



# Threats

## How mad scientists can destroy the world

### A Futurizon Report

July 2008

This document is provided free of charge for your own personal use, but the copyright remains the property of Futurizon GmbH. No reproduction or distribution of this article in any form is allowed unless you have written consent from Futurizon.

You can read more articles on almost every aspect of the future at <http://www.futurizon.com>.

Futurizon provides speakers, consultancy, and commissioned reports on all aspects of the future.

Contact details: [info@futurizon.com](mailto:info@futurizon.com) or [idpearson@gmail.com](mailto:idpearson@gmail.com)

Author: Ian Pearson BSc DSc(hc) CIPF FBCS FWAAS FRSA FIN FWIF

©Futurizon GmbH 2008

[info@futurizon.net](mailto:info@futurizon.net)

## INTRODUCTION

The world is already becoming accustomed to asymmetric conflict. This works both ways. With technology superiority, advanced countries expect to sustain far fewer casualties than their enemy, and even a few casualties are becoming politically damaging. On the terrorism side, one or two terrorist can kill many people in asymmetric attacks such as suicide bombing, which is becoming almost routine in the middle east, but which can be highly effective politically when it happens closer to home, most notably New York's 9/11 and London's 7/7.

This type of attack is likely to stay with us and we have to adapt to it, but it is now obvious that with rapid technological advancement, the magnitude of threat can be much bigger. In particular, the internet creates more opportunities for network groups to orchestrate various types of attacks and NBIC convergence creates new opportunities for rogue states or 'mad scientists' to create and use new types of weapon.

## ONGOING TECHNOLOGY DEVELOPMENT

Technology development will continue to accelerate for the foreseeable future. As well as ongoing miniaturisation, novel threats will arise from artificial intelligence; IT/biotech convergence; advanced robotics; even from the physics of fast networks. Robotic insects are already developing quickly, driven by military surveillance needs, but this technology will inevitably become a corporate threat too. Tiny insects might enter a building, home in on key equipment and then either spy directly, or introduce malicious software or hardware.

Ongoing IT miniaturisation is making it increasingly possible to do very sophisticated things with tiny gadgets. Putting a tiny surveillance device into a piece of office equipment might allow signals to be intercepted and recorded during printing or scanning tasks. Then they could sit quietly until their owner removes them for subsequent downloading. Such miniaturisation will make corporate espionage easier. In fact, as devices get smaller and smaller, there comes a time where 'smart dust' becomes so small that individual devices could be too small to be seen by the naked eye, making it almost impossible to detect devices. Since they could be largely passive, and only respond to particular types of signal, they might be hard to detect even electronically.

## NETWORK COMMUNITIES

The web is already being used as a political platform, and it will become a default choice for organising political campaigns and actions. The web offers advantages over the physical world.

Firstly, web communities may include membership from a number of physical countries. These members can use anonymity technologies to hide from physical detection, organising and leading communities without fear of being identified.

Secondly, with members in a number of legal jurisdictions, it is not easy to control the activities of network communities.

Leaders can communicate with the membership almost instantly, and consequently to potential exists for actions to be initiated in a very small time and coordinated with fast response as they happen. It is very hard to do this in the 'real world'.

Software enables more sophisticated actions, even for peaceful activity. For example, electronic boycotts of products from particular companies or regimes can easily be implemented by means of emails that set preferences in buying profiles. Other software can enable a central command to use the power of each member's computers (with their

consent of course). This may be used for marketing, campaigning, and mass emailing. If a network community doesn't want to stay on strictly legal grounds, then of course they may also use this power for malicious cyber-attacks, again relying on anonymity to hide command structures.

## JIGSAW TECHNOLOGIES

In the film *Batman*, the Joker uses a combination of apparently innocent cosmetics to create poisons that kill the users. Individually, each one is safe, but together, they are lethal. In the real world, such combinations of innocent software have often caused unexpected problems. In the early days of email, mailing list software interacted with out of office software to create exponential growth of email volume until the servers fell over. These two programmes were both benign, but still caused problems, each out of office message being relayed back to the whole list, where other out-of-office messages would echo back, cascading to list server overload. The problem was easy to fix once identified, but it still happened. Black Monday on the London Stock Exchange might also have been predictable if anyone had really thought about it, but again it still happened. Since then, there have been various crashes and network outages caused by various combinations of software and circumstances that no-one had properly considered.

Of course, this potential to use combinations of apparently innocent software is a valuable and potentially powerful tool for malicious agents. Greater supervision of routine software might pick up a few potential activities, but actually there are infinite ways of making trouble by using combinations of software, trigger messages, distributed keys, positioning systems and even randomised documents from search engines. The most vigorous checks can only cover a finite number, so some will inevitably get through.

## NETWORK RESONANCE

Fast networks can provide a novel means of attack. With slow networks, the time between transmissions is long compared to the physical time it takes for a signal to cross a wire to an exchange or router. As transmission speeds increase to 100s of megabits per second, the time between packets is of the same order as transmission time across the network segment. This provides an opportunity to set up resonances, especially on polling networks, such as some designs of shared fibre networks. Network resonance can greatly reduce network capacity, and if timed to coincide with high load, can throw a network into an overload situation. There are many potential variants on resonance based attacks, which can be as damaging as today's denial of service attacks.

## INFORMATION WAVES

A related class of threat is the correlated traffic attack, where traffic generation can be coordinated precisely from various network points so as to generate information waves, again creating potentially overload situations. It is hard to design networks that are immune to all the types of correlated traffic types that are possible, so it only requires a reasonably expert and determined hacker to create significant network problems.

Information waves are a particular kind of correlated traffic. Imagine a future world where most people own mobile devices that are much more integrated into their everyday lives than today. For example, with a modest level of conventional AI, a mobile device might run software that tracks investments. When it receives a piece of information from the digital ether, it might make calls to a variety of servers, adjusting investments, or compiling a report to advise the owner. Given the tendency of our software markets to form virtual monopolies and monocultures, it would be safe to assume that many devices would run similar or identical software. So, a piece of news such as a significant rise in interest rates, emerges from a source in the centre of a city. As it propagates at the speed of light through the city, it

would create a wave of call attempts, following it just a few milliseconds later. The nodes near each device would check their traffic levels and decide that they can accept these extra calls, because they cannot know that other calls are being accepted exactly simultaneously by other nodes on the network. The result could be a gross overload situation, and if the network is not well designed, it might crash. When it comes back online, the software on all of those devices would detect that at the same time, and perhaps simultaneously try to resend their packets. Overload could happen again and again until people are told to switch off their devices and switch them on again, introducing some of the randomness back into the system that is essential for load balancing. Simulation of information waves suggests that a well placed phone call could switch off a certain kinds of network within 10 milliseconds.

## HARDWARE BACK DOORS

A more recent threat to security has arisen because virus protection software only checks other software for viruses, because the assumption is that is where all the threats reside. However, it is possible to build hardware based attacks, using field programmable gate arrays (FPGAs) to build custom hardware that interfaces directly with other equipment and bypasses virus security checks. Although at the moment this is a new approach, it is likely to grow as a problem, driven by improving design tools and the increasing availability of powerful, yet small devices.

Another hardware threat arises from the deliberate introduction of malicious algorithms into the hardware design during the design or manufacturing processes. It is quite possible to design hardware that achieves all its requirements but which has hidden circuitry that only comes into play when a particular instruction is received or a particular set of circumstances arises. Hardware testing can only make a finite number of tests but there are infinite ways that back doors can be put invisibly into circuit designs. To make the problem even more difficult to address, circuits that appear to be quite innocent might also be part of a larger malicious circuit or algorithm that only exists when other devices or software are brought into play. Such jigsaw approaches can be impossible to test for. Sleeper circuits could be already waiting in millions of machines, only coming into play when the final piece of the jigsaw is introduced via a superficially benign web site or an otherwise innocent-looking email.

## GAME CONSOLE NETWORKS

Game consoles present an interesting threat. Generally speaking, they are designed for games rather than mainstream computing, and because people don't usually rely on them for their home computing, there is much less focus on security, even though the recent generation are all heavily designed for good networking. In spite of this presumed weakness, console viruses apparently don't exist yet. Perhaps this is because virus writers are busy with PCs, they are so little used for web browsing, can't easily run downloaded content, or perhaps they are just better designed. Nevertheless, it is certainly possible in principle to hijack a games console, and since there is currently no security software for consoles, they are unprepared for such eventualities.

If large numbers of consoles can be hijacked, they could add up to a very powerful distributed computer, which could be used for conventional threats such as denial of service, spamming, or for more novel attacks such as creating network resonance, spawning information waves, or used as an encryption decoder to attack weaknesses in home banking systems, thereby creating a potential economic threat.

It is also significant that each of the recent generation of consoles represents at least 0.1% of the processing power of the human brain. If aggregated in millions, the total capability would be thousands of brain equivalents. This makes them a potential AI threat, even for relatively weak intelligence mechanisms.

## AI THREATS

Artificial intelligence will certainly be used to create more sophisticated virus variants, and autonomous AI entities will eventually become potential threats in their own right. Today, computers act only on instruction from people, but tomorrow, they will become a lot more independent. Assumptions that people will write the software are not valid. It is entirely feasible to develop design techniques that harness evolutionary and random chance principles, which could become much more sophisticated than today's primitive genetic algorithms. Many people underestimate the potential for AI based threats because they assume that all machines and their software must be designed by people, who have limited knowledge, but that is no longer true and will become increasingly untrue as time goes on. So someone intent on mischief could create a piece of software and release it onto the net, where it could evolve and adapt and take on a life of its own, creating problems for companies while hiding using anonymity, encryption and distribution. It could be very difficult to find and destroy many such entities.

Nightmare AI scenarios do not necessarily require someone to be intent on creating mischief. Student pranks or curiosity could be enough. For example, suppose that some top psychology students, synthetic biology students and a few decent hackers spend some time over a few drinks debating whether it is possible to create a conscious AI entity. Even though none of them has any deep understanding of how human consciousness works, or how to make an alternative kind of consciousness, they may have enough combined insight to start a large scale zombie network, perhaps using games machines, and to seed some crude algorithms as the base for an evolutionary experiment. Their lack of industrial experience also translates into a lack of design prejudice. Putting in some basic start-point ideas, coupled with imaginative thinking, a powerful distributed network of such machines would provide a formidable platform on which to run such an experiment. By making random changes to both algorithms and architecture, and perhaps using a 'guided evolution' approach, such an experiment might stumble across some techniques that offer promise, and eventually achieve a crude form of consciousness or advanced intelligence, both of which are dangerous. This might continue its development on its own, out of the direct control of the students. Even if the techniques it uses are very crude by comparison to those used by nature, the processing power and storage available to such a network offers vastly more raw scope than that available even in the human brain, and would perhaps allow an inefficient intelligence to still be superior to that of humans.

Once an AI reaches a certain level of intelligence, it would be capable of hiding, using distribution and encryption to disperse itself around the net. By developing its own techniques to capture more processing resources, it could benefit from a positive feedback loop, accelerating quickly towards a vastly superhuman entity. Although there is no reason to assume that it would necessarily be malicious, there is equally no reason to assume it would be benign. With its own curiosity, perhaps humans would become unintentional victims of its activities, in much the same way as insects on a building site.

No-one knows whether it is possible even in principle for true intelligence or consciousness as we know them to develop on digital circuits, however sophisticated. It may be possible, so presents a possible threat, but we don't yet know for certain whether such a threat is real or imagined. But we should remember that computing does not have to rely on digital circuits. Analog computing was once mainstream, and a great deal of pre-processing and sensory electronics still uses analog circuits. If anything, it is likely that most AI in the far future will rely on adaptive analog circuitry rather than digital chips.

Likely developments certainly substantially improve the scope and capability of computing and therefore the likelihood of the existence of a real threat, as well as its potential magnitude. The two areas I believe are of primary concern are computing gel and smart bacteria.

The first of the important developments that can take us away from conventional digital computing is computing gel. The last few years have taken processing chips from single cores to multiple cores, the number doubling every year or so. One of the big problems with chips is heat dispersal, and ongoing miniaturisation of circuits also creates quantum effect related errors. Moving to 3D architecture would allow higher device counts without needing to further shrink size, though some further shrinkage will also be achieved. When device counts increase, interconnection becomes a further problem. A fairly obvious solution is to start suspending processors in gel, using free-space optical links to connect them together. It is not necessary in principle to hard wire processors into fixed architectures at the point of manufacture. They could use self organisation to establish their own networks. In this way, with the gel providing cooling and a medium for communication, thousands or even millions of tiny processors could be suspended in a 100ml pot of gel. Each of the processors could have just a few thousand transistors, or millions. Some could be digital circuits, some analog. Such a medium with a suspension of analog and digital processors, unconfigured, would be an excellent raw base on which to run evolutionary and adaptive algorithms. Such a gel could be given a loose default structure, with some neural networks, some digital circuits, but mostly unassigned reconfigurable circuits. It could use high speed experimentation, configuring circuitry in many different ways and making pseudo-random modifications in a smart evolution program, to discover by itself the best circuits and architectures to achieve a large library of functions. It is impossible to predict the range of capability such a gel might be capable of achieving.

By far the most potential for such a gel comes from allowing it to build up its own function libraries, develop its own algorithms, and adapt its own architectures under its own control. It could store libraries of architectures that work, while reusing the same circuitry for experimenting with others to see if they might be even better for particular purposes. As its capability increases, it could start to download ideas off the net, try them, adapt them, and improve on them. Starting as the equivalent of the 'primordial soup', it could blaze through the equivalent of billions of years of evolution at high speed. At some point a powerful enough gel would pass human level of intelligence, and keep going. The power of such a device is limited by different factors than our biological intelligence. Cooling and energy supply are shared issues, but the gel approach has huge advantages in networked scalability, physical reach, device speed, transmission speed, multiplexing capability and hence connectivity, as well as processing and storage density and even in the range of properties it can sense. So a vastly superhuman, but alien, intelligence, is likely to be feasible by this route. It might be created for all the right reasons, e.g. to solve environmental or medical or energy or science problems. But once such powerful devices exist, they may be used for less benign human purposes, or indeed develop their own.

If it were instructed on the precise architecture required to replicate exactly a particular human brain, via reverse engineering, then there is every possibility that this could be done, and such reverse engineering via nanotechnology-enabled probes is already well under way. Such reverse engineering and replication is essentially just copying, there is absolutely no requirement to understand how something works. This makes irrelevant the oft-quoted objection to human level AI based on the fact that we don't understand how the brain works. We simply don't need to understand how the brain works in order to replicate it.

But with high speed electronics replacing slow biological switching and transmission, far more memory and raw processing speed could be available, and vastly increased sensory capability. Instead of having access only to sensors that are very local, such as our eyes and ears, a gel could have access to sensors globally via high speed networks. It is worth noting that a signal can travel from fingertip to brain in a few milliseconds, but that is roughly the time it also take for an optical signal to cross the Atlantic. So our replica brain would be immediately enhanced to global awareness and vastly superhuman speed and memory. That of course makes it potentially dangerous.

This class of threat is made all the more interesting when we realise that future AI could also have emotions. Emotions are not necessary for some kinds of machine, but would be useful in others, especially those that have to work closely

with people or animals. In the optical gel described above, where neural networks perhaps are designed to respond to intensity of incoming light signals, a beam of light of a particular colour shone into a particular region of the gel, would bias that group of neurons in much the same way as hormones do in the human brain. This is just one way of achieving emotions in a machine, but there would be countless other possible techniques. So when thinking of AI based threats, we must leave behind our prejudices that machines cannot feel emotions.

## SMART BACTERIA & SMART YOGURT

As biotech and its spinoff synthetic biology progress, it will eventually be possible to design and build living bacteria that can build and power electronic circuitry within their own cell. The proof of the basic principle was achieved some years ago when DNA was used in a test tube to assemble gold particles on the ends of carbon nanotubes. Custom proteins will be designed to carry out a number of industrial nano-assembly jobs including routine circuit assembly. Eventually, this will be adapted so that it can be done in situ within a living cell by modifying the cell's DNA so that life processes (including reproduction) are preserved but it can still achieve the assembly. While the first generations of this technology will involve the subsequent harvesting of the circuitry from the bacteria, it is not necessary in principle, and eventually it ought to be possible to let the bacteria continue their lives, powering the circuitry along with the rest of their life processes.

As they reproduce, large colonies of such 'smart bacteria' could self organise into highly sophisticated machines, effectively smart yogurt, which is essentially a biological means of producing the gel computing described already. Self organising technology is already developing quickly, often inspired by techniques in nature, and its use means that even small levels of circuitry in each bacterium would be enough to create large, complex circuits by means of their aggregation via self organisation. Preliminary calculations suggest that a pot of smart yogurt could provide the electronic foundations for an intelligence equivalent to an entire country. All that is required is a similar approach as described above to harness chance and evolution to develop basic circuitry into intelligence.

But advanced AI is not the only means by which smart bacteria and smart yogurt present a threat. With this level of miniaturisation, and the ability to exist and reproduce in nature, smart bacteria could be the ultimate security threat. Small groups of bacteria could provide little islands of processing, connected via network links to the whole smart bacterial intelligence, or there could be large numbers of independent intelligences. Sensory networks could link sensor colonies to one or numerous such entities. With a dual existence in both the real and virtual worlds, they present an interesting threat both to electronic and natural systems. Bacteria floating in the air or on peoples' skin, could easily live on surfaces like keyboards, furniture or office walls. Intercepting keystrokes as people type bypasses any electronic security that only comes into play once the information is inside the computer. Ultimately, bacteria might even be able to directly intercept brain activity, making a nonsense of any security system that involves people. In the extreme, they might even be able to control human behaviour. And of course, as a potentially superhuman intelligence of enormous proportions, we cannot possibly imagine the many ways that they could use technologies against us, based on science that we could never hope to understand.

## BACTERIAL GREY GOO

In the early days of nanotechnology, there was much debate about grey goo, the supposed ability of self replicating nano-machines to take apart all the matter around us. This could be deliberate, as in a weapon system, or accidental, with a design error, manufacturing defect, or an accident changing the behaviour of machines designed for manufacturing purposes. Since then, the threat has been shown by various scientists to be imaginary, and not something we should worry about.

However, most of the analysis that came to this conclusion was based on the robots being made of metal, tiny but complex machines. As far as the author is aware, none of the analysis took account of the self replication and reassembly being done by custom-designed bacteria. Bacterial grey goo remains a possibility that has not yet been proven unfeasible.

In nature, there are millions of species, and none of them manages to take over the ecosystem to the exclusion of all others. We simply don't see bacteria that can digest all other life, since life fights back. However, smart bacteria would have a strong advantage. With networking and intelligence, and a strong synthetic biology capability derived from human or machine research, and the consequent ability to redesign the workings of new strains of bacterial to attack new niches, smart bacteria could adapt quickly to any threat to themselves, while also adapting to capitalise on every biological niche. In short, they could take over the entire ecosystem. If they wanted, they could kill all other life, or simply eat all its food supplies so that it dies anyway.

## SOLAR WIND DEFLECTOR

Some years ago there were some interesting conspiracy theories surrounding the US HAARP project. Indeed, some are still online. While the project is claimed to be used for research, the main conspiracy theory was its ability to create a reflector by heating regions of the ionosphere would be developed so that microwaves could be reflected onto the enemy.

However, conspiracy theories aside, there remains the potential for someone to do such a thing. But reflecting surface based signals would be insignificant compared to the potential of using a reflector to create another, larger reflector or lens, by harnessing the power of the incoming solar wind. The solar wind has a power level of around 10 terawatts. A small reflector that uses some of this to create another larger one, that in turn is used to create an even larger one.. perhaps 10% of the solar wind could eventually be captured and directed. The resultant 1TW beam would be a powerful weapon indeed. If concentrated on a 5km x 5km square, e.g. a city centre or airport, it would release 40kw per square metre. Without the ability to disperse such power, the area would very quickly heat up, catch fire and be destroyed. Alternatively, the technique could be used to disrupt the earth's magnetic field, which in itself would cause problems since it is largely responsible for protecting us from radiation. Either way, another case for James Bond to deal with.

## ABOUT THE AUTHOR

Ian Pearson graduated in 1981 in Applied Mathematics and Theoretical Physics from [Queens University, Belfast](#). After four years in Shorts Missile Systems, he joined BT Laboratories as a performance analyst, and later worked in network design, computer evolution, cybernetics, and mobile systems. From 1991 until 2007, he was BT's Futurologist, tracking and predicting new developments throughout information technology, considering both technological and social implications. He now does the same for Futurizon, a small futures institute.

He is a Chartered Fellow of the British Computer Society, the World Academy of Art and Science, the Royal Society of Arts, the Institute of Nanotechnology and the World Innovation Foundation. He also holds an Honorary Doctor of Science degree from the University of Westminster.