Scambaiters, Human Flesh Search Engine, Perverted justice, and Internet Haganah: Villains, Avengers, or Saviors on the Internet?

Andreas Zingerle

University of Art and Design
Linz, Austria
andreas.zingerle@ufg.ac.at

Abstract

In recent years, Internet users have been increasingly participating in so called digilantes or cyber-vigilante communities, becoming self-appointed avengers of justice who wade through the Internet to hunt down unlawful netizens. These groups see the legal mechanisms for criminal punishment as ineffective and use social networks to crowd-source both the prosecution and the execution of punishment. I conducted an experimental investigation into these justice-seeking activist groups to compare the 'scambaiting' anti-fraud movement and their methods and similar web-formations like 'Perverted justice', 'Human Flesh Search Engine' and 'Internet Haganah'. Each group's motives are presented with recent examples, and parallels are drawn to similar projects carried out by journalists, artists or activists. Mass mediated prosecutions entertain popular culture and are used to regulate social norms. It was found that vigilante communities use congruent techniques in gathering intelligence and use comparable prosecution methods like shaming, humiliation, cyber bullying, or doxing. Furthermore, moral concerns of these deviant actions and possibilities of governmentality are discussed.

Keywords

Vigilante online communities, digilantism, hacktivism.

Introduction

Vigilantes are self-appointed citizens who take the law in their own hands, when law enforcement is inefficient or not present. In 17th and 18th century feudal societies, personal vengeance was exacted through duels using swords or pistols. The two opponents agreed on matching weapons and to obey certain rules. In Western cultural traditions, the vigilante has been romanticized as the dissatisfied moral avenger: Robin Hood and his fellow outlaws fight the rich and share the booty with those in need. [7] During the Gold Rush in the 19th century, vigilante committees were formed in mining communities to deal with the rising crime rates and the lack of legal institutions. [18]

Nowadays, vigilante communities are seen as community service, like the national neighborhood watch groups 'USAonwatch' or 'Guardian Angels' in the United States or the 'Ourwatch' community in the United Kingdom. Often, members of these groups have difficult relationships with the police and prefer to take justice into their own hands and deal with the criminal themselves. In California, the Minuteman project is an activist group, which uses a network of webcams to patrol the US-Mexico border and monitor the influx of illegal immigrants. It describes itself as a 'neighborhood watch' on the borders. Volunteers report undocumented immigrants and smugglers to the Border Patrol. However, there have also been cases reported, where rights of Mexican citizens have been violated by shaming and humiliating them.

In one report, Bryan Barton caught a Mexican citizen crossing the border illegally, detained him and forced him to pose with a t-shirt saying: 'Bryan Barton caught an illegal alien and all I got was this lousy T-shirt'. [4] After the incident, the Border Patrol and the Mexican Consul agreed that no crime had been committed, but the Mexican immigrant was still publicly humiliated. When the Minutemen are not patrolling the border, they erect fences along the US-Mexican border.

Since the late 70s, groups of 'Real-life Superheroes' (RLSH) patrol cities. Inspired by their fictional archetypes like Batman, Guy Fawkes-masked V or the Watchmen group, they hide their identity by wearing masks and dress up in costumes to fight crime or perform public services, just like their comic book heroes. [31] Across the globe there are more than 200 registered RLSH that can be grouped into 'social workers with capes' and 'radical activists'. [19]

With the advent of mass-mediated online communication, vigilante groups establish themselves in web forums, discussion sites and Social Media platforms to express alternative public opinions on these new frontiers of the Web 2.0. Their actions have many names, 'DIY-justice', 'e-vigilantism', 'civilian policing', 'digilantism' or 'cyber-vigilantism', but in this article I will use the most commonly used name 'Internet vigilantism'.

Although a number of studies have examined vigilante communities like the 'Scambaiters' [38, 27, 32], 'Perverted justice' [12, 14], 'Human Flesh Search Engine' [37, 8] or 'Internet Haganah' [35, 9], a review of the literature indicates that there has not been a focus on mapping out parallels between those communities. With this paper I want to provide additional insights into the similarities in the use of tools and techniques and comparable prosecution methods used by these vigilante groups.

The paper is composed of four themed chapters:

- The first section of this article examines the scambaiting community 'Artists Against 419' (AA419) and outlines their prosecution tool 'Lad Vampire'.
- Chapter Two explains the 'Negobot' and 'Chatcoder' tools that the 'Perverted Justice' (PJ) movement uses.
- The third chapter is concerned with the 'Human Flesh Search Engine' (HFSE) and lays out their recent activities.
- The fourth section presents the findings of the research, focusing on the 'Internet Haganah' (IH) movement.

Finally, the conclusion gives a brief summary and critique of the moral and ethical issues concerning these groups and the government's proper role in online governance.

Crowdsourced Online Justice

Tatiana Bazzichelli states in her book *Networked Disruption* that artists, hackers and activist groups (AHA groups) use disruptive techniques of networking in the framework of Social Media and web-based services to generate new modalities for using technology, which in some cases are unpredictable and critical. This two way strategy in networking contexts can be used as a practice for generating criticism and can serve as a methodology for business innovation. These 'AHA groups' critically rethink interventions in hacking culture, art and technology; they accept that they must act from within the market scenario in order to change it, while ironically deconstructing it at the same time. This way, the goal is not to oppose frontally, but to trick them by becoming them and creating disruptive and ironic camouflages. [5] Similar tactics can be observed when investigating vigilante online communities and their practices. In recent publications, I mapped out vernacular tools used by scambaiters or how they use social engineering practices when communicating with Internet scammers. [39]

For this experimental investigation, I conducted autoethnographic research on different vigilante communities and documented the parallels to artistic and journalistic practices to map out correlations of their usage of technology and working practices. The 'Scambaiting' community is a global movement, which contacts Internet fraudsters in order to document their practice or jam their workflow. I observed the group very critically after conducting a 'scambait' myself without knowing of the existence of such an active online community. My initial intentions were based on curiosity, and the communication with the scammers gave me an opportunity to understand and document their working practices. By participating in scambaiting forums I encountered several subgroups, each following their own agenda ranging from wasting a scammer's time and humiliating them online to tech savvy activists who shut down fake websites, monitor scammers' email accounts or track down online groomers and romance scammers. [29] Members of these different subgroups were also involved in other vigilante communities: the 'Perverted Justice' movement, mass-mediated actions of the 'Human Flesh Search Engine' and the 'Internet Haganah' group, as well as their tools and techniques for obtaining background checks on website administrators. In the following paragraphs I want to introduce these communities and take a closer look at their methods and practices with the help of several case studies.

Scambaiting Communities Against Online Fraud

Scambaiters are online communities that take action against online advance fee fraud. They actively report scam emails to email providers, collect phone numbers or track IP-addresses of the senders and publish them on platforms to warn other Internet users. In order to be able to process a large number of emails, general warning platforms like scamwarners.com are assisted by more specialized forums that only document specific scam scripts, e.g. romance or employment scams or forums that document scam tactics like phone scams.

Some scambaiters create fake characters with email addresses and social media profiles and use these virtual personas to contact scammers. Often they act like gullible victims to give the scammer the feeling of an easy prey. Once the scammer takes the bait, the scambaiters document the scammer's working practice, for example by collecting identity cards and bank information, in order to document and jam the scammer's workflow. Some scambaiters specialize in reporting bank accounts or warning hosting providers about fake websites on their servers. In some cases, scambaiters manipulate scammers to leave their place of work and travel to remote areas, thereby actively jamming the scammers' workflow and making the travel as long and tedious as possible. [38]

Scambaiters use social engineering methods and several vernacular online tools to create trustworthy characters and believable storylines. Online tools like name-generators help to create fake characters with believable names and existing street addresses. When using VoIP telephony to be in direct contact with the scammers, they use voice morphers to pitch their voices or webcam add-ons to use pre-recorded videos that mimic live video feed. The scambaiters often ask for photographs of the scammers and ask them to pose with obscure signs or in humiliating poses. These photos are collected on online forums like the 'Hall of Shame', where they become memes or are virally shared with the public. Several forums document scambaits where users can comment on the stories and share tips on how to make them more humiliating and hilarious. [27] Forums like 419eater.com or thescambaiters.com specifically distance themselves from racist actions and claim to ban such users from their forums. Within the forum communities the members often challenge each other to submit photos of scammers in more and more hilarious positions. This is to prove to the community that the scambaiter has talent in persuading the scammer to believe their ridiculous stories. Therefore each forum member maintains a posting signature that is added to every posted message. Icons and animated gifs indicate their achievements: pigs indicate closed bank accounts, country flags represent shut down websites, hats for sucessfully sending a scammer on a travel.

Scambaiters try to unveil the real identities of the scammers and expose them to their friends and families. In order to do this, they request the scammers to submit scanned ID's or other documents and images to prove their authenticity. Images that come in the .jpg or .tiff format carry metadata that is stored as 'Exchangeable image file format' data (short Exif-data). When taking a photo, metadata like date, time and camera settings (e.g. camera model, aperture, shutter speed, focal length), GPS location information and a thumbnail of the image is saved and embedded within the image file itself. This is mostly done by default without the camera owner's knowledge. Scambaiters analyze the Exif-data and see whether a photo has been edited or when and where it was taken. This can often serve as additional information to prove the authenticity of a story.

There are reports where the scambaiters provided enough evidence to successfully catch the scammer, but

most times people just laugh at the scammers and feel superior to the petty criminals. Within these different motives and subgroups of scambaiters, the next paragraph is dedicated to the 'Artists Against 419' group and highlights some of their working tools.

The 'Artists Against 419' (AA419) is an international community that documents fake websites and tries to educate the public on how dangerous it can be to trust companies' online representations. Scammers often use fake websites and top-level domains like .com, .co.uk, or .net addresses to add credibility to their stories. AA419 started out by reporting fake bank sites that were used for phishing attacks. This was done by cross-checking the companies' websites with local regulator's lists. Back in 2003, a small group of net activists started using custom software to take down fake bank websites. They called these acts 'virtual flash mobs' and their programs were called 'Mugu Marauder', 'Muguito' or 'Lad Vampire'. [1] These programs repeatedly downloaded images from the fraudulent website until the bandwidth limit was exceeded and the hosting provider blocked the public access to the website for the rest of the month. This action can be considered as 'bandwidth hogging' and enabled the vigilante group to block access to fraudulent websites, if the hosting provider didn't react to their written complaints. The act of 'bandwith hogging' is often miscredited as a Distributed Denial-of-service attack (DDoS), but a DDoS attack targets the whole server, where normally several other websites are hosted and not just a single website. [24] The group provoked a lot of discussions and controversy with these illegal virtual flash mobs, so they discontinued the development of those particular software programs after September 14th, 2007. Since then their main focus is on writing complaint letters to hosting providers and establishing a reliable alliance with them.

Through a public database they publish fraudulent websites and link these entries to publicly available 'Domain Name Server' (DNS) entries. This DNS information shows the hosting provider's name and address, the date of registration or when the website was updated last. Besides banks they document all sorts of online businesses; international couriers, escrow services, insurance companies, online shops, construction companies, trading agencies, job or travel platforms. So far, the AA419 lists the biggest collection of fake websites, and the community actively maintains international relations with law enforcement, web hosting companies and domain registrars to get fraudulent websites removed from the Internet.

Hunting Online Pedophiles – The 'Perverted Justice Movement'

Perverted-Justice (PJ) is a civilian watchdog group and online community that tries to expose adult predators trying to contact minors through online chat rooms. [14] They setup sting operations by their members, who create fake profiles and pose as young teenagers, and log in to chat rooms and forums to make contact with predators. They document the chat transcripts and analyze the chat messages. Similar approaches include automatized software programs like 'ChatCoder' or 'Negobot' that analyze chat transcripts for inappropriate language. [21] Once a chat partner is unmasked as a predator, they play along and document the conversation. The chat message logs, phone conversations and real life meetings become part of the evidence to convict the predators. Since June 2002, over 588 predators have been convicted of abduction and molestation. Several members also regularly monitor social media platforms, like Facebook or Myspace, and actively report suspected profiles to the platform administrators.

Since Nov 2004 the Perverted-Justice community has become widely known due to their participation in the Dateline NBC investigative news program 'To Catch a Predator'. In this reality show, sting operations were set up to expose, humiliate, and arrest online predators. Members of the vigilante online community lured predators through online chat forums by setting up decoys. Once the decoys gained the predator's trust, they sent them to an empty house, where another young girl and the host of the show, Chris Hanson, questioned the suspects before investigators arrested them. Between November 2004 and December 2007 twelve such sting operations were carried out, over 286 people were arrested, and 103 (36%) were pronounced guilty. However, in the case of 150 incidents (52%) charges were dropped due to lack of evidence. The payments made by NBC to Perverted Justice created conflicts of interest within the online community. Also, local police departments criticized the vigilante working methods of the television show, which transformed from 'news reporting' to a 'news-making' agency.

In Nov. 2006 district attorney Louis W. Conradt was suspected of being a child molester. According to Perverted-Justice's documented message log files, Conradt, posing as a 19-year-old university student, engaged in sexually charged online chats with a person using the alias of a fictional 13-year-old named Luke. [29] Conradt persuaded Luke to exchange nude photos and after two weeks of ongoing file exchanges, the NBC team brought in an actor to play the fictional character Luke over the phone. After one phone call Conradt stopped responding to attempts to get in touch with him, leading the producers of the show to call in the local police. The producers and local law enforcement raided Conradt's house, where Conradt shot himself. Patricia Conradt, sister of the deceased Louis Conradt, sued the NBC network. The case was resolved amicably in June 2008. Due to this incident, there was heavy criticism of the producers' methods – public shaming, punishment, and social control as media entertainment.

Human Flesh Search Engine - Identifying and Exposing Individuals

The movement called 'Human Flesh Search Engine' (HFSE) originated in China with early incidents dating back to 2006. [34] The term was translated from the Chinese words 人肉搜尋 (Ren Rou Sou Suo), which broadly refers to 'an act of researching information about individuals or any subjects through the often viral and impulsive online collaboration of multiple users'. [36] Actual people, rather than computer-driven online searches, demonstrating citizen empowerment and civil participation, power the massive collaborations. Through the use of social media platforms, the wider public is involved in the fight against illegal behavior. By using progressive and interconnected search methods the knowledge of thousands of humans is used to uncover 'the truth' and identify any illegal behavior on

the part of an individual or a company. In China netizens of the human flesh search movement are also tagged as 'Red Guards 2.0'. [23]

Once the angry mob is released, the exploitation of private information or the leaking of classified information about the accused individual is impossible to avoid, due to the large number of people involved. This information is based on speculation or other low quality information, resulting in wrong accusations, flaming, cyber-bullying or even issuing death threats to innocent people by a crowd-sourced justice-seeking cyber-mob. Most outcomes include public shaming, exposing private information like home and work address, personal photos or video files, DDOS attacks, shutdown of personal websites, unemployment, fines or arrest.

Recent incidents include accusing and casting suspicion on innocent people after the Boston marathon bombings. 4chan and reddit users created 'photo think tanks' and crawled through the photos that were released by the FBI. The FBI planned to crowdsource to be able to gather more photo and video material from the incident. This worked out well and thousands of photographs were submitted to the FBI. In a second step, the 'crowd' was asked to identify the suspects, but the crowd already started their own investigations: a whole subreddit called 'FindBostonBombers' was dedicated to finding the suspects (see Fig. 1).

Figure 1. Photo posted on Reddit showing potential suspects in the Boston bombing

The crowd used several online tools to compare images. They used the 'Exif-Data' provided by many files to locate the exact camera position when the image was taken. Normally law enforcement use software tools like 'CrowdOptic' to carry out this kind of mass image recognition.

Another way to test the authenticity of an image is to use 'reverse image search' engines, which specifically search for matching images rather than finding content according to keywords, metadata or watermarks. When an image is submitted, a digital fingerprint is created that is compared to every other indexed image. The accuracy of different engines and plugins varies, from finding exact matches to similar images, including those that have been cropped, modified, or resized.

The findings of the analyzed images were published and discussed in subreddits like 'FindBostonBombers' and others.

The description of the subreddit stated:
'This subreddit is a place for people to post images, links, and thoughts about the potential identities of those responsible for the bombing. HOWEVER, please keep in mind that most or all of the 'suspects' being discussed are innocent people.'

The crowd fueled rumours and speculations and targeted people carrying backpacks – non-white, innocent people like Salah Barroom or Sunil Tripathi, amongst others, were accused and became public enemies. Some social media accounts of potential terrorists were leaked, and the innocent suspects received threatening calls. Different news stations contradicted the 'online witch hunt' and other news reports in order to bring the angry mob under control. Still, people were afraid to go on the streets. [22]

When the police reported the Dzahar brothers to be the suspects, news media also reported private information like their Amazon wish list and their favorite videos from YouTube. [30]

Internet Haganah – Confronting Islamists and their Supporters

'Internet Haganah' (IH) is a 'global open-source intelligence network' and web platform dedicated to confronting 'Internet activities by Islamists and their supporters, enablers and apologists'. Haganah, meaning 'defence' in Hebrew, was also the name of the early Israel Defence Forces who protected Jewish settlers in Palestine. Back in 2003 Adam Weisburd started blogging about offensive and dangerous sites and founded the organization. Over the years, the community has managed to shut down several thousands of radical websites. [6] On their website, they provide forums covering several issues, where community members post and discuss their collected intelligence on topics such as Reds in China, Russia, North Korea, Left/Right or Nihilist Wingnuts, Global Islamic Revolution, Islamists, Hamastan or Israel. [35]

Once a suspected website is found, it is posted on the Internet Haganah forum, where its relevance is then discussed by the forum members. The group uses online translation tools to translate the website's content. Offline versions and screenshots of the website are archived and used as evidence. Online archives like the 'Waybackmachine' are used to see the website's history. This way, they can create a timeline of the website and see the past publications and latest updates.

Furthermore, background data like the 'Domain Registry Information' is acquired to contact US-based hosting providers of jihad-supporting websites. If the hosting provider refuses to take down the website from their servers, they file further reports to U.S. National Defense Complaint Centers and provide information in the form of press releases and news articles to their media network. Within the network, cases are documented where hosting providers wouldn't cooperate and take the websites off their servers. In one case volunteers from the Haganah community figured out the hosting provider's administrator's private cell phone number and started to call his phone and put additional pressure on him until he took the site down. [6] In summer 2014 their website www.internet-

haganah.com went offline. Parts of the forum can still be accessed by the Internet archives 'Waybackmachine'.

Projects from Journalism, Art and Activism

The following section presents four projects from the fields of art and journalism that use disruptive techniques and other hacktivist methods to communicate their political messages. The different projects, a net-art performance, political activism, video installtion and subversive journalism, were selected because they use disruptive methods and software tools similar to those also used by the various vigilante communities. In 2011, Ian Paul created ‘Borderhaunt’, a net performance piece, where he cross-references a surveillance network database with a border deaths dataset to create a haunted commentary on the US-Mexican border situation. Electronic Disturbance Theater is a cyber-activist group using different software tools to shut down banks or governmental institution websites. The video installation Password:****** leaks email passwords of Internet scammers and shows how social engineering tactics can be applied to ‘deceive the human’ rather than ‘hacking the system’. The last project is by Mads Brugger, who documented his investigative journalistic approach to uncover diplomatic corruption in the central African state of Congo.

Borderhaunt - Cross-checking Databases for an Artistic Net Performance

The artist and theorist Ian Paul created a net-art performance called ‘Borderhaunt – A Border Database Collision’ [28]. The online performance took place on July 15, 2011 and was an attempt to merge two different databases associated with the U.S.-Mexican border. 667 participants from over 28 countries collected entries from the database that holds the names and descriptions of people who died trying to cross the border territory. This database is initiated by the Arizona Daily Star, who started compiling border deaths that were recorded by medical examiners in an effort to present an accurate number of people who died in the course of their attempts to cross into the United States illegally through Southern Arizona. These deaths are either caused by extreme weather conditions, violence of vigilantes, or abusive law enforcement officiers. [4] These entries were then sent to a database of the blueservo network, a private service company contracted by the Texas Sherriff's Border Coalition which crowdsources surveillance of the Texas-Mexico border, creating reports of ‘suspected’ undocumented border crossings. Volunteer users of the database watch livestreams of the border and submit ‘suspicious activity’ once they see an illegal immigrant crossing the border. For this the Department of Homeland Security installed 25m tall observation towers equipped with long-range radar, high-resolution cameras and an underground sensor network. One observation tower can detect the slightest movement in a 10km range along the Mexican border. [20] As a result of the performance, the border was symbolically haunted for the duration of the one-day action as the border police received over 1,000 reports of deceased immigrants attempting to cross the border.

The action was conceptualized as a kind of collective online performance and intervention for one day by cross-referencing the ‘Border deaths database’ and the ‘Blueservo surveillance network’, which reflected on border crossing deaths as well as disrupting the surveillance technologies used in the border territory (see Fig. 2).

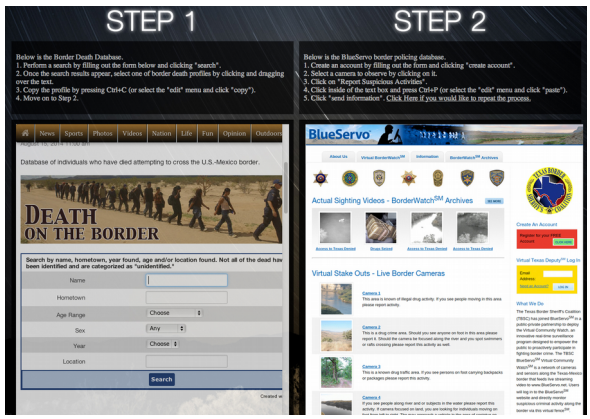

Figure 2. Screenshot of the project website: Step 1: Find a dead person in the database, Step 2: Report a suspicious activity in the Blueservo border policing database

Virtual Sit-ins: the Electronic Disturbance Theater

The Electronic Disturbance Theater (EDT) is a small group created by Ricardo Dominguez, Carmin Karasic, Brett Stalbaum, and Stefan Wray. These cyber activists and artists engaged in developing the theory and practice of ‘Electronic Civil Disobedience’. One of their software tools is called ‘Flood Net’, which is a URL-based software tool used to flood and block an opponent’s web site, enabling users to participate in collective electronic civil disobedience in solidarity with the Zapatista rebels of Chiapas (Mexico), a revolutionary group of indigenous people who were fighting against government oppression. [10] With these virtual sit-ins, members of the EDT slow down a website’s performance and drain the web server’s bandwidth until the website is extremely slowed down or even unreachable. [11] On April 10, 1998, they released a java applet called ‘FloodNet’ and performed non-violent actions against the Mexican president Zedillo’s website (’98 and ’99), several Mexican banks, the Frankfurt stock exchange, the U.S. Government and the Pentagon. The users were asked to create ‘bad URLs’, web addresses of nonexistent web pages at targeted sites, e.g. URLs that consisted of names of Zapatistas killed by the Mexican army. Each time such an website was requested, it was inscribed in the server’s error log. The Department of Justice counterattacked the EDT and destabilized the group’s infrastructure. Ricardo Dominguez, driving force behind the EDT group, claims that their actions are artistic experiments in ‘electronic civil disobedience’ rather than true acts of sabotage. By adopting the civil rights movement methods of ‘sit-ins’ to blockade the entrance of public buildings to block the Internet, they experimented with new ways to protest through the use of digital media. [13] In 1999, the group released the software to the public as part of the ‘Zapatista Disturbance Developer's Kit’.

The Video Installation 'Passwords: ****'**

The artist collective kairus.org referenced a scambaiting database, where activists publish scammers' usernames and passwords for their email accounts, and visualized popular passwords in a 6-channel videoinstallation. This sensitive data is gathered by using social engineering methods to persuade the scammers to share their login information. This can be done through the use of fake forms where the scambaiters ask for sensitive information that can reveal the scammer's email security questions, e.g. mother's maiden name or street adresses. Another method the scambaiters use is to offer a supposedly free web service to scammers. It is specifically advertised as a 'trusted and reliable infrastructure' that scammers can use for their businesses. The scambaiter sends out email formats of bulk messages in order to attract the interest of scammers to sign up for this service. During the application process, the scammer has to provide several alternative email addresses and a selection of passwords. Scammers who use several fake identities often use same or similar passwords for their email accounts. Once the scammer logs in to the newly generated account and tries to use the service for fraudulent activity, the email and password details are stored in a database. This database is shared amongst the scambaiting community to crowd-source the high number of scammers' account details. Scambaiters are asked to log in to the scammers accounts and to document criminal evidence. Often you can find fake documents, login information for other online services or gang communications. Scambaiters read through the emails and warn potential victims not to believe the rogue business and to stop communicating with the scammer. They continue monitoring the scammers account until the scammer loses interest and abandons the account. This makes it possible to learn from the scammers' practices and demoralize their attempts to gain any money from people who are ready to pay. The illegal act of accessing another person's account is justified by the efficiency of warning victims and collecting intelligence by accessing a criminal's 'virtual desktop', where important documents or passwords for other services can be found.

While looking at and experiencing the 6-channel video installation, the visitor reflects on issues of online security and questions one's personal password usage (see Fig. 3). The artwork stresses the 'online common sense' that passwords can be hacked as a result of security flaws like 'Heartbleed'; they can also be obtained by social engineering techniques. Securing personal data online with a strong password and constant security updates to avoid exploits is essential. However, people are still lax when it comes to securing their passwords and not sharing them with others. [39]

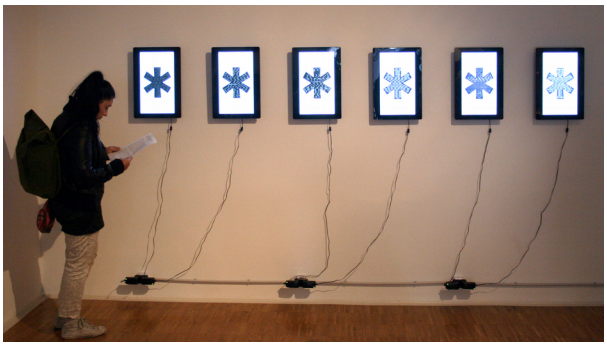

Figure 3. Exhibition setup 'Passwords:******'

Performative Journalism: Mads Brugger - The Ambassador

In his documentary film 'The Ambassador', Danish provocateur, journalist and filmmaker Mads Brugger impersonates a Liberian ambassador called 'Mr. Cortzen', who goes to the Central African Republic of Congo to expand diplomatic relations. He is able to buy a valid diplomatic passport from the state of Liberia over the Internet. Under his new name 'Mr. Cortzen', he is able to enter Liberia and establish diplomatic relations with other state diplomats. [15] His official agenda is to represent the state of Liberia, with a fake diplomatic passport, and set up businesses as a cover story, e.g. a match factory to employ a local tribe of pygmies. His second agenda is to provide an insight into corrupt politicians and to uncover the ongoing blood diamond trade. With hidden cameras he documents how he bribes his way up the social ladder and engages with government officials and other diplomats. Over-exposing his superior status as a white, ginger-bearded foreigner in a postcolonial outfit, Mads Brugger sees his production as 'performative journalism', uncovering the criminal potentials diplomatic immunity provides. In a fragile state like the Central African Republic of Congo, most white men have several hidden agendas, so he could avoid being questioned why he, as a white man, represents another African state. Because of the film, the Liberian press identified eight Mr. Cortzen-like diplomats in their corps. Today in many countries like Russia or China it is extremely dangerous for journalists to work and report from. Mads thinks it is necessary for journalists to use a new set of tools to research and report in such countries. [15] [17]

Discussion

There is an ongoing debate about the benefits these vigilante communities bring to net societies and law enforcement. Differences in the training of volunteers, the various state legislations and divergent sets of resources will undermine the communication process between the involved parties.

Vigilante groups invest a lot of time and commitment to their act of civil service. Members are often tech savvy and are open to sharing their findings with potential victims or law enforcement.

Could these vigilantes be used as a resource in the fight against cybercrime? Could a training by the police set certain standards and enhance the cooperation?

Since 1996 NGO's such as 'Ultrascan Advanced Global Investigations' (UAGI) operates by identifying, analyzing and predicting perpetrators of cross-border

fraud and the communications and support of terrorism from local or international religious extremists. [33] They offer a six-phase volunteering program, where helpers are coordinated to warn scam victims (phase 1), help them to file complaints (phase 2), visit working offices of scammers (phase 3), collect sensitive information (phase 4), report to the police (phase 5), and collect intelligence to get the scammers arrested (phase 6). Similar attempts to include civilians and private organisations to cooperate with law enforcement in the fight against cybercrime are undertaken by NC4 Cybercop or Project Vigilant. [2] [16] Such programs distribute the duty of policing and empower citizens to fight Internet crime.

Conclusion

By observing the working practices of the different online vigilante communities it is possible to map their working practices and the tool-sets they use that empower them to prosecute their 'opponents'. In general, these vigilante communities are very concerned about their anonymity and use fake profiles to camouflage their digital identities. Digital identity can be simply defined as the digital information that creates the image of an individually identifiable person. The groups use and misuse various vernacular online tools to gather intelligence. In the last few years, more and more artists have used net-activist tools for producing their artworks. This merges the activists, hackers and media-art movements into new genres, often referred to as tactical artists, hacktivists (hacker and activist) or artivists (artist and activist) cultures.

'Scambaiters' and members of the 'Internet Haganah' group use common practices to obtain background checks on the hosted websites to figure out their registration date, track down the administrators and get a physical address and phone number of the webmaster. Also, that way it is possible to obtain information regarding who the hosting provider of the website is and if local state laws or the hosting companies' 'Terms & Conditions' can be applied to the case in order to force the hosting provider to take the website off their servers. The activists use social media platforms, blogs and press releases to inform the public about their ongoing investigations and try to draw the public's attention to the case. They often cooperate with local NGO's that warn potential victims and extend their outreach.

Members of the 'Scambaiting' community use several online tools to create fake characters, track email IP-addresses or use image analyzers to extract Exif metadata from images. Software tools like 'Muguito' or 'Lad Vampire' are used for 'bandwidth hogging', reducing the server's capacity and limiting access to the website for potential victims. 'Scambaiters' and 'Perverted-Justice' communities use fake profiles to hide their identities and create 'honeypots' for online criminals and groomers. By using special software tools like 'ChatCoder' or 'Negobot' they can analyze chat transcripts for predatory language and distinguish faster between a potential criminal and a regular chat-forum user. The 'Human Flesh Search Engine' and the 'Internet Haganah' are heavily crowd-sourced sting operations, where lots of members are engaged in a single case, e.g. identifying suspects in the Boston Bombings or collecting evidence to shut down Jihadist websites. This massive user-driven approach to data gathering, analyzing and filtering information is discussed in forums where adaptive prosecution methods are also evaluated. In this process ethical discussions are often forgotten, and no ethical group guidelines are defined; common moral sense is thrown overboard. The act of online humiliation through offensive text messages or photo-collages and exposing sensitive data like phone numbers, private address, or occupation is a common form of self-justice. The viral prosecution can result in cyber-bullying, prank calls, physical harassment, and death threats, often also targeting the accused's friends, family members, or co-workers. These kinds of harassment and the pressure of increased media attention have also sometimes led to the loss of social status, e.g. study status or employment. As social media profiles are shut down in order to not provide a platform for harassment and bullying, individuals often find it impossible to give a statement in their defense. Once the Internet has found a victim, it becomes hard to counter any false accusations.

Collected evidence that is gathered through background checks or by documenting communication with the victim could be tainted and become unusable in court, or targets could be condemned as guilty when innocent, said Paul Kurtz, the executive director of the Cyber Security Industry Alliance, a coalition of chief executives of tech companies. 'When we all become "law enforcement officers", justice becomes very blurry.' Individuals and U.S. officials think that they can learn more about online criminals or terrorist operations by monitoring suspicious sites, which are operational. They can obtain background information that law enforcement cannot gather. Often, evidence is either gathered illegally or by morally questionable acts like hacking or social engineering. Nevertheless, every case has to be analyzed separately: how the provided evidence is gathered, whether it can be evaluated by law enforcement, or if it just interferes with their investigations.

References

1. AA419, 'Artists Against 419', http://wiki.aa419.org
2. Albertson, M. (2012). 'Secretive group expands role in cybermonitoring', http://www.examiner.com/article/secretive-group-expands-role-cybermonitoring
3. Arizona Daily Star, 'Border deaths database', http://azstarnet.com/online/databases/border-deaths-database/
4. Associated Press, (2005), 'Immigrant protests border volunteers'actions', http://www.nbcnews.com/id/7424693/ns/us_news-security/t/immigrant-protests-border-volunteers-actions/#.U1jSWMd4HLc
5. Bazzichelli, T., *'Networked Disruption', Rethinking oppositions in art, hacktivism and the business of social networking*, 2013, Digital Aesthetics Research Center, Aarhus, Denmark.
6. Cha, A., 'Watchdogs Seek Out the Web's Bad Side', http://www.washingtonpost.com/wp_dyn/content/article/2005/04/24/AR2005042401062.html
7. Chandler, J. (2006), 'Robin Hood: Development of a Popular Hero', http://d.lib.rochester.edu/robin-hood/text/chandler-robin-hood-development-of-a-

popular-hero
8. Cheong, P. H., & Gong, J. (2010). ‘Cyber Vigilantism, Transmedia Collective Intelligence, and Civic Participation’. *Chinese Journal of Communication*, 3(4), 471-487.
9. Conway, M. (2007). ‘Terrorism and Internet Governance: Core Issues’. In *Disarmament Forum* (Vol. 2007, No. 3, pp. 23-34). United Nations.
10. Dominguez, R., ‘Zaps Net Interface’, http://www.thing.net/~rdom/zapsTactical/zaps.html
11.EDT-Medienkunstnetz, http://www.medienkunstnetz.de/werke/flutnetz/
12. Egan, V., Hoskinson, J., & Shewan, D. (2011). ‘Perverted justice: A content analysis of the language used by offenders detected attempting to solicit children for sex’. *Antisocial behavior: Causes, correlations and treatments*, *20*(3), 273.
13. Electronic Civil Disobedience, http://www.thing.net/~rdom/ecd/ecd.html
14. Ethington, P. J. (1987). ‘Vigilantes and the Police: The Creation of a Professional Police Bureaucracy in San Francisco’, 1847–1900. *Journal of Social History*, 21(2), 197-227.
15. Hogan, Michael. ‘Mads Brugger, “The Ambassador” Director, Takes Exploitation To A Whole New Level’, http://www.huffingtonpost.com/2012/08/29/mads-brugger-the-ambassador_n 1840044.html
16. Huey, L., Nhan, J., & Broll, R. (2012). ‘“Uppity civilians” and “cyber-vigilantes”: The role of the general public in policing cyber-crime”. *Criminology and Criminal Justice*, 1748895812448086.
17. Isherwood, J., ‘Liberia to sue The Ambassador’, http://politiken.dk/newsinenglish/ECE1705916/liberia-to- sue-the-ambassador/
18. Kohm, S. A. (2009). ‘Naming, shaming and criminal justice: Mass-mediated humiliation as entertainment and punishment’. *Crime, Media, Culture*, 5(2), 188-205.
19. Kringiel, D. (2012), ‘Echte Superhelden - Der Maskenmann von nebenan’, http://www.spiegel.de/einestages/real-life-superheroes-echte-superhelden-a-947635.html
20. Kroft, S. (2010), ‘Watching the Border: The Virtual Fence’, http://www.cbsnews.com/news/watching-the-border-the- virtual-fence/
21. Laorden, C., Galán-García, P., Santos, I., Sanz, B., Hidalgo, J. M. G., Bringas, P. G. (2013, January). ‘Negobot: A conversational agent based on game theory for the detection of paedophile behaviour’. In International Joint Conference CISIS’12-ICEUTE´ 12-SOCO´ 12 Special Sessions (pp. 261-270). Springer Berlin Heidelberg.
22. Lindsay, J., Salah Eddin Barhoum, ‘Boston Teen Stunned To Be Portrayed As Bombing Suspect’, http://www.huffingtonpost.com/2013/04/18/salah-eddin-barhoum-bosto_n_3112892.html
23. Mac Kinnon, R., ‘RConversation: From Red Guards to Cyber Vigilantism to Where Next?’, http://rconversation.blogs.com/rconversation/2009/02/from-red-guards-to-cyber-vigilantism-to-where-next.html.
24. Miller, R., (2005), ‘Four Sites Targeted by Mugu Marauder Now Offline’, http://news.netcraft.com/archives/2005/02/28/four_sites_targeted_by_mugu_marauder_now_offline.html
25. Musgrave, J. (2010). ‘Cybercop earns fame but fuels skepticism’, Palmbeach Post online news, http://www.palmbeachpost.com/news/news/cybercop-earns-fame-but-fuels-skepticism-1/nL9HJ/
26. McNeal, G. S. (2006). Cyber Embargo: Countering the Internet Jihad. Case W. Res. J. Int'l L., 39, 789.
27. Nakamura, L. (2014). ‘“I WILL DO EVERYthing That Am Asked”: Scambaiting, Digital Show-Space, and the Racial Violence of Social Media’. *Journal of Visual Culture*, 13(3), 257-274.
28. Paul, I. (2011), ‘'Borderhaunt’, http://www.ianalanpaul.com/borderhaunt-2011/
29. Perverted Justice Archives, ‘The full evidence regarding Louis William Conradt, Jr.’, http://perverted-justice.com/?archive=Inxs00
30. Sheets, C., ‘10 Boston Marathon Bombing “Suspects” 4chan and Reddit Found”, http://www.ibtimes.com/10-boston-marathon-bombing-suspects-4chan-reddit-found-photos-1199213
31. Tangen, P., ‘The Real Life Superhero Project’, http://www.reallifesuperheroes.com/
32. Tuovinen, Lauri, et al. ‘Baits and beatings: vigilante justice in virtual communities.’ Proceedings of CEPE (2007): 397-405.
33. Ultrascan Advanced Global Investigations, http://www.ultrascan-agi.com/public_html/html/419_volunteers.html
34. Wang, B., Hou, B., Yao, Y., & Yan, L. (2009, October). ‘Human flesh search model incorporating network expansion and gossip with feedback’. In Proceedings of the 2009 13th IEEE/ACM International Symposium on Distributed Simulation and Real Time Applications (pp. 82-88). IEEE Computer Society.
35. Weisbud, A., ‘Internet Haganah’, http://internet-haganah.com/haganah/
36. Yi-Tsen, L. (2011) ‘The power of the online public opinion. A case study on Human-Flesh Search’.
37. Zhuo-chao, Y. (2009). ‘The legitimate limit of human flesh search engine’. *Public Administration & Law*, 7, 039.
38. Zingerle, A., & Kronman, L. (2013, October). ‘Humiliating Entertainment or Social Activism? Analyzing Scambaiting Strategies Against Online Advance Fee Fraud’. In *Cyberworlds (CW), 2013 International Conference on* (pp. 352-355). IEEE.
39. Zingerle, A., ‘How to obtain passwords of online scammers by using social engineering methods’. In *Cyberworlds (CW)*, 2014 International Conference on Cyberworlds. IEEE.

Authors Biography

Andreas Zingerle is a media artist from Austria. He works as a research assistant and PhD candidate at the ‘Time-based and Interactive Media Department’ at the University of Art and Design in Linz, Austria. He is researching anti-fraud strategies and demonstrates how they work in interactive narratives, artistic installations and creative media workshops. In the past he worked on art installations exploring creative misuse of technology and alternative ways of Human Computer Interaction. Since 2004 he takes part in international conferences and exhibitions, among others Ars Electronica, ISEA, Siggraph, Japan Media Arts Festival, File, WRO Biennale. Together with his partner Linda Kronman he started an artistic collaboration called 'Kairus'. More info: andreaszingerle.com, kairus.org.