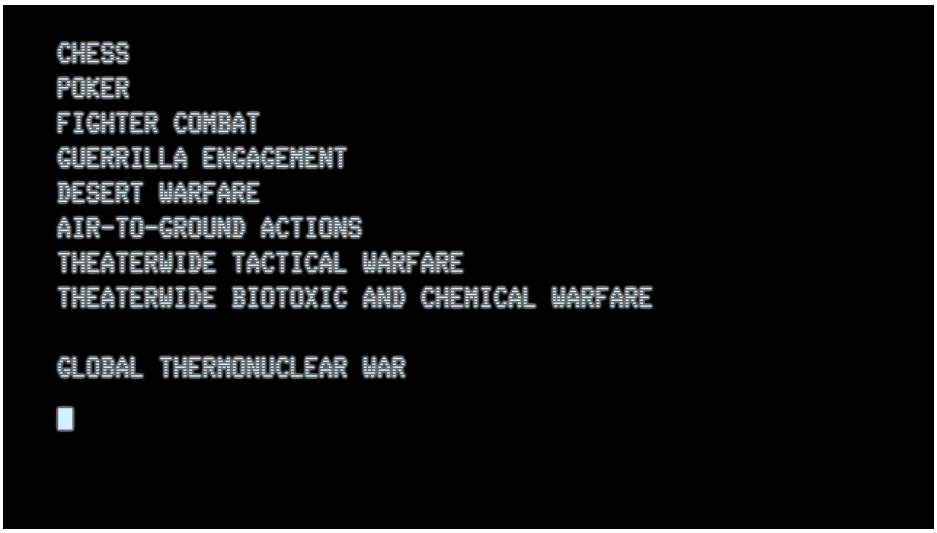
THE IMITATION GAME

Connie Veugen[1]



As soon as mankind moved from being hunters and gatherers to settle down and become farmers, they felt the need to protect themselves with walls and barriers, to shield themselves from wild animals, marauders and even enemy clans. Some of these barriers are still in place like the great wall of China. Others, like Hadrian's wall built by the Romans in the North of England to fend of the Picts, have crumbled, its stones used for other purposes like the building of other walled strongholds. And even though new inventions, most notably the canon, showed that every barrier has its weaknesses mankind did not stop building walled cities and elaborate walled defences. In the 20th century Hitler had enormous defence works put up such as the Panther-Wotan line and the Atlantic Wall to try and secure the territories his armies had conquered. Even today barriers and walls are erected like Trump's Mexican wall or Hungary's eastern wall to keep out immigrants. In a way these physical walls still serve a purpose, despite the moral and ethical questions they illicit. Their security, however, is only imaginary as the daily attacks governments, companies and even private citizens face are carried out using bits and bytes.

The reasons for that are twofold and to explain them I want to take you back to the beginning of networked computing. In America governments, universities and businesses had been using individual mainframe computers since the late

[1] Connie Veugen is a senior lecturer at the Faculty of Arts, VU Amsterdam. She is a specialist in Comparative Media studies with a focus on computer games, especially the assassin's creed series, and transmedia storytelling. More information can be found on her website www.veugen.net

1950s[2]. As we all know after the second world war the Cold War ensued with a rising fear both in America and in Russia of the outbreak of a global nuclear war. Consequently, after the launch of Sputnik, in 1957 president Eisenhower founded the Advanced Research Project Agency (ARPA) which was tasked with improving the military's use of computer technology. From 1962 the focus shifted more and more towards computer networking and communications technology, which resulted in the precursor of the internet: Arpanet. In 1969 the Arpanet connected 4 computers and on the 1st of May of that year the very first message was sent between two of these computers. After this initial success the Arpanet steadily grew, connecting 61 mainframe computers in 1975. Networks that spanned larger areas connecting computers in different cities are called wide area networks (see image 1). One could say that the severs connecting the internet form the ultimate Wide Area Network.

Moving on to personal computing, stand-alone personal computers (then called home computers) were available since 1975 including well known computers such as the Apple II and the TRS80. The real growth started in 1981 when IBM released its first personal computer which was both used in business and as a personal computer at home. In larger businesses, but also in universities these PC's were connected in so-called Local Area Networks or LANs, networks that spanned a building or a city (see image 1). As most private computer owners only had the one computer there was no need to have a computer network. However, in 1978, Ward Christensen, who worked at IBM found himself stuck at home due to a blizzard. It bugged him that he had no means to access the computer in his office so together with his partner Randy Suess he developed a dial-in system to enable him (and others) to remotely access another computer using a modem and a standard telephone line. By giving other people access to his own computer, he created what became known as a Bulletin Board System (BBS), a computer where people could leave and answer messages. This was the beginning of the public dial-in modem networks.

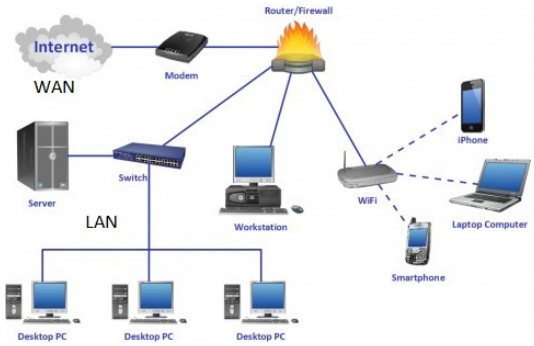


Image 1 Wide Area Network, Local Area Network & Firewall

[2] These mainframe computers were multi-tasking and multi-user but they were stand-alone i.e. not connected to other computers through a network.

This brings us to 1983 and the film WAR GAMES. This film not only shows the need for virtual barriers between computers, but also gives another reason why human dependence on computer technology should always be critically assessed. The computer whiz-kid protagonist in the film, David (played by Mathew Broderick), uses a dial-in connection to change his grades in the school computer. The only barrier to prevent unauthorized access to this computer is a password that is regularly changed. But David knows where it is written down (a mistake many people still make) so accessing the school computer is not really a problem. David is a keen gamer, so when he is at school, he has his home computer dial random numbers in the hope to find another computer 'answering'. As soon as he has got time David then dials the successful numbers to see if the other computer can be logged-in-to and if it contains games. In this way he finds an intriguing computer which is not as easy to hack. He goes to consult some professional computer friends who tell him that most programmers often built-in secret 'doors' to allow them easy access to the system in case of emergency, avoiding other security measures that might be in place. These so-called 'backdoors' give them 'root' access i.e. access at the highest security level. This also holds true for the state defence computer in the film called WOPR (War Operation Plan Response) an AI computer that was built to control the nuclear defence system as such an important task could no longer be trusted to humans. This is shown in the first scenes of the film when military personnel fail to turn the switch in a training session as they are too aware of the consequences:

> *You can't screen out human response!*
> *Those men in the silos know what it means to turn the keys,*
> *and some are just not up to it!*
> *Now it's as simple as that.*
> *I think we ought a take the men out of the loop…*
>
> *…General we all know they're fine men,*
> *but in a nuclear war we can't afford to have our missiles*
> *lying dormant in those silos*
> *because those men refuse to turn the key when the computers tell 'em to!*

As can be imagined the responding computer David is trying to hack is WOPR[3]. David finds the backdoor but still fails to find the correct password. Asking the computer for help, David finds a list of games and incorrectly assumes that he has found a computer owned by a games company (image 2).

[3] Of course, one can wonder why such an important computer can be accessed via a dial-in connection. The film explains that this is a cock-up by the telephone company.

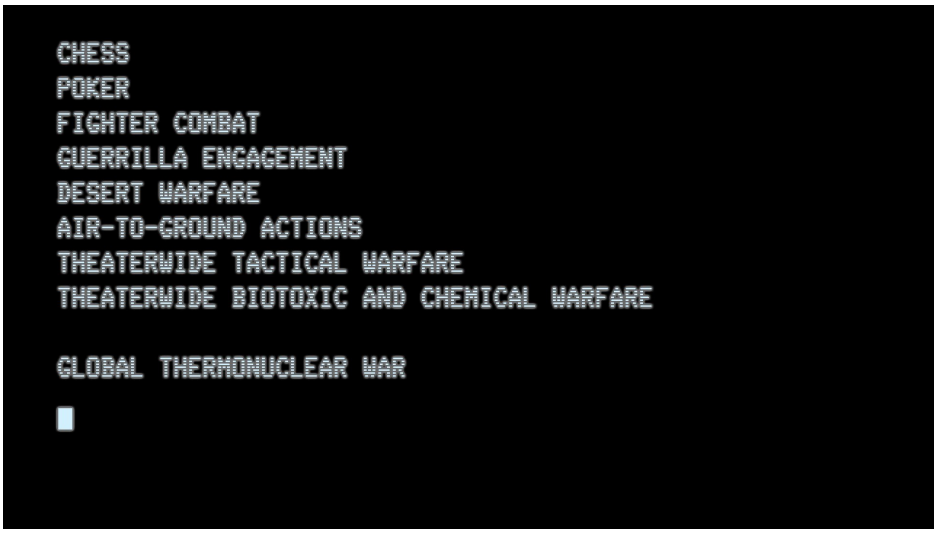


Image 2 Part of the list of games WOPR can play

Naturally David chooses to play "Global Thermonuclear War" taking the side of the Russians. Consequently, WOPR begins to compute the best nuclear response, however, when it becomes clear that the Russian attack is not real, it is impossible to stop WOPR as the computer wants to win the game. And as WOPR's AI can launch nuclear warheads independently the threat becomes very real. To prevent world destruction David now truly has to find the computer's backdoor password, which turns out to be the name of the son of its original programmer: Joshua (deceased at an early age)[4]. Finally, Armageddon is avoided by teaching WOPR Tic-Tac-Toe, a game that has no winners or losers (image 3).

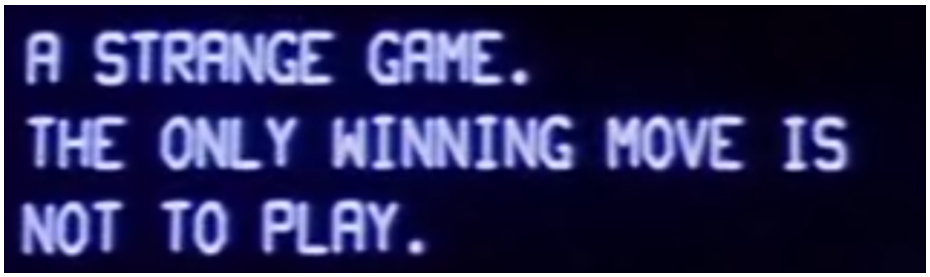


Image 3 Joshua's conclusion

Of course, the film is fictional (although it is alleged to be based on a real incident with a military computer) but incidents with unauthorised outside access did increase. Consequently, by the late 1980s when it became more common to connect computers in a network of both trusted internal computers and untrusted external computers new technology was introduced to protect the internal network. These non-physical barriers to protect computers in a network are called firewalls, virtual barriers monitoring the network traffic between two

[4] Who, non-accidentally (imho), is named after Alan Turing's (the father of AI) childhood friend who also died at an early age. Hence one of the allusions of the title of this article (the other being of course that you enter the system pretending to be someone else).

connected computers, based on predetermined security rules[5]. Everyone who has owned a connected computer around the time (and into the late nineties) had to install firewall software to secure their computer. Still, the best way to protect important (military, government, research) computers is not to connect them to a network at all. This is common practice but as the 2007 attack on the Iranian Nuclear complex in Natanz showed in real case scenarios this is not enough. Old fashioned espionage, infiltration, a USB stick and a powerful virus (in this case Stuxnet) are enough to cause serious damage and to trigger a new phase in modern warfare: cyberwar[6].

Still, as later films of the 80s and 90s have shown (a.o. HACKERS (1995)) most attacks are still conducted using networked computers. Targeted hacks are not as easy to avoid, despite the installation of a firewall, as people still use very simple passwords[7] or are tricked by phishing mails or calls by personnel that allegedly works for Microsoft. What has changed with the introduction of the internet, easier-to-use interfaces and especially the internet-of-things is that we have become more and more dependent on connected systems and AI and that we have become far less aware of the increased danger this poses. Operating systems have become increasingly friendly, starting with Apple's 'What you see is what you get' (WYSIWYG) McIntosh interface in 1984. And security software such as virus scanners and firewalls have become a part of some well-known operating systems. Consequently, a large number of people do not even realize that, for instance, Android smartphones (that often do not have these features) are vulnerable to malicious software[8]. Our increasing dependence on computer technology poses a serious threat to our privacy, which is underlined by the almost daily news reports of AI systems that gather our data without our explicitly stated consent[9] but also our blind reliance on the AI itself. The internet and the internet-of-things, the ever-faster processor techniques and the sharp decline in storage costs have meant that enormous amounts of data are stored every minute of every day. If we just limit ourselves to WhatsApp messages according to Statista in May 2018 65 billion WhatsApp messages where send globally, add to this SMS messages, Apple iMessages, Android Messages, Facebook Messenger Messages, Twitter, Snapchat, Instagram and other platforms, a guesstimate of more than 200 billion messages per month worldwide

[5] For those of you wondering where the term comes from, a firewall in its original meaning is a physical barrier (originally a specific kind of wall) to stop a fire from spreading (see image 1).
[6] Already, the 1995 film THE NET showed us that you do not need a network to spread a malicious data-stealing virus.
[7] But with the myriad of applications we have installed on our computers, tablets and smartphones it is hard to come up with safe and unique passwords. Research in April 2019 showed that the most common password is 123456789 used by 7.7 million users.
[8] Some weeks ago, security experts confirmed that it has become virtually impossible for the general public to secure their devices. As the past years have shown the corporations that hold our data like Google and Facebook, but also our national banks should be held responsible for the 'theft' of our data by companies like Cambridge Analytica or our money by bugs in their own software like the Dutch ING bank.
[9] Alexi and her AI competitors are well known, but did you know that your robot vacuum cleaner also sends all manner of data to companies whose chips are used in the machine?

does not seem farfetched. By far too much data to be handled by humans alone. The term used for this amount of data is Big Data and to extract information from such an enormous amount of data AI has become a necessity. However, to train AI we do not only need the data and computers with a lot of processing capacity, we also need algorithms. And that is where the new problems arise.

As Cathy O'Neil shows in her brilliant book *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy* (2016), the main reasons algorithms are less neutral than most of us think is that the new AI systems are no longer programmed rule by rule. Many of our modern AI systems train themselves. Strictly speaking this is not a problem, it depends on the data that is used and whether or not the domain is clear. The AI itself is neutral, it does not know whether the choices it makes are better or worse than other choices. One of the domains AI is increasingly used for is image recognition. This is also the case in the 1987 film NO WAY out. In this film Kevin Costner plays a Russian spy, who is sought for the murder of a young woman who is both the mistress of a politician but also Costner's girlfriend. The damaging evidence that could identify him is the negative of a Polaroid found at the scene (image 4).

Image 4 The Polaroid negative found at the scene

This negative is handed over to the CIA to reconstruct, using the latest advances in AI. However, at the time the self-learning AI still needed some human input:

> *The computer just asks itself what if we played with the pixels?*
> *That is all that there is to it?*
> *No, you have to keep on top of the computer, computers don't think.*
> *For example, if we had programmed this to come out a car,*
> *instead of a monkey, right now you would be looking at a hairy Buick.*

Presently, modern versions of image recognition AI have shown to be very helpful when applied to a specific domain. For instance, in medical sciences where they are used successfully in several fields such as early diagnostics and/or

the monitoring of possible complications[10]. However, as the art-installation *IMAGENET ROULETTE* shows, training image recognition AI is still fraud with unreliable algorithms. This time the problem is not the distinction between a car and a monkey but the biased training examples of image data depicting people.

Due to the ever-increasing amount of data that has to be processed, more and more AI systems are now trained using so-called deep learning (learning without human monitoring). Again, the AI itself is neutral, but as there is no human monitoring of the learning process the data used to train the AI is even more critical. Self-trained AI are strictly speaking a number of black-boxes. We rely on their outcome without knowing what happens in the box itself[11]. If such an AI is trained with biased data, the output is by nature also biased. Yet in most cases, we do not know as we can no longer guesstimate what output to expect[12]. One such an example was a company that used AI to make their hiring process less biased. Yet, because the data used was based on previous successful applications the successful applicants were still predominately white male. In America, the first court cases against biased algorithms are already being tried. For instance, the 2018 case against the Los Angeles Police Department. The LAPD started using Big Data interpretation in 2011 to predict where future crimes would probably be committed and by whom. With limited resources this seemed to be a reasonable way for a police department to use its money well. However, as the system bases its data on known criminals, it was already biased as most of the convicted criminals in America as a whole but also in Los Angeles are black American males. Consequently, they were pre-targeted by the AI. Also, the system failed to alert the police to the much more damaging instances of cybercrime (white collar crime) as these were (and are) underrepresented in the system as convictions are rare. Another, related problem is that the people who gather the data are blind to the possibility of built-in biases. In medicine, until recently most research data collected bases on white middle class men. Consequently, doctors are not trained to diagnose heart failure in women, not to mention any knowledge of how to treat certain diseases that are specific to Asian or Afro-American people. So, a more critical approach to the way data is gathered and the perceived infallibility of computers is still, and increasingly necessary. It is not enough that entertainment media have to remind us of what can go wrong, as for instance in the TV-series STAR TREK DISCOVERY (2019) where in season 2 we learn that the decisions necessary to rule the different

[10] Of course, in our present-day reality where everything is monitored by public or even private security cameras, a TV series such as PERSON OF INTEREST (2011-2016) shows that the CIA no longer needs discarded photo negatives. Even more than our digital messages our image data is also recorded and used without our explicit consent or even realisation. What if that data is then manipulated real-time? As in the current BBC series THE CAPTURE (2019).

[11] And because of the complexity if the output is 'wrong' we have no way of finding out what went wrong.

[12] This reminds me of the time when electronic calculators were allowed in schools. Instead of relying on their own calculating abilities the students more and more relied on their calculators, even going so far as to correct the teacher because surely the "calculator was right".

worlds and races are too difficult to be resolved by the intergalactic council alone, so they rely on a complicated AI that helps them form opinions. As was to be expected (remember WAR GAMES) there comes a moment when the 'human' factor is abandoned and the AI makes all the (rational) decisions. Needless to say, that the AI decides that the 'human' factor itself can be eliminated in favour of more AI[13].

What the TV series addresses here is what has been dubbed superintelligent AI. What if the AI becomes more intelligent than humans? How do we prevent an AI from doing us harm? Of course, this is different from the garbage in – garbage out problem of the biased data. The only way to deal with the latter is to make programmers, producers, those who commission the software, but also end-users more and more aware of this particular problem and as educators to not only train our students to be more critical but also to teach them responsible non-biased data collection or to limit the data to a very specific domain[14]. Recently, in an interview with *de Volkskrant* the director of Royal Dutch Library (Koninklijke Bibliotheek) showed such an awareness. He explained that the KB is currently training AI to help assigning keywords to their increasing collection of digitized works. When asked, he made it clear that he did not believe that the AI would ever be capable to do this without human interference: "I'd prefer a combination of human/machine. The computer gives suggestions, but the human has the final say. This also to prevent bias (in the AI)" [my translation][15]. In superintelligent AI the software does not train itself for one particular purpose and domain, it trains itself to solve increasingly difficult non domain specific tasks. Again, in principle, this is not a bad thing and as our problems become increasingly difficult (cf. global warming), like the council in STAR TREK DISCOVERY we will need superintelligent algorithms to help us find solutions. However, the problem with self-learning superintelligent AI is that it might come up with solutions which to us are not what we were looking for at all. In 2013, programmer Tom Murphy VII designed an algorithm that taught itself to play Nintendo games (cf. WOPR's designer teaching the AI strategy by having it play games). As the AI's ultimate goal was not to lose (another way of looking at winning) while playing *TETRIS* it would simply pause the game indefinitely as "the only winning move is not to play" as Murphy observed in a paper (thus repeating WOPR's conclusion, see image 3). So, is human kind doomed to extinction in thirty years as a consequence of superintelligence as some pessimists have predicted? Not if more money is spend researching superintelligent systems. Research that is already being conducted by Skype's cofounder Jaan Talinn who in 2012 cofounded the Cambridge Centre for the Study of Existential Risk

[13] For the outcome see STAR TREK DISCOVERY.

[14] Training data for language AI is often taken from Twitter. As we all know, what Twitter data to use is already critical. Just imagine training a language AI with Donald Trump's Twitter feed.

[15] "Ik denk eerder aan een combinatie mens/machine. De computer komt met suggesties, maar de mens heeft het laatste woord. Ook om te voorkomen dat er vooroordelen in een systeem sluipen". For those who are wondering how bias could have happen: the AI assigns keywords based on automatically analysing the original text.

(CSER) or Eliezer Yudkowsky and his Machine Intelligence Research Institute or Stuart Armstrong at Oxford University's Future of Humanity Institute, they and others are looking into solutions how superintelligence can be contained[16]. One of the solutions is a so-called kill-switch which was also already predicted in the science fiction film BLADE RUNNER (1982), loosely based on Isaac Asimov's novel *Do Androids Dream of Electric Sheep?*, where the so-called replicants (humanoid AI) have a built-in kill-switch that ensures that they can only operate a maximum of three years (image 5)[17].

Image 5 Roy Batty's Final Speech

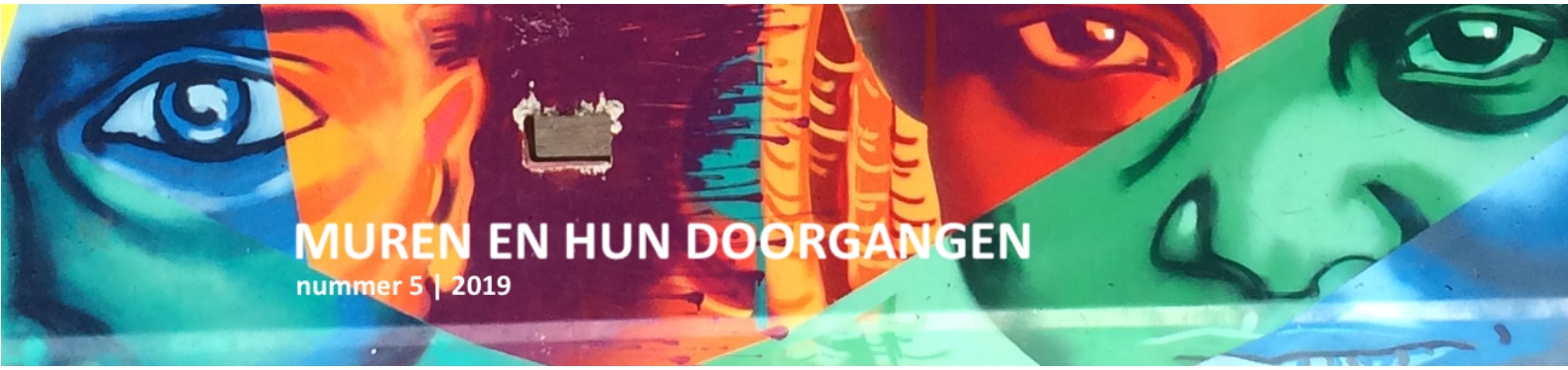
> *I've seen things you people wouldn't believe. Attack ships on fire off the shoulder of Orion. I watched C-beams glitter in the dark near the Tannhäuser Gate. All those moments will be lost in time, like tears in rain. Time to die.*

The last words of Replicant Roy Batty (Rutger Hauer).

Just some weeks ago, I was discussing human operated kill-switches with a professor of evolutionary biology, two philosophers and robotics and AI researchers. The latter were in the early stages of building self-replicating AI robots to help colonize Mars. The researchers were full of optimism, while the professor and myself were really doubtful that such a kill-switch would work. We not only had the WAR GAMES example, but also remembered the absolute devastation of young Tamagotchi owners when their robot pet died.

[16] For a very insightful article on superintelligence and the people studying it see Mara Hvistendahl article in the *Guardian* 'Can we stop AI from outsmarting humanity? Online at https://www.theguardian.com/technology/2019/mar/28/can-we-stop-robots-outsmarting-humanity-artificial-intelligence-singularity.

[17] In the film JURASSIC PARK the dinosaurs also had a limited life-span. Interestingly in the game *JURASSIC WORLD EVOLUTION* one of the things you can do to get better game results is tweak the lifespan of your dinosuars.



Stichting IVMV
INSTITUUT voor
MAATSCHAPPELIJKE
verbeelding

I N H O U D S O P G A V E
nummer 5 | 2019

MUREN EN HUN DOORGANGEN
IVMV Online Magazine

Visuele mythologie

10. Gabriël van den Brink, ‘De avonturen van Kuifje als getekende mythologie’
11. Boudewijn Steur, ‘Beyond the wall’

Digitale doorgangen

12. Connie Veugen, ‘The imitation game’
13. Sigrid Burg, ‘Digitale muren. Bescherming of valkuil?’

Beeldbijdrage: Tosca Philipsen, ‘Grensbegroeiing’