

IT RESEARCH AND DEVELOPMENT: SHOULD THERE BE CONTROL?

John Weckert

jweckert@csu.edu.au

Centre for Applied Philosophy and Public Ethics, Charles Sturt University, and
The Australian Institute of Computer Ethics

ABSTRACT

Two views are prevalent with respect to suitable topics for scientific research. One, typically held by researchers, is that any topic of research is fair game for the researcher. The result of research is knowledge, and knowledge in itself is neither good nor bad, therefore there can be no *moral* reasons for restricting research in any area. The proviso of course is that the research is undertaken in an appropriate manner. The other view is that science, even pure science, should serve the needs of the broader society, and therefore the society should have a say in what scientific research is conducted. Some should be avoided or even forbidden because it is harmful or will most likely be put to some harmful uses, because it is useless and therefore a waste of public assets, or because other research is more important. This debate rests on a tangled web of theories, reasoning and assumptions. This paper will explore some of these underlying issues and examine the responsibilities of professional researchers, with particular reference to research in artificial intelligence.

INTRODUCTION

Nearly twenty-five years ago Joseph Weizenbaum made the following strong statement that there are some kinds of research in which computer scientists should not engage:

There are ... two kinds of computer applications that either ought not be undertaken at all, or, if they are contemplated, should be approached with utmost caution.

The first kind I would call simply obscene. These are ones whose very contemplation ought to give rise to feelings of disgust in every civilized person. The proposal I have mentioned, that an animal's visual system and brain could be coupled to computers, is an example. ... I would put all projects that propose to substitute a computer system for a human function that involves interpersonal respect, understanding, and love in the same category. ... The second kind ... is that which can easily be seen to have irreversible and not entirely foreseeable side effects. (1984, pp. 268- 270)

More recently, Bill Joy, co-founder and Chief Scientist of Sun Microsystems, said this:

It is most of all the power of destructive self-replication in genetics, nanotechnology, and robotics (GNR) that should give us pause. ... In trust, we have had in hand for years clear warnings of the dangers inherent in widespread knowledge of GNR technologies - of the possibility of knowledge alone enabling mass destruction. ... The only realistic alternative I see is relinquishment: to limit development of the technologies that are too dangerous, by limiting our pursuit of certain kinds of knowledge. (*Wired*, 8.04, 2000)

If Weizenbaum and Joy are right then there is some research that computer scientists ought not do, and if they do, they can be held morally responsible for the consequences of that research. Or so it would seem. There are however a number of issues which need to be sorted out before we can be confident in affirming this. One concerns the difference between pure research on the one hand and technological development on the other. The first question here is whether there is any pure research that should not be undertaken. This issue is often discussed under the heading of "forbidden knowledge"; is there any knowledge that we should not attempt to discover? A related question is whether there is any technology that should not be developed. The final question is the moral responsibility of scientists and developers of technology, and the related question of their freedom. Are scientists and technologists morally responsible for the consequences of their discoveries and developments?

FORBIDDEN KNOWLEDGE

We need to draw a tripartite distinction. The knowledge itself must be distinguished from both the method of gaining that knowledge, and from the uses to which it is put. Clearly certain methods for gaining knowledge are wrong, for example those which harm people. Similarly certain uses of knowledge ought to be avoided if those uses cause harm. (In particular situations it might be that there is some greater good that makes some degree of harm permissible.) The question of whether the knowledge itself ought to be forbidden is, or seems to be, quite a different matter. The knowledge is neither morally good nor bad in the way that the methods or uses might be.

It is difficult to make sense of the claim that there is some knowledge that is morally bad in itself regardless of any consequences. It is certainly difficult to find any examples. It is easy to find examples of knowledge that have harmful consequences. Joy himself seems to acknowledge this when he said that knowledge of nanotechnology should be limited in order to limit the technologies that would be developed from that knowledge. If knowledge has harmful consequences should it be forbidden? Not always, because many types of knowledge can be put to both beneficial and harmful uses, and we do not want to automatically rule out the beneficial. To help to clarify matters, the consequences of knowledge can be divided into two groups, physical

and mental. Physical consequences are always uses to which the knowledge is put. The knowledge that is of most concern is that which will almost certainly be used in harmful ways. There may be no way to prevent the harm without preventing the knowledge in the first place. A case could be made that that knowledge is not the fit subject of research and ought to be forbidden. Mental consequences do not necessarily involve uses. Some knowledge is such that simply knowing it has negative consequences on life. Nicholas Rescher puts it this way:

Some information is simply not safe for us to have – not because there is something wrong with its possession in the abstract, but because it is the sort of thing we humans are not well suited to cope with. There are various things we simply ought not to know. If we did not have to live our lives amidst a fog of uncertainty about a whole range of matters that are actually of fundamental interest and importance to us, it would no longer be a human mode of existence that we would live. Instead, we would become a being of another sort, perhaps angelic, perhaps machine-like, but certainly not human. (1978, p. 9)

Suppose that as a result of research in IT it became known how to build machines which in behaviour were indistinguishable from humans, and moreover, that it was obvious that these machines were purely deterministic and without freewill. If we knew this, we would obviously have to see ourselves in a new light. Would we, in our present stage of evolution, be able to cope? If the GNR technology discussed by Joy develops in the manner that he fears, would we be able to continue to live happy and satisfying lives? If some knowledge has profound effects on the way we see ourselves, should it be forbidden? It seems that here, just as in the case of knowledge that almost inevitably leads to harm, a plausible argument can be made for forbidding it.

It has been argued that it can be justifiable to restrict or prohibit research, but who should do this restricting or prohibiting? According to Johnson (1996, pp. 197-217), while it is justifiable to prohibit knowledge (or certain kinds of research), there is nobody to do the forbidding in the way that God was able to forbid Adam and Eve from eating the fruit in the Garden of Eden. However, we must seek the potential forbidding!

Four possibilities are considered by Marcello Pera; the state, committees of experts, the scientists themselves, and the people. He rejects state intervention for a number of reasons, the most plausible being that it is a threat to individual freedom. The committee of experts is rejected because there are no experts in ethics and because it would lower the sense of responsibility of scientists, and he also rejects the scientists themselves, because they would serve their own interests. He concludes that “the supervision of science should be granted to everyone” (1989, pp. 58-72).

Pera’s conclusion is hardly satisfactory given that we are not told how “everyone” would make the decisions. The only possibilities seem to be numerous referenda, which would be impracticable, or through the state, which he rejects. The most plausible suggestions are the State or the scientists themselves. There are problems with state intervention, but with appropriate legislation these can be minimised. There are also problems with self-regulation by the scientists, but again despite Pera, this could work. With respect to the state, Singer writes

In general, the social consequences of attempting to put some areas of knowledge beyond reach will be worse than the consequences of gaining the knowledge, and doing our best to see that it is used responsibly (1996, p. 229)

Agazzi also has worries about state intervention in science:

If we were to accept social control of science, meaning total planning of scientific research to make it completely goal-oriented, we would also have to accept two highly undesirable consequences. First, we would have to limit the freedom of science ... Second, ... we could not avoid the question: “Who will determine the goals?” (1989, pp. 215-216)

A strong statement for the freedom of science from political control is supplied by David Baltimore:

First, the criteria determining what areas to restrain inevitably express certain sociopolitical attitudes that reflect a dominant ideology. Such criteria cannot be allowed to guide scientific choices. Second, attempts to restrain directions of scientific inquiry are more likely to be generally disruptive of science than to provide the desired specific restraints. (1979, p. 41)

A number of arguments are offered to support these claims. First, is what Baltimore calls the “Error of Futurism”, that is the supposition that we can predict the consequences of any research accurately enough to make any sensible decisions. The second argument is a version of one of John Stuart Mill’s. Freedom of speech and expression allows the development of new ideas, increases the choices in life and generally renews and vitalises life and makes it richer. A third argument is that repression in scientific research is likely to lead to repression in other areas and so will increase fear rather than strength in society. A fourth argument is based on the unpredictability of science. Even if some research is not allowed, the knowledge to which it may have led might emerge from other research quite unexpectedly.

This argument is aimed at pure or basic research, and Baltimore admits that the further one moves toward applications of research, the weaker these arguments become. If these arguments hold for pure research, then all of the responsibility for undertaking worthwhile research rests on the scientists themselves, which is where, according to Baltimore and others, it ought to rest (we will return to some of these others later). It must be noted here that it does not necessarily follow that the scientists’ responsibilities extend to the *uses* to which the knowledge is put. They create or generate the knowledge from their research but others decide how it is to be

used. Whether this separation of responsibilities is ultimately sustainable is another matter, but for the moment we will accept it.

We now return to Baltimore's arguments that research should not be externally controlled. His first argument, the "Error of Futurism" is that prediction is too unreliable to provide the basis for any restrictions on research. Consider for example Weizenbaum's prediction that research into speech recognition could have no useful consequences (1984, p. 271). It now appears to be an important tool in HCI for users with certain disabilities. Prediction is certainly fraught with danger, however we often must base our actions on predictions, on what we believe may happen, and it is not clear why the case of research should be any different. The second argument is that freedom of speech and expression allows the development of new ideas, thereby increasing lives' choices and generally making it richer. This is true, and this form of freedom is undoubtedly an important good, but it is not the only one, and can be in conflict with others. In general we are restricted in the performance of actions which will, or are likely to, harm others. Again it is unclear why research should be treated differently. The third argument is that repression in scientific research is likely to lead to repression in other areas and so will increase fear rather than strength in society. While we can agree that repression is not good, many things are restricted, or repressed in a civilised society, for example, certain research using human subjects and driving under the influence of alcohol, but these restrictions would surely reduce rather than increase fear. If probable harm is as closely associated with knowledge as suggested earlier, there is no reason why pure research should be treated differently from other aspects of life. The final argument was that, because of the unpredictability of science, the undesired knowledge might emerge unexpectedly from research other than that which was disallowed. This is true but not to the point. While it may not be possible to ensure that some undesirable knowledge will be discovered, it is almost certainly possible to reduce that probability.

It is then, permissible or even obligatory on occasions to restrict or forbid research on the ground that mental or physical harm is likely to result from it. However, this should not be done lightly, because freedom in this area is important, not only for the good of science but also for the good of society. There should be a presumption in favour of freedom of research. If research is to be restricted the burden of proof should be on those who want to restrict it. However, there seems to be a conflicting intuition here. If a *prima facie* case can be made that some particular research will most likely cause harm, either mentally or physically, then the burden of proof should be on those who want the research carried out to demonstrate that it is safe. The situation then appears to be this. There is a presumption in favour of freedom until such time as a *prima facie* case is made that the research is dangerous. The burden of proof then shifts from those opposing the research to those supporting it. At that stage the research should not begin or be continued until a good case has been made that it is safe.

The conclusion to this point then is that the case against the state having a role in the control of scientific research has not been made, but that such control has dangers and it should not be embraced lightly. The argument so far has focussed primarily on pure research, because that is where it is most difficult to make a case for control. However, it is not obviously much different in the case of technological development, one of the fruits of pure research. Just as a scientist can say that he or she is just adding to knowledge and therefore has no responsibility for the use to which that knowledge is put, so technologists can say that they are just developing tools, and it not up to them how those tools are used.

RESPONSIBILITY

We will leave this now for the moment and turn to the responsibilities of scientists.

According to Agazzi again:

A responsible person may be defined as one who is aware of his duties and ready to fulfil them. Scientists are already accustomed to respect some duties in their profession. These are mostly connected with what we call "intellectual integrity," but scientists have been less accustomed to considering other duties related to the 'context' of their research activity. The time has come for them to do so ... (1989, p. 217)

This needs closer examination.

Scientists are obviously responsible for the knowledge that they discover. If the knowledge has profound effects on the manner in which we see ourselves, whether this be beneficial or harmful, then the scientists must be responsible for that change. They must also take responsibility for the harmful consequences of knowledge that almost inevitably leads to harm.

However, some clarification of "responsibility" is in order. Hart isolates four such senses 1970, (pp 211-12): role-responsibility, causal-responsibility, liability-responsibility, and capacity-responsibility. These can be demonstrated as follows:

1. Role-responsibility.

The network manager is responsible for managing the computer network..

2. Causal-responsibility.

Dick is responsible for the virus on this computer because he knowingly copied an infected file.

3. Liability-responsibility.

Legal - Dora hacked into the university's student record system and changed some grades, so she is legally liable for this action.

Moral - in the same circumstances, Dorothy is liable in the sense of being morally blameworthy.

4. Capacity-responsibility.

Dick and Dora are responsible for their respective actions in the sense that they had the ability to choose to do or not to do them.

While all of these senses are important, the central one here is 3, liability-responsibility, which is closely related to accountability, the next concept that we need to consider. Accountability concerns who should be praised or rewarded if some good action is performed, or blamed or punished if a bad one occurs. It is not enough however, that one performed an action to be held accountable for it. When is a person P accountable for some event E? Three conditions must be satisfied: (1) P must have caused E, or knowingly allowed E to happen when he could have prevented it; (2) P must have intended to cause E or intended to allow it, or allowed it through neglect or carelessness (which implies that P did not have any intention *not* to cause or allow it) (duty of care); (3) P must have been free to choose to cause E or free to allow it.

Duty of care is a bit like accountability, but it is concerned primarily with carelessness or negligence. It is essentially the duty, in certain circumstances, to take care. One account states that 'whenever a person is so placed in relation to other people that failure to exercise care may foreseeably cause them injury, a duty of care is owed' (Creyke and Weeks, 1985, p 1). Perhaps the most quoted statement on duty of care is this, by Lord Atkin in 1932:

The rule that you are to love your neighbour becomes in law, you must not injure your neighbour; and the lawyer's question, Who is my Neighbour? receives a restricted reply. You must take reasonable care to avoid acts or omissions which you can reasonably foresee would be likely to injure your neighbour. Who, then, in law is my neighbour? The answer seems to be - persons who are so closely and directly affected by my act that I ought reasonably to have them in contemplation as being so affected when I am directing my mind to the acts or omissions which are called in question (Balkin and Davis, 1996, p. 200).

The emphasis here is not so much on refraining from actions which intentionally harm others, but rather on taking care to avoid actions which are likely to unintentionally harm others. So while accountability covers more than just duty of care, this duty highlights one aspect of accountability easily overlooked. Those who cause, usually indirectly and unintentionally, injury to others through negligence or carelessness, are accountable because they satisfy the conditions for accountability mentioned before. In particular, they satisfy the third alternative in condition (2).

In this section we have outlined the link between professionalism and duty of care. Claiming professionalism is claiming some special status in society, particularly with respect to trust, and with this status come special responsibilities. Because one is responsible in a certain way, one is accountable for one's actions and for the results of those actions if one does not take proper care. Professionals do have a duty of care, so to the extent that people are professionals, or to the extent that they claim to be, they have this duty, and this applies to scientists as well.

CONCLUSION

Should there be control of IT research and development, or put another way, should any knowledge or technological development be forbidden? The answer is complex, but we have seen that there is a case for saying that there should be some restrictions. It must be acknowledged that this answer is not without dangers. Control of science by the state can easily be abused, and has been. For this reason the State should only have a role in extreme circumstances. A much more satisfactory situation is one in which the scientists and developers themselves make the decisions on what to research and develop. However, they should not have the view that they cannot be held responsible for the results of their discoveries and developments. They have as professionals, it was argued, a duty of care. The situation is summed up well by Barry Allen:

Rather than an extra-scientific prohibition we should think about how scientific training undermines any nascent moral sense students may have of their responsibility for the knowledge entrusted to them. No prohibition, no forbidding of knowledge can begin to address a problem that can only be solved through changes in practice, especially in education... (1996, p. 308)

The concerns of Weizenbaum and Joy should be taken seriously. This paper has considered the issues generally, and not specifically in relation to IT. This is a necessary first step, but only a first step. The effects, or likely

effects, of IT research and development must be examined carefully. Perhaps an analogy can be drawn with the field of genetics. Research and development in genetic engineering continues, but at least there is open discussion of ethical questions, particularly regarding possible and likely consequences for the lives of humans, and life in general. So far this kind of discussion has been lacking in IT, but it is surely time that it began.

REFERENCES

- Agazzi, E. "Responsibility: the genuine ground for the regulation of free science", in W.R. Shea and B. Sitter (eds) **Scientists and Their Responsibilities**, Canton, MA: Watson Publishing Company International, 1989, pp. 203-19.
- Allen, Barry, "Forbidding knowledge", **Monist**, 79, 1996, pp. 294-310.
- Balkin, R.P. and Davis, J.L.R. **Law of Torts**, 2nd. ed. Sydney: Butterworths, 1996.
- Baltimore, David, "Limiting science: a biologist's perspective", in G. Holton and R.S. Morrison (eds) **Limits of Scientific Inquiry**, New York: W.W. Norton, 1979, pp. 37-45.
- Creyke, Robyn and Weeks, Phillipa. **Duty of Care: Law and the Direct Care Worker**. Canberra: Shield Press, 1985.
- Hart, H.L.A. **Punishment and Responsibility**. Oxford: Clarendon Press, 1970.
- Johnson, Deborah, "Forbidden knowledge and science as professional activity", **Monist**, 79, 1996, pp. 197-217.
- Joy, Bill, "Why the future doesn't need us", **Wired** 8.04, 2000.
- Pera, Marcello, "Should science be supervised, and if so, by whom?", in W.R. Shea and B. Sitter (eds) **Scientists and Their Responsibilities**, Canton, MA: Watson Publishing Company International, 1989, pp. 58-72.
- Rescher, Nicholas, **Forbidden Knowledge and Other Essays on the Philosophy of Cognition**, Dordrecht, Holland: D. Reidel, 1987.
- Singer, Peter, "Ethics and the limits of scientific freedom", **Monist**, 79, 1996, pp. 218-229.
- Weizenbaum, Joseph, **Computer Power and Human Reason: From Judgement to Calculation**, Harmondsworth: Penguin Books, 1984 (originally published San Francisco, W.H. Freeman 1976).