Internet Filtering: A Solution to Harmful and Illegal Content?

Marie Eneman  
University of Gothenburg  
Sweden  
marie.eneman@gu.se

Abstract  
There is an increasing public concern about harmful and illegal content, such as terrorist and extremist content, child abuse material available on Internet. As a response, many countries have implemented Internet filtering as a tool to regulate certain types of harmful and illegal content on Internet. This paper will focus upon the use of Internet filtering for child abuse material. On the one hand, Internet filtering is described as a tool to regulate child abuse material, on the other hand Internet filtering is associated with concern for important ethical questions related to effectiveness and censorship strongly connected to important democratic values and rights such as freedom of expression, privacy etc. Therefore, this paper addresses the question how democratic societies can recognise, evaluate, and address ethical issues raised by the implementation and use of filtering systems to prevent and control the distribution of harmful and illegal content such as child abuse material.

Keywords: Harmful and illegal content, Technological regulation, Internet filtering, Child abuse material, Critical information systems research

I. INTRODUCTION

There is an increasing public concern about harmful and illegal content, such as terrorist and extremist content, child abuse material, available on Internet, which calls for further regulatory attempts [1, 27]. As a response, many countries have implemented Internet filtering as a tool to regulate certain types of harmful and illegal content on Internet [2, 3]. This paper will focus upon the use of Internet filtering for a specific type of harmful and illegal content, i.e. child abuse material. On the one hand, Internet filtering is described as a tool to regulate child abuse material, on the other hand Internet filtering is associated with concern for important ethical questions related to effectiveness and censorship strongly connected to important democratic values and rights such as freedom of expression, privacy etc. Therefore, this paper addresses the question how democratic societies can recognise, evaluate, and address ethical issues raised by the implementation and use of filtering systems to prevent and control the distribution of harmful and illegal content such as child abuse material.

The term child abuse material (sometimes also referred to as child pornography) refers to documented material depicting the sexual exploitation of children ranging from images of children posing (usually naked) to material portraying physical sexual abuse [4]. The availability of child abuse material on the Internet challenges traditional legal regulation and a response of this is the growing partnership between governments and the Internet industry [5]. Internet Service Providers (ISPs) have, in many countries, been encouraged by their governments to implement what is called voluntarily filtering systems in different forms such as ethical codes of conduct, self-regulatory measures, policies and practices [25]. There is an ongoing debate about the implementation and use of Internet filtering to regulate content on the Internet referring to that Internet should be open and vibrant [1]. A central argument against it is that it will not achieve its ostensible goal, namely the avoidance or reduction of harm to children while simultaneously setting the dangerous precedent of censorship. Another important counterargument is that filtering practises in most cases is organised in a non-transparent way and therefore not in line with accountability and transparency [24]. The rapid development and changes of technology is also a challenging dimension when designing regulatory initiatives [26].

As this paper will show Internet filtering is a complex and controversial issue, with arguments for and against. Most people agree that the law should regulate child abuse material, but there are critical voices raised about the use of technology (here Internet filtering) as regulator. The present paper takes a different perspective on the issue of Internet filtering. Rather than engaging in one of the main positions on the topic, the paper will recount these in order to ask a wider and more fundamental question. It will use Internet filtering as an example of what can be called an "emerging technology". Such technologies exist but have a large degree of uncertainties surrounding them. Filtering technologies in relation to child abuse material will be discussed as an example that is pertinent because of its timeliness in the light of current debates but also because of the multiple technical, legal, social and ethical complexities surrounding it. These complexities range from the protection of innocent children to the big question of the relationship between state and individual. While Internet filtering raises some specific questions, it can nevertheless be seen as a representative problem in that ontological and epistemological uncertainties concerning technologies in question as well as lack of agreement on underlying norms are typical for emerging technologies. In order to give an answer to the question what an appropriate way of addressing ethical issues of emerging technologies may be, the paper will use some of the ideas of critical social theory. Critical theory is characterised by its critical intention to promote emancipation [6]. This normative
premise of the approach establishes a close link to ethical questions. At the same time, critical theory provides a theoretical basis that claims to be relevant and implement changes. If this claim is valid, then critical theory could provide indications of how ethical issues of emerging technologies could be addressed, including the question of an appropriate way of dealing with the problems leading to Internet filtering.

II. INTERNET FILTERING

The term Internet filtering refers to a wide range of different techniques that can be used to block access of Internet content [5, 25]. The rationale to implement filtering mechanisms is to regulate users' access of certain content that for one reason or another is regarded inappropriate by the state or the organisation deploying the filtering [1, 8]. Content that is subject of Internet filtering ranges from terrorist and extremist, politics to sexuality to culture to religion etc [7]. Many countries are also blocking the intermediaries, i.e. the tools and applications that could be used to circumvent the filtering mechanisms [10]. A growing trend is to deploy filtering mechanisms to protect economic interests, which is illustrated by the blocking of low-cost international telephone services. This paper focuses upon ethical issues surrounded ISPs filtering of child abuse material. The main function of ISPs, although they cover different characteristics and services, is the commercial provision of Internet access services to users. The ISPs key role in enabling access to material provided by a third party has given rise to the question of whether they have, or should have, any liability in relation to the material they provide access to [7] This issue has been specifically highlighted in relation to the availability of illegal material such as child abuse material [5]. Many states, and national law enforcement agencies consider ISPs to be part of the distribution chain of child abuse material. Three common approaches for filtering practices are: (i) inclusion filtering, (ii) exclusion filtering and (iii) content analysis [2, 3, 5]. These different approaches can be used in combination to achieve desired effect. The first approach allows users to access web sites that have been checked and approved in advance. The inclusion filtering approach is very limited since it allows access to only pre-approved web sites and blocks all other content and is therefore not widely used. The second approach, called blacklisting, refers to the processes of blocking pre-identified web sites that occur on compiled lists. All other information, not found on a blacklist, is accessible. It should be noted that the creation of updated blacklists is a complex task, especially since content moves and Web sites change their IP addresses. When users attempt to access blocked content, they are subsequently blocked. There are various degrees of the transparency when using this technique. In some cases, the filtering system returns a 'blockpage' that informs the user that the content requested has been blocked [7]. Most ISPs use this method in their filtering practice of child abuse material. There are however cases where the blocking is deliberately hidden by the filtering regimes to disguise the fact that they are blocking web sites. The third approach, content analysis, is a growing filtering technique. The concept is to avoid pre-compiled lists and to focus on analysis of the requested content. One of the advantages with content analysis is that this approach distinguishes between different types of content as opposed to filtering entire web sites, it is however a complex and expensive method.

Arguments for and against the use Internet filtering

There are a number of arguments used in the debate about Internet Filtering. The main argument used to support implementation of Internet filtering is that certain content is harmful and/or illegal and that citizens should be protected from it [2, 5, 11]. There is a variety of harmful and illegal content that already is, or is suggested to be, subject to filtering systems. Child abuse material is commonly considered to be the most harmful content and is currently subject to organisational Internet filtering in many states [9]. Besides being harmful, many actions surrounding child abuse material are also criminal offences in most jurisdictions [3, 4]. Another area of concern, frequently used in the debate to justify Internet filtering, is the issue of national security and international terrorism and extremism [1]. After September 11, there has been an increased implementation of filtering and surveillance systems to counteract international terrorism [12]. Other topics that frequently are targets of Internet filtering are: terrorist and extremist, pornography, violent content, political contents, gambling, instructions how to commit suicide, copyrighted material, hate speech and racism [7]. Germany and France are also blocking websites with content denying the Holocaust or promoting Nazism. The main argument against Internet filtering is that it is a form of censorship that constitute a threat to important civil liberties particularly freedom of expression and privacy which are considered to be important foundations of democracies [9]. A related argument to this is that harmful and illegal content should be regulated by International Law or national legal systems [1]. There are also concern raised that Internet filtering threatens the end-to-end principle, which is considered to be the basic principle of securing Internet to be free, open and vibrant [3]. Another argument is that no state or organisation has managed to implement a perfect system [13]. There is today no example of implemented filtering systems that neither underblocks nor overblocks content. Furthermore, it is possible to circumvent existing filtering systems. There is a wide range of circumvention tools that can be used to circumvent filtering systems [11].

III. CRITICAL INFORMATION SYSTEMS RESEARCH AS THEORETICAL PERSPECTIVE

Definition

The term "critical" has many different meanings. This paper concentrates on a particular tradition, which is sometimes called "critical theory", "critical research" or “critical information systems research” (CISR). For the purposes of this paper we will concentrate on two main characteristics of CISR: first, its intention to change social reality and promote emancipation, which is the heart of the critical enterprise and second the ontological view of technology underlying the critical approach, which is important for the purpose of exploring issues of emerging technologies. We contend that CISR is most centrally characterised by one main feature, namely the intention to change social reality and promote emancipation, which sets it apart from other research approaches and traditions. Emancipation is an ongoing and central theme within critical research, which is relatively independent of particular issues, topics, theories, or
methodologies [14]. This links directly to the intention to facilitate change. Critical social theory aims to "bring about real change in the human condition" [6]. This intention to change reality is based on the recognition of the problems caused or perpetuated by the status quo, of structural contradictions and existing restrictions, oppressions, and domination. The starting point of the critical approach is thus not to gain an unbiased view of an external reality but the perception that the reality we live in is not perfect and can be improved [15]. There is thus an important aspect that refers directly to the researcher's motivation for doing research and his or her perception of the world. To put it differently, "critical stance is focused on what is wrong with the world rather than what is right" [16]. What is wrong about the world from the critical perspective is that human beings are not given the opportunity to live the best possible lives they could or to achieve their potential. This aim of promoting emancipation raises a range of questions and issues. The most obvious one is that of the definition of emancipation. The definition of emancipation alluded to in the preceding paragraph is somewhat problematic because of its open nature. There are many attempts to provide a more detailed definition [17, 18].

Given that the central aim of promoting emancipation may be practically problematic and epistemologically impossible to ascertain, critical scholars tend to focus on more modest goals which are nevertheless conducive to emancipation. In a first step they typically aim to lay the groundwork of social change by exposing the status quo from an orthodoxy position. Restricting and alienating conditions need to be understood in a first step [19] in order to then be exposed as such. An underlying assumption of this approach is that individuals, in many cases members of organisations, fail to see that alternative social realities are conceivable. Once these are pointed out, they are nevertheless capable apprehending them and using their newly found view of the world to develop to their own advantage. Exposing alienation, domination, and oppression implies a challenge that contains the opportunity to overcome them. Given the exposure of problems of the status quo, a range of possibilities of addressing these and changing social reality are conceivable. Many of these (e.g. political revolutions) raise problems in their own right and contain the seed to different forms of alienation and oppression. Critical scholars therefore usually aim to bring about change in a non-violent way that is sensitive to the perceptions and preferences to the people affected, without necessarily accepting the extant world-views. In practice this means that the aim of change is realised by contributing to "transformative redefinitions" [6]. Such transformative redefinitions that allow alternative views of social realities that are more emancipatory than the status quo need to be co-created with the individuals and organisations in question.

Internet Filtering as an Emerging Technology

CISR raises a number of theoretical and practical problems. The one problem that this paper addresses concerns the question how the emancipatory intention can be implemented with regards to technologies that are uncertain or emerging. Part of the answer to this question requires a more detailed understanding of the concept and ontology of technology that underlies the critical approach. The dominant ontological position adopted by critical scholars is that of social constructivism or constructionism [14]. The relevance of this ontological position is that it gives a foundation for the process of transformative redefinition. Only if reality is socially constructed, can there be a point in trying to develop alternative construction. At the same time, however, critical scholars realise that there are limits to the ability to construct and re-construct reality. Alternative realities are possible but they are constrained by accepted views and perceptions which are often difficult to change [20]. The constructionist ontology thus raises problems because it has to contend with the existence of an external reality, which may be open to social intervention but which also has a prima facie objective existence. This can raise significant problems for critical research because it limits the possibility of social change.

In the context of this paper one needs to ask whether Internet filtering technologies should be viewed as stable or not. The discussion of filtering technologies provided above as well as the majority of the arguments for and against their application and very much carried by an objectivist and sometimes even determinist understanding of such technologies or socio-technical systems. From the critical perspective one should note that this is a conceptualisation of technology that is theoretically problematic and therefore hard to support in practice. Developments of technologies over the last 20-30 years have shown that uses and implications of technologies are notoriously hard to predict. The path that led from the introduction of the world wide web to electronic commerce was short in terms of time but not foreseen by anyone. But even on a smaller scale, it is often very difficult to predict relatively obvious consequences of new technologies. A prominent example of this problem is that of privacy. Despite a wealth of privacy legislation in most industrialised countries, privacy issues remain central problems and no clear way of addressing them has been found. The UK government, following others, notably Canada, have recently recommended Privacy Impact Assessment [14] and it provides templates for doing them. Despite the obvious nature of these consequences and existing legislation surrounding them, privacy issues tend to remain unresolved to the point where a recent review suggested that even a considerable percentage of UK government databases are in breach of privacy legislation [3]. The point to note here is that it is very difficult to predict which actual consequences the different alternatives of filtering Internet connections with a view of preventing access to child abuse material will have. It is plausible to assume that the immediate aim of making such access more difficult for a majority of users will be successful. But predicting side effects and other consequences of the use of such technology is more difficult. In this particular case, an added difficulty is that of the potentially shifting moral perception of the underlying issues. Sexually abusing children is a moral taboo of the strongest order in most societies and this is unlikely to change in the predictable future. However, this general agreement on the moral evaluation of the crime per se should not lead us to ignore the fact that it is related to a number of grey areas. A primary example of such a grey area is the definition of a child where age limits are somewhat arbitrary and certainly subject to cultural differences [11]. A related problem is that of the ethical justification of the moral views in question. While there may be agreement on some moral issues, there may be much less agreement on the reasons for holding these positions. In view of these problems, one could argue that it is not appropriate to view Internet filtering and related regulatory activities as objectively given and described but to view them as emerging. They are subject to discourses, some of which this paper has recaptured. This
constructivist view of Internet filtering that results from taking a critical perspective raises the question how to evaluate the debate surrounding Internet filtering if it is based on ontological assumptions about the technology that may be untenable.

IV. DISCUSSION

Internet filtering can be described as an emerging trend. As indicated earlier, there are a number of legislative and voluntary processes underway that will increase filtering in many western democracies. Filtering is even more widely used in non-democratic states. This raises a number of interesting questions. Why do some countries filter and others not, even under similar legislation regarding child pornography? What do Internet filtering practices mean for the future of the Internet as well as the future of markets, social norms and modes of governance? This paper will not be able to give comprehensive answers to these and similarly complicated questions. What it sets out to do is give an indication how they are to be evaluated from a critical perspective and what the practical consequences of such a critical view could be. Before this is investigated, it is necessary to briefly assess the ethical implications of filtering.

Ethics of Filtering

Filtering regimes based on blacklisting requires the pre-identification of the material so that it is known to the system. The compilation of these lists is a challenge and complex task and involves weaknesses with the system, such as subjective and moral opinions among the persons creating the lists, the constant change of web content and IP addresses. One of the conditions for this system to work is the pre-identification that means that ‘new’ material that has not yet been blacklisted is accessible. It requires a lot of efforts to keep these lists updated. It should be noted that blocking mechanisms only block the users access to the material, in many cases the material remains available on the web site although the specific access point to the Web site has been blocked. This means it can be accessed when using a different type of technologies or ISP. Law enforcement agencies investigate the possibility of removing the content but in many cases, this is not possible due to juridical issues. Difficulties arise for example when a Swedish ISP block access to a website containing child abuse material and the server hosting this content is placed in another jurisdiction with different legislation regarding child pornography. This highlights the need of national and international coordination to achieve desired effectiveness with Internet filtering. A serious ethical issue of filtering mechanisms is the risk of overblocking. Underblocking can also be a problem with implemented systems, i.e. when the systems do not block all content that it is intended to do. Overblocking on the other side is when the filtering mechanisms block more that it should. Another important ethical issue with Internet filtering is the potential risk of ‘hiding’ the problem. In many cases the child abuse material is still there, it is only the access that has been blocked. Another potential risk in relation to this is the serious consequences for example researcher and journalists investigating issues that are subject of filtering. A further ethical issue is the basis of the debate about filtering and censorship. If filtering poses a threat to the free and open information environment, then the ethical value of free speech and freedom in general may have to be weighted against the ethical advantages of blocking access to child abuse material.

Emancipation

Current filter mechanisms do not seem to hinder persons who are intent on accessing child abuse material, the systems could however be used to prevent the public from getting in contact with child abuse material [2, 11]. Consequently, it could be argued that filtering mechanisms may have the effect of preventing potential offenders from starting to access such material. Regulation models that are designed to require extra steps for the users to gain access of child abuse material may prevent persons who may try and access this type of content based on curiosity. Finally, it could be argued that the use of filtering mechanisms may lead to a reduced demand of child abuse material and consequently reduces the victimisation of the abused child. If critical research is about changing the status quo with a view to promoting emancipation, then one should ask how Internet filtering can affect emancipation of different stakeholders. The most obvious answer to this is that children may be emancipated from primary or secondary abuse. Society as a whole can be emancipated from the fear of child abuse. Offenders may be emancipated from their problematic desires. The list of possible links to emancipation was elaborated on before [17, 18] and need not be reiterated here. What is important is how emancipation can be hindered by threats ideologies or hidden agendas.

Ideologies and Hidden Agendas

Critical researchers pay attention to ideology and hidden agendas. Language usage is a central theme when investigating ideological assumptions. One could question and reflect upon the current discourse of Internet filtering. The term censorship has been replaced by the term filtering. Although that filtering is a form of censorship, they carry different moral connotations. Censorship has primarily been associated with oppressive regimes, but today when western countries implement similar systems they are referred to as filtering. Another remarkable aspect that can be found in the discourse of filtering is the use of the term voluntarily. Available policy documents describe ISPs implementation of filtering systems as voluntary filtering, which also could be questioned since there are examples where governments have threatened to introduce a new legislation if the Swedish ISPs did not take their responsibility and implemented filtering systems. Another aspect that needs to be further investigated is the economic context of Internet filtering, which is considered central for the understanding of social phenomena [6]. The ISPs are currently using commercial software in their attempt to regulate access of child abuse material, which gives rise to valid concern. These actors have economic incentives to ensure that there is high demand for their products. Finally, filtering is clearly an exercise of state and political power. It can therefore be used to promote particular political agendas. This is undeniably often the case in non-democratic countries and arguably also sometimes to be observed in democratic ones.

Emancipation and Emergence of Technology

All the arguments above are united by the implicit assumption that the technology in question can be unambiguously described and the consequences of its use can be predicted. As
outlined earlier, CISR scholars tend to doubt this claim. Instead, the critical view of technology emphasises the local, contextual, and constructivist nature of technology. Taking this perspective serious means that one must move away from linear descriptions and predictions and attempt to question the sources and implications of such linear views. This perspective allows a different evaluation of the entire filtering debate that was recounted earlier. Instead of perceiving the debate to be one of protection of children and public order versus censorship and illicit use of power, it can be recast as one regarding different perceptions and predictions of technology. The use of filtering technologies is not clear-cut and predictable. This is not because of an inexact understanding of the technological components but rather because the nature of technology and its future consequences is emergent. Predictions are thus not exact scientific statements but value judgments. This view does not posit that there is a better way of anticipating technology and its use. It means that the whole debate around emerging technologies, including filtering technologies needs to be recontextualised. Contentious questions like Internet filtering are not about proving truth claims but about exploring possible options, considering possible futures and deliberating which ones of those one wants to achieve. This implies that the discourse in question is understood as having more levels of abstraction than the simple one in which the question is whether one position is true or another. These questions remain important and require technical as well as social, ethical, and legal expertise. It is important to see, however, that there are other levels of abstraction that allow for a different view of these discourses, for example the critical discourse questioning the very assumption of technical determinacy that tends to underpin both sides of the argument for and against filtering. And it is also important that all of these constitute societal discourses around technology and that these discourses have implicit or explicit ethical references to the way members of a society collectively want to organise their lives.

Practical Recommendations

The preceding section may be of conceptual interest but for practical political questions one can justifiably ask which practical recommendations arise from it. Critical research, despite its general aim to have practical consequences, cannot normally give algorithmic advice on how to address particular problems. In the light of what was said so far, it does become clear, however, that some practical advice can be given; at least in the sense that certain procedures are more amenable for the critical approach than others. The argument that even seemingly clearly defined and currently used technologies such as Internet filtering are better understood as emerging and uncertain implies a different view of discourses. Instead of fighting about absolute truth claims whether certain measures will have clearly defined consequences, it would be desirable to institute more reflective discourses. This means that the implications and assumptions about technology should be spelled out clearly and made part of the discourse itself. Furthermore, critical theory of technology [21, 22, 23] suggests that for reasons of epistemological clarity as well as ethical legitimacy decisions on technology should have a strong participative aspect. That means that discourses need to be institutionalised that allow all of the stakeholders, i.e. those who are affected by the technology to have a voice in making decisions about it [6]. One can ask how the critical view of modern technologies inform national and international policy makers with regards to filtering could be implemented. One example would be to allow discourses at the point of usage rather than concentrate on abstract principles. This would mean that technological devices are not used to cut discourses off but remain open to them. For example, instead of trying to render access to suspected child pornography sites impossible, it might be possible to require explicit identification and justification for the access of such sites. Justifications might be a mis-classification of the site or a legitimate desire to access material, for example for research or other legitimate purposes. This would mean allowing users to override technical limitations if they demonstrate that they have understood them. It would also imply that those individuals who choose to override concerns would have to be in a position to justify themselves or face appropriate sanctions. This is just one suggestion of what might be possible or desirable to do. The recommendation raises new questions and issues about technical implementation or institutional support. This is not a major problem from the critical perspective because it can be understood as a contribution to a discourse whose overall aim is to improve the way we use technology to improve our social lives. It is mindful of the fact that individual contributions to discourses cannot give final answers but must be measured from the perspective of their overall contribution to society.

V. Conclusion and Implications

Based upon the societal problem with harmful and illegal content available on Internet, this paper has discussed the use of Internet filtering as a regulative mechanism to control the access and distribution of child abuse material. The topic was chosen because it is a topical and illustrative example when using technology to address social problems. Internet filtering is being promoted by a number of states for a number of reasons. The most prominent reason given by western democracies is that it can prevent one of the crimes that these societies find most abhorrent, namely the sexual abuse of children. This paper has recounted the debate surrounding the advantages and disadvantages of Internet filtering. Using a critical lens, it has argued that these debates may be better understood if one moves away from the confrontational style in which both sides want to show that the other is misinformed or driven by doubtful motives. Instead, the paper suggested, a better way to understand the debate may be to question the underlying assumptions, notably those concerning the nature of technology. By suggesting that Internet filtering technologies are not stable but emerging and interpretively flexible, the paper has opened an avenue for a different conceptualisation of the technologies in question but also of the debates surrounding their use. Furthermore, drawing on critical ideas, the paper has suggested that a way forward may be to attempt to keep open discourses and provide stakeholders a way to contribute to their understanding of technology. This refers to high level political discourses as well as the ability to engage in discourses on a detailed and technical level. They are not easily applicable, they do not offer a guarantee of consensus or of a better use of technology. They are expensive because they require participation on a greater scale and they need to be linked to institutions in ways currently not clearly developed. Despite of these and other disadvantages, I
believe that the current paper has shown that a participative approach, even to such difficult and morally challenging issues as Internet filtering offer promise. The better understanding of the problem in question and the increased legitimacy of eventual decisions are worth paying the price. In addition, a critical and participative approach to technology embodies democratic ideals that are worth upholding in democratic societies, even if they require more complex and costly arrangements than less democratic alternatives. Finally, I want to highlight that to date in the debate about regulation of harmful and illegal content on the Internet, most attention has been focused upon the important and popular rights such as freedom of expression and privacy. These two rights have received a significant deal of attention in the debate of individual’s rights in our contemporary digital society. However, even though that these two rights are fundamental important rights for any democratic society and deserve to be protected, they should not automatically take precedence over all other rights such as the important right of the child not to be sexually abused in our society.

REFERENCES

[27] Age-verification for online pornography to begin in July – GOV.UK, available at: https://www.gov.uk/