Preface

Internet trolling is one of the fastest spreading pieces of computer jargon of the 21st century. Barely a day goes by where trolling and “trolls” are not in the news. Internet trolling has come to refer to any form of abuse carried out online for the pleasure of the person causing the abuse or the audience to which they are trying to appeal. The simplest definition of Internet trolling is the posting of provocative or offensive messages on the Internet for humourous effect. The word ‘troll’ when used to refer to persons who try to provoke others originated in the US military in the 1960s prior to the realisation of the Internet for mass communication, with the term, *trolling for MiGs*. This was reputed to be used by US Navy pilots in Vietnam in their dog-fighting, popularised by the film starring Tom Cruise called Top Gun. Such a practice, of trying to provoke the opposing fighter pilots was not an authorised operation, but was defended by pilots in order to identify their “strengths and weaknesses.” Following this military use of the term trolling, the advent of the Internet soon led to similar phrases like *trolling for newbies*, which became a common expression in the early 1990s in the Usenet group, alt.folklore.urban. One of the first recorded attempts to define trolling was in the mid-1990s with the launch of the book, *netlingo*, which is now also online where one can see their definition of trolling. Others have described a Troll as someone who mostly initiates threads with seemingly legitimate questions or conversation starters, but with the ultimate goal of drawing others into useless discussions.

We now know something about what Internet trolling is and where it came from, but what do we do about it? This book starts the process of answering this question. The first section on social, legal and ethical issues in relation to Internet trolling gets the ball rolling.

SOCIAL, LEGAL, AND ETHICAL ISSUES IN INTERNET TROLLING

Sutirtha Chatterjee’s chapter on ethical behaviour in online environments is particularly poignant in the discussion of Internet trolling. Sutirtha describes a number of unethical behaviours that inhibit participation in online communities. This includes flaming, swearing, insults, and deception. Such behaviours are now grouped under the header of Internet trolling. Even though flaming is the posting of offensive messages online, they are not always as negative as one might first think. For instance, there are users known as Troller Rollers, who are trollers who troll other trollers. Sutirtha shows how flame trolling, which is the posting of messages which are not only offensive but provocative, is caused by a process called deindividuation. In this process, a person, such as a participant in an online community, becomes estranged from the mainstream members and this results in misbehaviours from them. Many of the mainstream trollers that are talked about in the media have been men in their 20s, who more generally have been hit
the hardest by the lack of jobs due to the global recession and pension crisis. Serial flame troller, Sean Duffy, for instance was known to have difficulties in social interaction with others and being unable to develop social relationships among his peers and others. Two of Duffy’s most prominent victims were the families of Natasha MacBryde, who died at 15, and Sophie Taylor, who died at 16. In the case of MacBryde, Duffy created a YouTube video called “Tasha the Tank Engine” which had MacBryde’s face superimposed on a locomotive. Disturbed by Duffy’s actions, MacBryde’s father said, “This person was hiding behind a computer. For me you can’t see him, you can’t do anything. It is very hard for a father. You all try and protect your kids.” The next chapter in this section by Calin Gurau discusses how the rules of online communities like message-boards often explicitly prohibit actions like “flaming” and “trolling,” particularly where they start “flame wars” which is a barrage of abuse between members of online communities in a particular thread. Sometimes these are started by members of a website called Elders who are out-bound from the community, who “troll for newbies,” by posting messages that will obviously incite a flame war from unsuspecting new members.

Calin’s chapter also discusses how communities based around anonymity, such as the community called 4chan which has members who abuse others and inaccurately call themselves “trolls” is a perfect example of this. Calin says how users are often asked to refrain from providing information that might lead to them being identified. Indeed, in online communities like Wikipedia there is a detailed policy on Outing, where one can’t reveal the identity of another user. This often results in another sin at Wikipedia, Conflict of Interest, and the two are rarely reconciled, resulting in flaming, like that seen on the Talk Page of the article about Barry Wellman displayed in Figure 1 where a user accuses the editors of the article of using it to promote the subject they are alleged to know.

Figure 1. An example of flame trolling on Wikipedia resulting from perceived conflict of interest
Calin’s chapter is followed very appropriately by Georgios Michaelides and Gabor Hosszu’s chapter on privacy and security in virtual worlds and the threats to look out for. They raise important issues in the study of Internet trolling, including cyberstalking and cyberbullying, and the effect availability of personal information in online communities and on social networking services can be used to aid Internet abuse, including by flame trollers. The victims of these trollers, who are often variants of “Snerts” (i.e. Sexually nerdish egotistically rude trolls) if they have a grudge against a general group that person is in, or “E-Vengers” who have a grudged against them personally. Herogios and Gabor explain the effects Snerts and E-Vengers have on users of online communities, including loss of self-esteem, increased suicidal ideation, and feelings of being scared, angry, frustrated, and depressed. A notorious flame troller who carried out such abuses over a four-year-long period against a number of victims was Maria Marchese of Bow, London. Her campaign as a Hater troller resulting in the break-up of an engagement of a respected psychiatrist, Dr. Jan Falkowski following allegations of rape he was exonerated from with forensic evidence. She also drove an accountant, Miss Deborah Pemberton to the brink of suicide telling her to “dig her own grave.” Judge John Price at Southwark Crown Court sentenced Marchese for three counts of harassment under the Protection from Harassment Act 1997, including a restraining order. He told Marchese, “It is difficult to imagine a more serious case,” as her “uniquely disturbing” obsession as a cyberstalker caused “enormous pain and suffering.” Whilst they are at risk of ‘Reporter Trollers’ who report any content they disagree with, Georgios and Gabor recommend in their chapter using “Report Abuse” buttons in the same way one would normally use “Contact Us.”

Georgios and Gabor’s chapter leads suitably on to cyberstalking more specifically, with Alok Mishra and Deepti Mishra taking the problem head on by identifying it not only as a social problem, but a risk to web security more generally that needs to be taken seriously. Alok and Deepti concur with the findings of Georgios and Gabor about the long-term damaging effects cyberstalking has on the victims, who are mostly young women, they find. This is certainly evident in the news coverage on Internet trolling in general. Colm Coss was one of the first to be sentenced for Internet trolling for the posting of offensive content on a memorial website of celebrity and cancer victim Jade Goody, aged 28 when she died. Coss defaced the memorial page of former Big Brother star Jade Goody, becoming one of the first recorded cases of “RIP Trolling.” Coss’s actions shocked the nation, where he was reported to have ransacked memorial pages, including with boasts that he’d had sex with dead bodies. Manchester Magistrates Court was presented with clear and convincing evidence in the form of photographs that Coss shows neighbours as proof of him being a “troll.” Magistrate Pauline Salisbury said to Coss, “You preyed on bereaved families who were suffering trauma and anxiety,” and “We know you gained pleasure and you aren’t sorry for what you did.”

PSYCHOLOGICAL AND WELLBEING ISSUES IN INTERNET TROLLING

It is clear from the first section that as well as a number of legal, ethical, and security issues, the outcome of the flame trolling has significant effects on the psychology and wellbeing of the victims which is explored in the second section, along with chapters that seek to understand the trollers themselves. The first chapter in this section presents an ethnographic study that identifies the different troller character types that become apparent in online communities where trolling occurs within. It followed a single case study by early career researcher Whitney Phillips of the 4chan website whose members identified personally as trolls to justify their abusive behaviour for their own entertainment. An interview with Phillips,
who wrote that case study for an online journal, citing mainly hyperlinks, led to the convenient adoption of the term “troll” in the British media to refer to anyone involved in any form of abuse or harassment online, which then spread across the world. This chapter on the other hand shows that there is in fact a rich array of trollers, some of them who act provocatively to entertain others such as the Trolls and also those who act offensively for their own sick entertainment, like the ones interviewed on 4chan, who are primarily made up of the obnoxious type of troller known as a Snert, as they try to provoke humour like a Troll but in an offensive way for their own sick purposes. A depiction of the users on 4chan who call themselves ‘Anonymous’ is in Figure 2.

The next chapter in this section is by Wanda Cassidy, Karen Brown, and Margaret Jackson who seek out ways to address problems created by Snerts and other flame trollers in educational settings. They suggest it is not only important to try to curtail the flame trolling that forms part of cyber-bullying, but also to foster kudos trolling, through what they term, cyber-kindness. They show that through what has become known as Web 2.0, such as blogs and other social media websites, young people who are often both the victims and perpetrators of flame trolling can make positive contributions to Cyberspace to promote self-esteem and confidence in them and their peers. A notable example of such trolling is that Liam Stacey, who was 21 at the time of his flame trolling offence. Stacey was convicted for a racially-aggravated public order offence following a tirade of racist abuse on Twitter after he was rebuked by others after he mocked the cardiac arrest of Bolton player, Fabrice Muamba. Stacey was sentenced to 56 days in jail by district judge John Charles, who summed up saying, “It was racist abuse via a social networking site instigated as a result of a vile and abhorrent comment about a young footballer who was fighting for his life.”

This chapter is followed appropriately by one on online empathy by Niki Lambropoulos. Niki finds that trollers lack the same propensity towards forming factions as other participants and that those participants who empathises with others, such as having similar outlooks are most likely to participate. This might explain why online communities have traditionally been founded around mutual social ties, such as a topical interest or common geographical location. In this regard her finding that those most likely to troll are more remote from other users shows that flame trolling is more likely to occur where participants lack empathy with other members, which in the case of E-Vengers can be down to the community not tolerating their personality or other aspects of their identity in the first place. A notable example of such a flame trolling is that of Sean Duffy who first started trolling in his 20s, and is considered one of the

Figure 2. Anonymous trollers on 4chan justify their abuse as Snerts by calling themselves trolls (Courtesy: Wikipedia)
most prolific flame trollers to have surfaced on the Internet. He targets his victims, the family members of teenagers who have died, by posting offensive pictures and videos on their memorial pages, making him a *R.I.P. Troller*. Duffy has been sentenced to 13-week prison terms a number of times under the Malicious Communications Act 1988, but this has not stopped him from flame trolling his many victims.

The final chapter in this section, by Brian Whitworth and Tong Liu, takes the issue of empathy a step further. They argue that not only should users try to be more polite and considerate of others, but the software which they utilise should be designed for promoting sociability as much as traditional factors like designing for usability. “If software can be social, it should be designed accordingly” is how they put it. They argue that by designing out factors that promote anti-social behaviour typical in flame trolling, benefits can be achieved for all participants. Polite computing as they call it will, they argue, increase the number of legitimate interactions, including by kudos trollers, reduce anti-social attacks by flame trollers, and increase synergistic trade and software use. They present research which shows that politeness can prevent some of the worst types of flame trolling, like resentment where some are given greater opportunities to participate than others. Lack of politeness can breed mistrust and can sometimes result in claims of harassment, even if they are unfounded. One example is that of Angela Martindale, who was 39 at the time, was found not guilty of harassment at Prestatyn Magistrates Court following taking retaliatory action against former cagefighter Adam Finnigan. Finnigan told the court that he and his partner, Suzanne Rogovskis, had been bombarded with abusive text messages and voicemails between August and October 2010, following Martindale and Finnigan ending their relationship. Rogovskis told the court she recognised Martindale in one call shouting, “I’m going to rip your head off.” Clearing her of harassment, Magistrate Peter Oakley said there was no clear evidence that she had made the calls, and said the police had not investigated the matter fully.

**TRUST AND PARTICIPATION ISSUES IN WEB 2.0 SYSTEMS AT RISK OF INTERNET TROLLING**

This third section extends on the last section about psychology and wellbeing by looking at the related issues of trust and participation. It begins with a chapter by the editor looking at the role of defriending and gamification for increasing participation in online communities, where trolling and lurking are both commonplace. Trust is not just an issue in e-commerce based online communities, but in all of them. It is this lack of trust that can result in many people staying, or becoming, Lurkers –people who don’t feel they can fully participate beyond browsing. Fear of being flame trolled is one of the barriers to posting in an online community, and this needs to be overcome. Chapter 9 extends upon the character theory in Chapter 5 by introducing a twelfth type of character - the Elder. This addition is important as means one has to recognise that Lurkers are a type of participant - who take part in online communities for the enjoyment of surveillance. Elders are naturally their opposite, as they take part for escapism using their expertise to ‘troll for newbies,’ knowing they are on an outbound basis. The gamified flow of persuasion model in this chapter seeks to show the different transitional stages that a user has to go through to attain their optimal level of participation. In Web 2.0 systems based around technologies such as the Circle of Friends method used on Facebook and MySpace, online community participation can be dashed by defriending, which is where a friend stops someone access their page by deleting or blocking them from their buddy list. In the paper, the author presents an analysis of the typical things said on weblogs when someone has defriended another, to make it easier for both participants and systems operators to look out for the tell-tale signs of what can lead to defriending.
The next chapter, by Vanessa P. Dennen, continues the theme of understanding the narrative structure of weblogs and their impact on online participation. Vanessa looks at the unique, shared, and community problems that need to be addressed in order to negotiate a shared meaning and purpose on blog-based platforms. She finds that blogs are often based on a reflective process, like as discussed in the last chapter, but also they can go beyond this and seek advice. Such blogs are ripe environments for flame trollers who may seek to dash someone’s confidence when they are looking for sympathy and empathy for the situation they are in. The problems are not as apparent as in message boards because in many cases the owner of the weblog can delete comments in ways in which they could not if it was on such an open commenting platform. Vanessa also shows how it is possible that the visitors to blogs can reduce lurking and encourage kudos trolling by linking the different blogs they visit together, where others are going through a similar narrative.

In the next chapter, Miranda Dandoulaki and Matina Halkia look at the role Web 2.0 systems play in the reporting and misreporting of crisis situations. They show how the global village that exists through the availability of multiple mass media via the Internet means what happens in one part of the globe can have an effect in another part. A disturbing example of Internet trolling that has an effect on the Gaza and wider world is that of depictions of desecration of religious texts like the Qur’an. Whether they are videos of their burning or the posting of this text being vandalised, what might seem a joke or protest has resulted in unnecessary incitements of hate. Probably the most notable example of religious hate posted in the blogosphere was that by Paul Z Myers, pictured in Figure 3, who was a well-known Junior Lecturer that posted pictures on his website of banana skins over the Qur’an. Both this and the Gaza case presented by Miranda Matine demonstrate the influence of social media technologies on global events and in the last of the latter, even when the disaster or crisis area has huge inadequacies in technological infrastructure.

Trying to justify his actions, Myers said, “By the way, I didn’t want to single out just the cracker, so I nailed it to a few ripped-out pages from the Qur’an and The God Delusion. They are just paper. Nothing must be held sacred. Question everything. God is not great, Jesus is not your lord, you are not dis-

Figure 3. Paul Myers commits regular and serious acts of flame trolling (Courtesy: Wikipedia)
ciples of any charismatic prophet.” Myers’ university refused to take action, saying that as it was on his personal website they have no jurisdiction. Myers’ abuse towards women and people of a faith other than Atheism led to a whole raft of counter-trolling, where people posted equally abusive or insightful comments about him. One person who was banned from Myers’ free thoughts blog for free thinking made an entire video, posted to YouTube, showing a lot of compelling evidence and concluding that he had an “utter lack of academic integrity.” The whole episode shows how the effects of trolling go beyond the environment in which they originated.

In the following chapter, Paolo Massa describes how using reputation systems on blogging platforms can resolve such abuse by involving the website users. They talk about the Slashdot website and how a moderation point can be spent within 3 days for increasing the score of a comment by 1 point, choosing from a list of positive adjectives (insightful, interesting, informative, funny, underrated), for decreasing the score of a comment by 1 point. They can also choose from a list of negative adjectives (off-topic, flame-bait, troll, redundant, overrated) with moderation points being added or subtracted to the reputation of the user who submitted their opinion on a comment.

In the final chapter of this section, Eric M. Rovie continues the theme of trust, with particular reference to anonymity. Anonymity often allows users to create pseudonyms for the online community they are in to hide their true identity. A pseudonym acts to remove barriers to flame trolling as these trollers think they are safe from prosecution and detection. Eric points out however, that with the right technology or a court order the details are obtainable. This was found to be the case in Great Britain where a victim of obscene flame trolling, Nicola Brooks, brought a case in the UK High Court after she suffered “vicious and depraved” abuse following posting a comment on Facebook supporting a former reality TV contestant Frankie Cocozza when he left the X-Factor show the year before. The High Court, in the first judgement of its kind in the UK, ordered Facebook to reveal the identity of those anonymous users who flame trolled Brooks. Eric explains how trust is the key to any online transaction, especially on an e-commerce site where people have to trust another to fulfil their side of the bargain. Eric shows how people who do not wish to take part in give-and-take, which he calls Hobbesian Fooles, can remove a lot of the trust in an online community. Built around the Chatroom Bob character, these users get what they want from others, regardless of the consequences. In an e-commerce website, Eric suggests that guarantee schemes can act as a disincentive, as it will put opportunists off as they could lose any money they make and their trading status in general. Above all he suggests that removing anonymity like that seen in 4chan will help those who would otherwise not use the community feel confident in taking part by knowing the identities of the other participants. Eric’s chapter leads nicely on to the fourth and final chapter on possible solutions to reducing flame trolling and encouraging kudos trolling.

POSSIBLE SOLUTIONS FOR DEALING WITH INTERNET TROLLING

Finally, the fourth part of the book looks at what options are available for tackling the problem of flame trolling and encouraging kudos trolling. The first chapter of this section, by Chen, Ryan, & Olfman, takes psychological theories around social cognitive theory. They suggest that behaviour changes online, including learning to tolerate others, and taking part in kudos training involves four processes—attentional, retention, production, and motivational. They make a clear distinction between face-to-face behaviour modelling and online behaviour. To put their theories into action they present an experiment that assesses the effectiveness of an online behaviour management system on educational outcomes. They show that
comparable levels of behaviour change between participants in the online group as the offline group, providing a strong basis for the development of behaviour management systems to combat flame trolling.

Lakshmi Goel and Elham Mousavidin discuss the importance of knowledge management for designing sustainable communities online in the second chapter of this section. They argue strongly that moderation is essential to combating flame trolling. To achieve this, they argue that every online community needs a dedicated community manager with the authority and resources to manage and act for the benefit of the community. These managers, also known as systems operators, or sysops, should seek out independent experts to provide their experiences to the community in order to reduce flame trolling and general conflict. These are two among a number of recommendations that Lakshmi and Elham make, and their chapter is an important read for supporting the running of online communities so as to design out the likelihood of flame trolling.

The next chapter in this section is by Ross A Malaga, which looks at the role of Retaliatory Feedback in affecting behaviour in online communities. Retaliatory feedback can result in discouraging honest feedback and the unfortunate encouragement of flame trolling in response to honest feedback. Ross proposes an innovative solution to the problem, which is a feedback escrow. One of the most noteworthy cases dealing with the effects of retaliatory feedback and lack of reliable feedback was that of Shumon Ullah, who was jailed for 32 months at Burnley Crown Court for fraudulently receiving £38,500 in three weeks from eBay. Ullah built up a base of kudos from other uses by selling many affordable goods to become a “reliable trader.” He then used his neighbour’s details to “break out” and fraudulently offer electrical equipment worth nearly £230,000, but failed to deliver them. Customers started flaming him by giving him negative feedback, which resulted in fight flame with flame abuse from him, including claiming he was bankrupt. He was then reported by customers to the police who carried out what the presiding judge, David Heaton, QC, called a “a very comprehensive, time-consuming and no doubt expensive operation.” Sentencing him, Judge Heaton said it was necessary to “punish” Ullah in order to “deter” others. The case shows the importance of reputation systems, but like Alberto Ochoa, Julie Ponce, Alberto Hernandez, and Felipe Padilla say in the next chapter in this section, artificial intelligence can be used to pick up on the tell-tale signs sooner.

Alberto, Alberto, and Felipe present a novel multi-agent based system that can help detect and respond to flame trolling. They show that abusive behaviour need not always be about the posting of abusive message but social exclusion also. Called a Social Blockade, people, especially young people, have their contributions ignored, which can then have severe consequences. They could either become a Lurker and not participate, or become an E-Venger and wreak havoc in the name of justice to avenge the way they were treated. A prominent case of this kind was that of Gavin Brent from Holywell, Flintshire. When he was aged 24, Brent was fined £150 with £364 costs by Mold Magistrates Court under the Telecommunications Act 1984. The case came about after Brent suffered what he called “mis-treatment” by the police who refused to listen to his concerns, the court heard. The court was told that this instigated the online abuse by Brent towards police detective, D.C. Lloyd, resulting in Brent posting “menacing” messages against Lloyd, which Brent thought was a legitimate expression of “freedom of speech” in the circumstances. Brent’s messages against Lloyd included, “P.S. - D.C. Lloyd, God help your newborn baby” and were later removed. Alberto, Alberto, and Felipe discuss how multi-agent systems can intervene to resolve such postings, as by operating autonomously according to a fixed set of routines, they can help intervene to achieve desired behaviours.
In the final chapter of this section and the book, Zheng Yang and Silke Holtmanns look at trust modelling and management to enable the movement from social trust to digital trust. They introduce a comprehensive solution to reducing the barriers caused by flame trolling or fear of flame trolling called, trust modelling and management. They suggest that someone is most likely to gain a sense of trust from a system if they feel a sense of confidence, belief, gratification, and disposition. They argue that an online community that is most able to convey that it is able to show benevolence, motivation, honesty and faith is most likely to generate trust in participants.

CONCLUSION

It is clear for all to see that Internet Trolling is not going to go away, as while it is possible for the human race to exist, we will be seeking out ways to be humourous and mischievous to entertain both ourselves and others. Transgressive forms of humour, like that done by Snerts who abuse others for their own sick entertainment, are a reality for most types of online community. This book provides many ways for understanding and responding to these problems, varying from technical approaches to more social and behavioural interventions. It is clear that the impact of the offensive kind of Internet trolling – flame trolling – can have huge effects on the wellbeing and psychology of Internet users, which the mangers of online communities, known as sysops need to take more seriously. However achieved, online empathy will be essential to building a shared meaning about users of an online community, whether a traditional message board or more recent innovations like weblogs. As Internet trolling is likely to stay, practitioners should read this book to understand the problem more and the potential solutions available.

This book, with its descriptions and solutions for dealing with flame trolling in controlled environments, like e-commerce websites and learning and social networking platforms, will raise a number of other issues. For instance, based on the findings of the chapters in this book, organisations will have to ask about the effectiveness of their information security policies, and how to implement statute and case law for dealing with flame trolling. In organisations based around a strong Web presence they are going to have to review the extent and limitations of their webmaster’s powers over trolling (i.e., sysop prerogative) and potential new technologies for dealing with the problem. Organisations need to consider the experiences of those users who have been subjected to flame trolling in assessing consumer satisfaction as their brand will be as dependent on customer experiences on the Web as much as in the real-world settings.

The issues raised in this book will be challenging for political parties and politicians. For example, the use and misuse of “trolling” for electoral gain, and mishaps in electoral campaigns due to social networking trolling bloopers will become an important consideration that can’t be avoided. Technologies like those discussed in the final chapters will have to be taken seriously. Information systems, particularly on the Internet, will have to be designed so as to encourage kudos trolling and at the same time reduce flame trolling. The political, ethical, security, privacy, and legal issues discussed in this book will need to be put into practice. New guidelines will have to be developed in order for systems operators to exercise proper control over their members who want to take part in flame trolling, and the role of state intervention and regulation on flame trolling needs to be considered.

Jonathan Bishop
Centre for Research into Online Communities and E-Learning Systems, UK