. 2011). Those who produce and sell computer artifacts are to act in a manner that is worthy of trust. But unlike the previous analyses, The Rules require that both computer professionals *and users* be trustworthy to society as a whole, reinforcing the inclusion of LoAS in the analysis of computer ethics.

A Model of Trust

We will use the following principles of trust developed by (Taddeo, 2008) and (Taddeo, 2009) and the model of trust in (Grodzinsky et al., 2011) (see the next section) as the basis for our discussion of trust relationships and The Rules.

1. **Trust is a relation between *a* (the *trustor*) and *b* (the *trustee*).** NOTE: ***a*** and ***b*** can be human or an artificial agent (AA). A relation (certainly in the mathematical sense, but also in the sociological sense) can involve both.

2. **Trust is a decision by *a* to delegate to *b* some aspect of importance to *a* in achieving a goal.** NOTE: We rely on the notion that an artificial entity ***a*** includes "decisions" (implemented by, for example, IF/THEN/ELSE statements), and we assume that ***a***'s decisions are designed and implemented with the assumption that there is a high probability that ***b*** will behave as expected.

3. **Trust involves risk; the less information the trustor *a* has about the trustee *b*, the higher the risk and the more trust is required.** NOTE: this is true for both artificial and human entities. In AAs, we expect that risk and trust are quantified or at least categorized explicitly; in humans, we do not expect that this proportionality is measured with mathematical precision.

4. **The trustor *a* has the expectation of gain by trusting the trustee *b*.** NOTE: With respect to AAs, "expectation of gain" may refer to the expectation of the AA's designer in moving toward a particular goal, or it may refer to an explicit expression in the source code that identifies this expected gain, or both.

5. **The trustee *b* may or may not be aware that trustor *a* trusts *b*.** NOTE: If ***b*** is human, circumstances may have prevented ***b*** from knowing that ***a*** trusts ***b***. The same is true if ***b*** is an AA, but there is also some possibility that an AA trustee ***b*** may not even be capable of "knowing" anything in the traditional human sense.

6. **Positive outcomes when *a* trusts *b* encourage *a* to continue trusting *b*.** NOTE: If ***a*** is an AA, this cycle of *trust → good outcome → more trust* could be explicit in the design and implementation of the AA, or else it could be implicit in data relationships, as in a neural net.

We contend that these principles are attributes that belong to any type of trust, be it e-trust or face-to-face (f2f)-trust.

An Object Oriented Model of Trust

In a previous paper, we developed a model of trust based on object oriented software development principles (Grodzinsky et al., 2011). The superclass of the trust model is"XYZtrust" which represents what is common to eight subclasses of trust. In each of these eight subclasses, we can locate a set of possible trust relationships. The letters X, Y, and Z in XYZtrust stand for important aspects of the trust relationship. X and Y classify the trustor and trustee respectively as either human (H) or artificial (A); and Z specifies if the trust is established either face-to-face (P, or "physical") or electronic (E). These three binary variables establish the eight possible subclasses of XYZtrust, and each subclass is more concrete than the more abstract superclass XYZtrust. Each subclass is still an abstract description analogous to a template, and that template represents many possible instantiations of trust relationships. A list follows that enumerates the eight subclasses, along with an example of a trust relationship that fits into that subclass.

HH-Ptrust: traditional notion of human, "face-to-face" trust; most discussions of "trust" presuppose this subclass

HH-Etrust: humans trust each other, but mediated exclusively by electronic means

HA-Ptrust: human trusts a physically present AA, for example, a robot (no electronic mediation)

HA-Etrust: human trusts an artificial entity (for example, a web bot) over the Internet

AH-Ptrust: an AA, perhaps a robot, trusts a physically present human

AH-Etrust: an AA, perhaps a web bot, trusts a human based on Internet interactions

AA-Ptrust: an AA trusts another AA in a physical encounter; because this is Ptrust, the AAs might, for example, use sign language

AA-Etrust: an AA trusts another AA electronically, e.g., two web bots communicate via the Internet

Each instance of an XYZtrust subclass is either all P or all E. Thus if two particular humans, participate in HH-Etrust and also participate in HH-Ptrust, then we intend that these instances of trust be modeled with two separate instantiations, one in HH-Etrust and the other in HH-Ptrust.

HHP	HHE
HAP	HAE
AHP	AHE
AAP	AAE

Table 1: Eight distinct subclasses of XYZ Trust

Table 1 lists the subclasses of XYZtrust. In this arrangement, the first column requires physical proximity, and the second column is strictly electronic. In this paper, our analysis of The Rules focuses on the types of trust relationships found in the first row. Although the other rows (each of which involves at least one AA) have relevance to computing professionals (see Grodzinsky, et al, 2011), the Rules emphasize the human relationships at stake in sociotechnical systems that involve computing.

The Rules, Trust and Society

The Rules are an attempt to change people's minds about moral responsibility for computing artifacts. The people who have signed The Rules thus far are people aware of the computing details known by developers. However, if The Rules are to have a wider impact, that impact will be observable at LoAS. Whenever computing professionals take seriously their collective and individual responsibilities for computing artifacts, they act in ways that enhance public trust in computing. Insofar as The Rules successfully encourage computing professionals to act this way; The Rules will enhance public trust.

But The Rules can also be seen as a challenge to computing professionals, a call to account for the artifacts they design, develop and deploy. Seen in this way, The Rules could be interpreted by some as a reason for society as a whole to trust computing professionals less, especially if computing professionals resist acknowledging the responsibilities outlined in The Rules. The extent to which The Rules have either of these effects on society (or any discernible effect at all) can best be examined (either theoretically or empirically) at LoAS.

Although computer users may often act as if they are trusting machines, The Rules insist that the people who design, develop and deploy these machines are directly responsible for the machines and their effects. In addition, The Rules state that people who use these machines are responsible for how they are used. We can restate that emphasis in the following way: those who design, develop and deploy computing artifacts are placed in a relationship of trust with those who use the artifacts. The computing artifacts and the sociotechnical system that the artifacts are embedded in act as a medium through which a trust relationship is played out between people.

The Rules also state that users are responsible for the consequences of their use of computing artifacts. In this relationship, users are in a trust relationship with people who are affected (or potentially affected) by the use of a computing artifact (the "penumbra"). Again, the computing artifact and the sociotechnical system of which it is a part serve as a medium through which trust relationships are played out among people.

Rule 1: The Major Players in Computing Trust

Rule 1 identifies two groups of people: those who create a computing artifact and those who use a computing artifact in a sociotechnical system: the creators of the artifact (designers and developers) and the users (who deploy the artifact). Thus, there are three trust relationships that must be considered: trust among the designers and developers themselves; trust between users and creators; and trust between users and the penumbra. Both the group and individuals in the group that create computing artifacts must take on moral responsibility for the artifact. Individuals must trust that other team members in the group are willing to accept the moral responsibility that Rule 1 calls for. That is, there is an accepted goal among the creators of the artifact to examine the effects of that artifact on society and to perform their functions with the appropriate standard of care. This is a familiar theme in professional ethics and addresses the problem of "many hands" that Helen Nissenbaum cites as a problem of computing accountability (Nissenbaum, 2007).

In the second trust relationship, between the users and the creators of the artifact, the users who trust developers will buy their products and use them with confidence.

In the third relationship, it is less essential that those who use a particular artifact in a particular sociotechnical system develop any sort of trust with those in the penumbra who use that particular artifact in the same or any other sociotechnical system. Nor is there any strong need for the development of trust among people

employing different computing artifacts in the same sociotechnical system. More essential is the response of the users and creators when it is evident at LoAS that the artifact is causing an unanticipated problem.

Rule 2: Many Hands and Trust

Rule 2 emphasizes the need for a high degree of trust between users of a computing artifact and the creators of a computing artifact. It is intuitive that the users of a computing artifact will need to trust the creators of the artifact. The users trust that the artifact meets its specification. That is, the artifact should do what it is supposed to do, and using the artifact in its intended way does not result in unintended negative consequences. Ironically, most software developers, both proprietary and free and open source developers, disclaim all warranties for software, even when the software is used for its intended purpose.

But this user → creator trust has an inverse. Creators also have trust issues with their users in addition to the ones cited in Rule 1. For example, developers trust (or at least hope) that users will honor the relevant licensing agreements. (Notice that this is true for both proprietary software and for Free and Open Source software.)

Furthermore, developers need to trust that users will not use the computer artifact in bizarre, unanticipated ways. With the moral responsibility that Rule 1 brings to bear on the artifact creators for the foreseeable effects of the use of the artifact, the developers need to trust that users will not use the artifact in some unforeseeable, negative or potentially dangerous way. Such use may bring needless risk to society. Honest statements by the creators of an artifact regarding the intended use of the artifact build should be honored by users to further the trust that both groups have in each other. A user, who intends to use an artifact for an unintended purpose, ought to seek advice from its creator regarding the new, unanticipated use, in order to ascertain whether the intended use is consistent with the moral responsibilities outlined by The Rules.

One concern regarding using a computing artifact for an unintended purpose is the damage such use might bring to the fabric of trust that has been built between computing artifact developers and society. An artifact user bears the responsibility of weighing the potential damage any unintended use of the artifact may bring to this important trust relationship.

Rule 3: Users, Trust, and Society

Rule 3 mitigates some of the concerns regarding trust and the unintended use of a computing artifact. However, even though according to The Rules, the knowing user of the computing artifact bears moral responsibility, we are still concerned about the damage to the developer/society trust relationship such use would bring. Users also have a wider responsibility towards society. The power of a computing artifact brings with it the responsibility to use that power in a way that respects other people. Thus, users of a computing artifact must act in a way that is trustworthy with respect to society; this trustworthiness should be observable at LoAS.

Rule 4: Sociotechnical Systems and Trust

Computer artifacts are "connections" among people; the emphasis in Rule 4 widens this focus. Sociotechnical systems include computing artifacts, but also encompass customs, laws, regulations, cultural norms, infrastructure, and any other factors that make the development and use of the artifacts possible. But developers and users are responsible for their choices to participate in these sociotechnical systems. Society expects these systems to benefit society, and trusts both computing professionals, as well as other people who participate in these systems, to develop and use them responsibly.

Joseph Weizenbaum (1984) and Bill Joy (2000) both argued that developers have a choice in the kinds of computer science research to undertake and the kinds of artifacts to develop. They both suggest that there may be certain areas of research that we should be careful exploring, or at least, we should be honest about the limitations of the artifact because often we cannot predict the effects of the artifact on society. Our increasing societal reliance on socio-technical systems, the enormous reach and impact of these systems, and the accelerating rate of changes in these systems, make this choice a difficult one for computing professionals. The difficulty does not absolve computing professionals from their responsibility to make their choices deliberately and wisely.

Rule 5: Truth-Telling and Trust

Issac Asimov is given credit for an interesting quote about humans and computing artifacts: "Part of the inhumanity of the computer is that, once it is competently programmed and working smoothly, it is completely honest." (Asimov)

Trust and honesty are deeply linked. Rule 5's statements about not deceiving others go a long way toward building trust. As honest statements about computing artifacts are made, and it becomes clear that the statements are true when the artifacts are deployed (observed at LoAS), society builds trust in the claims of artifact creators. Overstatements or a lack of candor about the weaknesses or dangers of particular artifacts and sociotechnical systems will erode that trust. For example, when an Intel processing chip had a hardware error that occurred in certain kinds of division, the company tried to stonewall and said that it had an infinitesimal probability of occurring, and it would be expensive to fix this error. Under pressure from engineers who feared that these chips would not be reliable—imagine them in an avionics system where deaths could occur—Intel eventually relented. Trust in the company was restored when Intel owned up to the error, redid the masks and issued substitute chips at no cost to the users.

While we acknowledge that no software is error free, we observe a lack of trust in companies such as Apple or Microsoft at LoAS when users refuse to buy first generation products. Consumers continue to buy products after software failures, but this is often seen as a necessary evil, not a reflection of trust. Lawsuits about software failures reflect this lack of trust. (For example, see (Weier, 2008), (King, 2010), and (Vijayan, 2010).)

WikiLeaks: A Case Study of Electronic Trust (and Distrust)

In 2010, the website WikiLeaks released numerous documents classified as secret to by the United States government. As the ramifications of these documents became apparent, numerous financial companies broke their ties with WikiLeaks, depriving it of its main source of funding. In retaliation, WikiLeaks supporters posted distributed denial of service (DDoS) software on Facebook and encouraged people to download that software and install it on their computers. Once installed, the originators of the software used those computers to send disruptive messages to retaliate against the financial institution's computers, making their websites almost completely inaccessible by anyone intending to conduct business with the financial institutions.

There are numerous entities that come into consideration in this story including WikiLeaks, the developers of the DDoS software, the developers of Facebook, the sociotechnical system created by Facebook, those who installed the DDoS software on their computers, the users of the financial institutions websites, and the financial institutions themselves. Each of these entities is part of some trust relationship and serves to illuminate the points we have made.

Prior to the WikiLeaks incident, DDoS developers had done much to damage the trust between software developers and users of software. DDoS software was typically installed on a "user's" computer without the user's knowledge and then used by the developer to cause harm to a third party. This sort of misappropriation of a user's computer led to mistrust of the developer of the operating system. Rule 3 is clear about laying responsibility for the damage at the feet of those who deployed the DDoS since the user rarely knows about the DDoS software existing on the user's computer.

In the WikiLeaks incident, DDoS developers provided like-minded people the opportunity to participate in an apparent act of civil disobedience. In this context, however, users of the DDoS software chose to install the software on their computers. Rule 3 in this case, applies to the users, since they knowingly installed the software and then knowingly relinquished control of it to those who launched the software. In this context, the user trusts the developer to accomplish a goal for the user. Thus, the same sort of software, depending on the sociotechnical context, facilitates and erodes trust between developers and users.

When considered at the LoAS, both uses of DDoS do little to facilitate trust between software developers and society. The users of the DDoS software have damaged the ability of people to access their financial information at the targeted financial institutions. General trust in software developers and the systems they have developed is eroded. Further, since many users of financial institution's websites may not understand that the DDoS is taking place, they may begin to lose trust in the financial institutions themselves. Another concern surrounding this incident is that unlike traditional acts of civil disobedience, the participants in the DdoS—both those who allowed their computers to be used and those who orchestrated the attack—did little to make their identities known.

What If We Don't Trust? Applying The Rules to Cloud and Quantum Computing

We asserted in our introduction that the existence and proliferation of computing artifacts and the growing sophistication of sociotechnical systems do not somehow insulate us against the need to trust and the obligation to be trustworthy. Instead, we are convinced that the power and complexity of these systems require us to be *more* dependent on trust relationships, not less. This becomes clear when we consider trust in new and emerging technologies. We analyze The Rules and trust as they apply to the relatively new cloud computing and then to the emerging field of quantum computing.

Cloud Computing

Cloud computing is a term that encompasses many different technologies and sociotechnical systems. However, most of these systems feature computing resources (both hardware and software) that are remote from the users of the resources. Users access these resources via Internet connections and rely on the providers of the resources to make the resources accessible and secure.

In most cases, the users and the providers of cloud computing will not physically meet. The relationships between users and providers are, then, of the electronically mediated type described above. The emphasis in The Rules on responsibility and honesty seem particularly appropriate to these HHE relationships. First, cloud computing users must trust providers to take care of their data and software; this is fundamental to the implicit social contract of cloud computing. Providers who are careless or cynically indifferent to issues of accessibility and security are clearly not living up to the responsibilities described in The Rules.

Similarly, if concerns about the security and accessibility (both trust issues) discourage some from exploiting cloud computing, and if cloud computing gives its users a competitive advantage (as its

advocates claim), then these trust issues will have economic repercussions. If those who trust and commit to cloud computing are disappointed by a serious breach of security or a lack of accessibility at an urgent moment, then their trust will have been betrayed.

The Rules suggest another, perhaps less obvious, responsibility relevant to cloud computing: users of cloud computing should not exploit computing resources from the cloud to do harm to others. For example, it would be a clear violation of The Rules for someone to sign up for a particular cloud computing service, and then use that access to steal data or disrupt the resources for other customers of that service. Both providers and users are called to account by The Rules for responsible behavior.

Quantum Computing

Quantum computing is a relatively new field. The fundamental nature of quantum computation has been known for over thirty years, yet physicists are still trying to develop quantum computers that are of any significant size. One significant difference between quantum computers and digital computers is that quantum computers are probabilistic in nature. That is, there are times that the output is not a correct solution to the problem for the given inputs. Another thing that is known about quantum computers is that a quantum computer of a given size will be significantly more powerful than digital computers of the same size. It is important to note that there are small quantum computers available in the marketplace for highly specialized applications. Trust and The Rules will play an important role as bigger quantum computers are developed and they become commercially available.

Rule 1 ascribes moral responsibility to the creators of quantum computers for their foreseeable effects. Given the potential power of this technology, quantum developers would be well-served by developing trust among potential users and in society regarding these artifacts. As a nascent technology, developers will have an economic interest in developing a user base. That user base needs to trust the claims of the developers and society needs to trust that developers and users of quantum computers will bear responsibility for their general use. Computer and information ethicists have a role to play here. We can help analyze the impact that this new technology might have.

We have begun this process for quantum computation (Wolf et al., 2011). We identified a potential problem with secure communication when there are those with access to quantum computers. Today's most common cryptographic methodologies are subject to attack by quantum computation. Those with quantum computers will be in a position to decode any encrypted message they capture. The Rules suggest that quantum developers are morally responsible for this situation and the impact it will have on the social and economic systems of the world.

Conclusions

In this paper we have analyzed The Rules and the role trust plays among computing professionals, users and society as a whole. We explored cloud computing and quantum computing as two examples of computing paradigms in which trust and responsibility are key. This work on emerging computing paradigms builds on earlier work by Weckert and Moor who examined how best to proceed in nanotechnology research when there might be adverse effects at LoAS. Using the "precautionary principle" (Weckert and Moor, 2004: 12) assert: "If some action has a possibility of causing harm, then that action should not be undertaken or some measure should be put in its place to minimize or eliminate the potential harms." Their analysis focuses primarily on the relationship of LoA2 and LoAS and examines the various types of harms and risks of artifacts that could cause adverse effects at LoAS. Their claim about the necessity of weighing benefits and harms is supported by The Rules. In addition, using models such as those developed by

(Floridi, 2008) and (Grodzinsky et al., 2008) helps computing professionals, users, and society (LoA2, LoA1 and LoAS) better understand issues of responsibility, risk and trust for particular computing artifacts and sociotechnical systems.

Concerns such as those expressed by Weckert, Moor, Joy and Weizenbaum call into question the kinds of research or development that should be pursued and at what cost. The Rules offer a way of assessing risks so that if an artifact is a large risk for serious harm, development of that artifact can be curtailed or at least more carefully managed. Although The Rules echo the sentiments of computer scientists and philosophers that the burden of proof that an artifact is safe lies with the developer and is present at LoA2, The Rules also advocate that *all* people involved with computing artifacts act in a manner that is trustworthy and responsible. If people don't act responsibly as they produce, promote and use computing artifacts, then The Rules suggest that society should neither trust those people nor their artifacts. We contend that this is true for *all* computing artifacts and for the sociotechnical systems in which they are embedded.

We are convinced that this need for trustworthiness with respect to computing artifacts and sociotechnical systems will be visible at LoAS. Our future research plans include examining empirical evidence at LoAS that demonstrates the importance of trustworthiness from participants at LoA1 and LoA2.

Acknowledgement

A previous version of this paper was presented at ETHICOMP 2012.

References

The Ad Hoc Committee for Responsible Computing (2010), Moral responsibility for computing artifacts: The rules, online at edocs.uis.edu/kmill2/www/TheRules/ accessed 01.03.2011.

Asimov, I. Quotes.net, online at www.quotes.net/quote/145 accessed 01.03.2011.

*Frances Grodzinsky is a professor of Computer Science and Information Technology at Sacred Heart University where she is co-chair of the Hersher Institute of Ethics. She is also a Visiting Scholar, Research Center on Computer Ethics and Social Responsibility, Southern Connecticut State University, New Haven, CT and serves on the board of INSEIT (the **In**ternational **S**ociety for **E**thics and **I**nformation **T**echnology).*

Keith W. Miller is the Schewe Professor of Computer Science at the University of Illinois Springfield. His research areas are computer ethics and software testing, and he is a member of INSEIT, IEEE SSIT, and ACM SIGCAS. Email at kmill217@gmail.com.

Marty J. Wolf is a professor of Computer Science at Bemidji State University. In addition to his computer ethics research, he has also published papers in graph theory, computer science education and the philosophy of information.

Floridi, L. (2008), The method of levels of abstraction. Minds and Machines, 18, 303-329. doi:10.0007/s11023-008-9113-7.

Floridi, L. and Sanders, J.W. (2001), Artificial evil and the foundation of computer ethics. Ethics and Information Technology 3, 55–66.

Grodzinsky, F. S., Miller, K. and Wolf, M. J. (2008), The ethics of designing artificial agents. Ethics and Information Technology, 10, 2-3, DOI: 10.1007/s10676-008-9163-9.

Grodzinsky, F. S., Miller, K. and Wolf, M. J. (2011), Developing artificial agents worthy of trust: "Would you buy a car from this artificial agent?", Ethics and Information Technology, 13:1, 17-27. doi: 10.1007/s10676-010-9255-1.

Joy, B. (2000), Why the future doesn't need us. Wired 8:4.

King, L. (2010), BP oil spill monitoring systems a failure says £14bn US government lawsuit, ComputerworldUK, online at www.computerworlduk.com/news/it-business/3253838/bp-oil-spill-monitoring-systems-a-failure-says-14bn-us-government-lawsuit/ accessed 01.03.2011.

Nissenbaum, H. (2007), Computing and accountability, in J. Weckert, ed. Computer Ethics. Aldershot UK: Ashgate, 273-80. Reprinted from Communications of the ACM (1994), 37, 37-40.

Taddeo, M. (2008), Modeling trust in artificial agents, a first step toward the analysis of e-trust. In Sixth European Conference of Computing and Philosophy, University for Science and Technology, Montpelier, France, 16-18 June.

Taddeo, M. (2009), Defining trust and e-trust: from old theories to new problems. International Journal of Technology and Human Interaction 5, 2, April-June 2009.

Vijayan, J. (2010), Deloitte sued over failed ERP project, ComputerworldUK, online atwww.computerworlduk.com/news/it-business/20575/deloitte-sued-over-failed-erp-project accessed 01.03.2011.

Weckert, J. and Moor, J. (2004), Using the precautionary principle in nanotechnology policy making. Asia Pacific Nanotechnolgy Forum News Journal 3:4, 12-14.

Weier, M. (2008), SAP software a 'complete failure,' lawsuit claims. Information Week, online at www.informationweek.com/news/global-cio/trends/showArticle.jhtml?articleID=207000149 accessed 01.03.2011.

Weizenbaum, J. (1984), Computer Power and Human Reason: From Judgment to Calculation. New York: Penguin Books.

Wolf, M.J., Grodzinsky, F. and Miller, K. (2010), Artificial agents, cloud computing, and quantum computing: Applying Floridi's Method of levels of abstraction. To appear in Luciano Floridi's Philosophy of Technology: Critical Reflections, H. Demir, ed. Springer, forthcoming in 2013.