Ethics Inf Technol (2016) 18:199–210
DOI 10.1007/s10676-016-9406-0

ORIGINAL PAPER

Doxing: a conceptual analysis

David M. Douglas[1]

Published online: 28 June 2016

Abstract Doxing is the intentional public release onto the Internet of personal information about an individual by a third party, often with the intent to humiliate, threaten, intimidate, or punish the identified individual. In this paper I present a conceptual analysis of the practice of doxing and how it differs from other forms of privacy violation. I distinguish between three types of doxing: deanonymizing doxing, where personal information establishing the identity of a formerly anonymous individual is released; targeting doxing, that discloses personal information that reveals specific details of an individual's circumstances that are usually private, obscure, or obfuscated; and delegitimizing doxing, which reveals intimate personal information that damages the credibility of that individual. I also describe how doxing differs from blackmail and defamation. I argue that doxing may be justified in cases where it reveals wrongdoing (such as deception), but only if the information released is necessary to reveal that such wrongdoing has occurred and if it is in the public interest to reveal such wrongdoing. Revealing additional information, such as that which allows an individual to be targeted for harassment and intimidation, is unjustified. I illustrate my discussion with the examples of the alleged identification of the creator of Bitcoin, Satoshi Nakamoto, by Newsweek magazine, the identification of the notorious Reddit user Violentacrez by the blog Gawker, and the harassment of game developer Zoe Quinn in the 'GamerGate' Internet campaign.

✉ David M. Douglas
d.m.douglas@utwente.nl

[1] Department of Philosophy, University of Twente, Enschede, The Netherlands

Keywords Doxing · Internet harassment · Anonymity · Journalism · Hate speech · Privacy

A spectre is haunting the Internet—the spectre of doxing.[1] Doxing, sometimes spelt 'doxxing' or 'd0xing', involves releasing someone's personal details onto the Internet in an easily accessible form. These details may include full legal names, residential addresses, unique identifiers for governmental records and services (such as social security numbers in the US), business records and documents, and personal photographs of one's self and loved ones. These details may already be publicly available, but in difficult to access forms or distributed across various sources that obscure them from casual discovery. These details might also be government, company, or organization records obtained via a security breach. In some cases, they might even have been obtained directly from the person herself, either willingly or unknowingly. Doxing can occur to anyone, from high-profile public figures to obscure everyday people. All that is necessary to become a victim of doxing, it seems, is to be of interest to someone else on the Internet.

There are various motives for doxing someone. It may be motivated by a desire to expose wrongdoing and to hold the wrongdoer to account. It may be used to humiliate, intimidate, threaten, or punish the identified individual. It is often a tool for 'cyber stalking', as the information may be released in a context that would cause a reasonable person to fear for her life (Citron 2014). It can also serve as a tool for Internet vigilantism, where those opposed to someone's actions retaliate by revealing her identity and personal information, leaving the victim open to public ridicule, harassment, and vilification (Solove 2007). And information released onto

[1] With apologies to Karl Marx and Frederich Engels.

the Internet is easy to access and difficult to remove: entering a doxing victim's name into a search engine may reveal her personal details and the abuse associated with the doxing attack for years. The potential for harm and disruption are obvious when a person's professional life and reputation depends on her visibility on the Internet (Citron 2014).

As the above suggests, doxing may have a devastating impact for its victims. It assists in harassing and stalking individuals, both physically and on the Internet (Citron 2014). Such stalking creates significant distress and increases the risk of physical harm, especially if the personal information is used to encourage others to abuse the victim. A parallel can be drawn with sexual harassment on the Internet. Mary Anne Franks (2012) lists three factors that contribute to the harm online sexual harassment causes: the harassers' anonymity, the amplification of the harassment caused by the accessibility of the harassing content which may encourage further harassment, and the permanence that results from the difficulty of removing harassing content from the Internet. Even if doxing is not used as a tool for sexual harassment, these factors also contribute to the harms of having personal information revealed on the Internet.

Despite these harms, doxing is sometimes presented as a tool of protest and for exposing wrongdoing. Corruption by Chinese government officials is often the target of the so-called 'Human Flesh Search Engine', composed of Chinese Internet users who search for and release evidence of private and public transgressions and wrongdoing (Gao and Stanyer 2014). For example, an investigation into two Chinese local government officials was launched after documents listing travel expenses for research trips to the US and Canada were anonymously released onto the Internet. These documents provided evidence that public funding had been used to pay for trips to tourist attractions (Gao and Stanyer 2014).

This paper is an attempt to untangle the intertwined concepts and issues raised by doxing. I present and justify the claim that significant differences exist between various cases of doxing that justify placing them into different categories. I call these categories deanonymization, targeting, and delegitimization. Deanonymizing doxing releases personal information establishing the identity of a formerly anonymous or pseudonymous individual. Targeting doxing discloses personal information that reveals specific details of an individual's circumstances that are usually private, obscure, or obfuscated. Finally, delegitimizing doxing reveals intimate personal information that damages the credibility of that individual. I use this classification to highlight the significant differences between three cases of doxing: the alleged identification of Bitcoin creator Satoshi Nakamoto, the identification of the notorious Reddit Internet forum user Violentacrez, and the harassment of several female game developers in the 'GamerGate' incident. I conclude that in cases where exposing wrongdoing is in the public interest, deanonymizing and delegitimizing doxing is permissible only to the extent necessary to reveal that wrongdoing has occurred. Using any form of doxing to humiliate or threaten the subject, and revealing more information than necessary to establish wrongdoing, is unjustified.

Defining doxing

The term 'doxing' comes from the phrase 'dropping documents' or 'dropping dox' on someone, which was a form of revenge in 1990s outlaw hacker culture that involved uncovering and revealing the identity of people who fostered anonymity (Honan 2014). The term is already prominent enough to be included in formal dictionaries. For instance, the Oxford British and World English Dictionary defines doxing as to "[s]earch for and publish private or identifying information about (a particular individual) on the Internet, typically with malicious intent" (Oxford Dictionaries 2015). As the Oxford definition suggests, doxing does not necessarily have to be motivated by malice. Several high-profile incidents of so-called 'doxing' involved journalists revealing the identities of formerly pseudonymous Internet identities (Chen 2012a; Goodman 2014). Despite this, doxing is a term with negative connotations: labeling these accounts as 'doxing' suggests that the journalists involved have acted wrongly in revealing personal information about a pseudonymous individual (Beaujon 2014). Examining the concept in more detail by considering the different kinds of personal information that may be released will help to determine whether doxing is necessarily or primarily a malicious act.

A more nuanced account of doxing can begin by considering what it actually establishes: it removes some degree of anonymity from a specific person. Marx's (1999) concept of identity knowledge offers a useful tool for this task. The seven broad types of identity knowledge Marx describes are listed in Table 1. Perfect anonymity, according to Marx (1999), is the inability to be identified according to any of these seven types of identity knowledge.

Being identified by some of these types will be greater threats to anonymity than others. For example, being identified as an adult male in a large European city does little to reduce my anonymity, as it does not easily allow someone to gain other types of identity knowledge about me. However, being identified by name and address makes maintaining my anonymity more difficult as this information can be easily used to establish other types of identity knowledge. Knowing my name allows someone to search

Table 1 Types and examples of identity knowledge [based on Marx (1999)]

Type	Description
Legal name	The name under which someone is known for official and legal purposes
Locatability	Information that reveals where someone lives or where she can be contacted personally, such as an address or a telephone number
Pseudonyms linked to name or location	A name or code that represents a single individual (such as a bank account number) in a system that is related to their legal name or some other potentially unique characteristic (such as an address)
Pseudonyms not linked to name or location	
(a) For policy reasons	(a) A name or code representing someone in a system that is not related to her legal name, such as an anonymized medical record
(b) Audience is unaware it is a pseudonym	(b) A name someone uses instead of her legal name as a disguise or for deception, such as an alias
Pattern knowledge	Someone who can be recognized by her regular public actions or habits, such catching the same bus every morning at the same time
Social categorization	Information that can be used to place someone into social categories (or stereotypes), such as physical appearance, accent, style of dress, and so on
Symbols of eligibility/non-eligibility	Possessing artifacts or knowledge that identifies someone as being entitled to particular privileges and treatment, such as a uniform, a password, or a train ticket

public records and databases (to say nothing of the Internet) for further information about me. Knowing my address allows others to encounter me in person and observe my movements, habits, physical appearance and characteristics. In Marx's classification, these observations establish *pattern knowledge* and *social categorization* identity information about me.

Using Marx's categorization, I suggest that doxing should be understood as releasing publically a type of identity knowledge about an individual (the *subject* of doxing) that establishes a verifiable connection between it and another type (or types) of identity knowledge about that person. The verifiability of doxing distinguishes it from other forms of exposure and publicity. As the origins of the term 'doxing' ('dropping documents' or 'dropping dox') suggest, it utilizes documentary evidence of identity knowledge.

Different types of identity knowledge are documented in different forms. Identity knowledge relating to personal details used for administrative purposes may be recorded in official records or documents, such as birth certificates, tax returns, employment records, and so on. Such documents may reveal legal name, locatability, and pseudonyms that are connected to an individual's name or location. Documents that describe unique characteristics possessed by an individual in a pseudonymous record that is unrelated to her name or location may reveal further identity knowledge if it can be cross-referenced with other information. This possibility exists where medical records are not sufficiently anonymized. Symbols of eligibility may document themselves (such as railway tickets) or may be documented through records of such symbols being granted to an individual, such as graduating from a university. Similarly, official documentation will exist for symbols of eligibility being withheld or taken from an individual.

Other types of identity knowledge, such as pattern knowledge and social characterization, are documented in other ways. Frequently updated location information, such as stored by mobile devices that record their location, may reveal an individual's daily routine, and so establish pattern knowledge about that individual.[2] Social characterization may be established through photographs and imagery recorded about a person and her behaviour. Such characterization will often be up to the interpretation of the observer, and may be misleading if the images are taken out of context or presented in a biased manner. This is especially the case with activities that are invested with social or symbolic significance, or which challenge entrenched beliefs and expectations. For example, images

[2] The possibility of revealing pattern knowledge is why mobile device metadata (information about its usage) is so sensitive. For an example of how metadata analysis can reveal pattern knowledge about an individual, see Ockenden and Leslie (2015).

of a woman wearing revealing clothing or expressing her sexuality may be used to mock or humiliate her for not conforming to traditional notions of female behavior and gender roles (Poole 2013).

Doxing should be distinguished from related concepts such as blackmail, defamation, and gossip. Unlike blackmail, doxing does not involve making a demand to the subject to prevent information being released. A blackmailer only releases information if the victim does not comply with the blackmailer's demands. While the threat of doxing can serve as blackmail, doxing itself is not blackmail.

Defamation also involves the public release of information with the intention to humiliate, threaten, intimidate, or punish the subject. However, for information to be defamatory it must reveal something damaging to the reputation of the person (or people) described. Doxing does not necessarily have to reveal something questionable or embarrassing about the person involved. As I will describe later, while one form of doxing aims to harm the subject's reputation, doxing itself does not necessarily involve releasing such information.

Finally, doxing differs from gossip (even malicious gossip) in that it relies on releasing actual (or believed to be actual) identity knowledge rather than suggestion, hearsay and innuendo. Bok (1989:93) defines gossip as "informal personal communication about other people who are absent or treated as absent" (numbering of features omitted). While doxing can be formal or informal (i.e. consist of official documents or records, or accurate informal accounts), it is the difference between communicating information *about* someone and communicating information *of* someone. To illustrate this with a benign example, consider the difference between claiming 'X wore a pink tutu at a funeral' and releasing a photo of X wearing a pink tutu at a funeral. The first is an instance of gossip, while the second is a form of doxing.[3] The photograph serves as documentation of the claim being made about X, and is evidence that can be verified. Under Marx's classification of identity knowledge, it is social characterization knowledge as it documents X's apparent disregard for social norms. Merely telling a friend about X's poor taste in funeral attire does not provide this documentary evidence.

The value of anonymity and obscurity

Before examining the different forms of doxing in detail, I will establish the value of what doxing endangers: the subject's obscurity and anonymity. Doxing undermines what Ruth Gavison calls "our concern over our accessibility to others: the extent to which we are known to others, the extent to which others have physical access to us, and the extent to which we are the subject of others' attention" (1980:423). The subject no longer controls some aspect of identity knowledge about her, which reduces her ability to decide what she reveals about herself and to whom she reveals it. This control is an important aspect of a person's identity. We reveal some aspects of ourselves to some people but not to others. Our relationships with others are shaped by what we choose reveal to them and what they decide to reveal to us. Our identities and the social value attached to them (i.e. reputation and public persona) are difficult to build and easy to lose. Even forfeiting some degree of such control is a way of establishing one's own identity. Choosing to publicly document one's experiences and movements are decisions individuals make about how they wish to present themselves to others. Influencing how others perceive you is a vital part of establishing who you are (and crucially, who you are *not*) as a person.

To further illustrate the value of anonymity, I again turn to the work of Marx (1999), this time for his list of the rationales for anonymity. These are listed in Table 2. There is much to say about the significance of each of these rationales and whether they should be accepted in all cases. For reasons of space and scope, however, I will only make a few general comments here.

As Marx's list suggests, anonymity and obscurity are both forms of protection.[4] It can disguise attributes that may prejudice how others receive someone's work and ideas, such as gender, race, ethnicity, or class. It is anonymity's protective value that makes doxing particularly harmful in Internet communication, as it removes the subject's anonymity without an equivalent loss of anonymity for the attacker.

Collecting different types of identity knowledge about an individual can be regarded as building a 'dossier' on that person.[5] Dossier building involves the "compilation, maintenance, use, and dissemination of personal information on individuals" (Reichel 1977: 265). The opportunities created by the Internet for gathering personal information have effectively democratized dossier building. Now almost anyone with the desire and the time to search for another's personal information has the tools and information sources available for her to do so.

[3] Specifically, this is a form of what I call delegitimizing doxing, as I will describe later.

[4] This protection can of course be abused, as 'poison pen' letters and anonymous inflammatory comments on the Internet demonstrate.

[5] I thank Michael Nagelborg for this point.

Table 2 Rationales for anonymity [based on Marx (1999)]

Rationale for anonymity	Explanation
1. Facilitating the flow of information	Encourages information to be disclosed where there may be risks and penalties associated with doing so (such as whistleblowing)
2. Obtaining personal information for research	Allows individuals to be honest in their responses without fear of being punished or stigmatized if the information became public
3. Encouraging attention to the message content instead of the messenger	The identity of the informer may prejudice the reception of the information
4. Encouraging reporting, seeking information, and self-help	Individuals can report activities or to seek out information without fear of being stigmatized or victimized if others knew they were seeking certain information or were reporting certain activities
5. Obtaining resources or encouraging actions that involve illegality	Encourages individuals to seek help for problems that are linked with illegal actions (such as illicit drug addiction) or hand in illegal items (such as amnesties for contraband goods)
6. Protecting donors or those taking controversial but socially useful actions	Encourages individuals to contribute goods or actions without fear of intimidation, harassment, or creating future obligations
7. Protecting strategic economic interests	Allows someone to interact in the marketplace without their identity affecting their transactions (such as being charged more if someone is known to be wealthy)
8. Protecting one's time, space, and person	Allows someone to maintain their isolation from unwanted attention or interruption by others
9. Aiding judgments based on specified criteria	Promotes the unbiased assessment of something without being influenced by the identity of those involved
10. Protecting reputation and assets	Prevents an individual's personal information being used by someone else to deceive and defraud others
11. Avoiding persecution	Allows individuals to avoid harms that may result from belonging to a particular group (such as belonging to a persecuted minority)
12. Enhancing rituals, games, play, and celebrations	Allows individuals to interact in particular contexts without affecting their status and relationships in other contexts
13. Encouraging experimentation and risk-taking	Allows individuals to experiment with different behaviours and actions so that they may explore different ways of living without affecting their current relationships, commitments, and reputation
14. Protecting personhood and autonomy in sharing information	Allows individuals to control who has access to information about themselves and what information they choose to share and when
15. Traditional expectations of anonymity	Individuals expect that certain interactions do not involve revealing personal information about themselves to others, such as paying for items with cash

Types of doxing

I propose categorizing doxing into three types: deanonymization, targeting, and delegitimization. Each attempts to remove or damage something different from the subject: anonymity, obscurity, and credibility, respectively. Each type of doxing also creates new possibilities to further interfere with the life of the person involved. Deanonymization makes it easier to obtain other types of identity knowledge about the subject, and so creates greater opportunities for the other types of doxing to occur. Whatever advantages or protection the subject sought to gain by seeking anonymity or adopting a pseudonym will be lost. Targeting doxing creates the possibility that future harassment may take a physical form, with the uncertainty and risks of harm that it brings. The subject may be harassed and inconvenienced by others using her personal information to impersonate her. Finally, delegitimization presents a motivation for carrying out harassment and potentially further doxing by detailing how the subject is somehow unworthy of respect. These categories are listed and summarized in Table 3.

I now describe each category of doxing in further detail.

Deanonymizing doxing

Deanonymizing doxing releases information that reveals the identity of the person (or persons) who has previously been anonymous or known by a pseudonym. It also covers instances where someone's identity is revealed publically regardless of whether she has deliberately sought to conceal her identity or not.

Deanonymization is the broadest of the three categories of doxing as it can affect every type of identity knowledge and negates every rationale for anonymity. Depending on the subject's rationale for anonymity and the type of

Table 3 Types of Doxing

Type of doxing	Description	Loss to the subject	Examples
Deanonymization	Reveals any kind of identity knowledge about a person	Anonymity	Revealing the legal identity of someone using a pseudonym
Targeting	Reveals information that allows an individual to be physically located	Obscurity	Revealing someone's home address
Delegitimization	Reveals information intended to damage an individual's credibility, reputation, or character	Credibility	Evidence of supposed immoral activity, hypocrisy, or willful deception

identity knowledge released, it may not cause significant harm to the subject, and there may be plausible public interest justifications for disclosing it. For example, there is at least a *prima facie* public interest justification for revealing someone's identity when anonymity or a pseudonym is being used to deceive others for personal gain (a con artist impersonating someone else to gain money or prestige, for example). Literary hoaxes are an example that I will return to later in discussing the potential justifications for doxing.

A famous instance of deanonymizing doxing is the reveal of the supposed identity of the person behind the pseudonym 'Satoshi Nakamoto'. Satoshi Nakamoto is the name adopted by the creator (or creators) of the Bitcoin crypto-currency (Nakamoto n.d.). The true identity of Bitcoin's creator is still uncertain. An article in *Newsweek* identified Dorian Satoshi Nakamoto as Bitcoin's creator, a claim he has repeatedly and consistently denied (Goodman 2014). The creator of Bitcoin has a number of clear rationales for anonymity: avoiding interference, protection, and a desire not to draw attention away from the creation itself.

Another example is the deanonymization of the notorious Reddit moderator 'Violentacrez', who was revealed to be Michael Brutsch by the blog *Gawker*.[6] Brutsch was a volunteer moderator who contributed to and oversaw various forums (or 'sub-reddits') on the Reddit website. Violentacrez was heavily involved in deliberately provocative sub-reddits such as 'creepshots' (which featured voyeuristic photographs of unsuspecting women) and 'jailbait' (which featured photographs of girls under the age of consent) (Chen 2012a). Brutsch claimed in a television interview that he treated his activities on Reddit as a game (Chen 2012c). This is the rationale of play from Marx's list of rationales for anonymity. Violentacrez was an example of an Internet 'troll': someone who deliberately flouts social norms and provokes others for her own amusement, often under some form of anonymity (Phillips 2012).[7] The pseudonym allowed Brutsch to entertain himself and others by deliberately offending people and breaking social taboos with the material he posted on the website. Protecting his reputation (and employability) is another important justification (and another of Marx's rationales for anonymity), and an accurate one given that Brutsch lost his job as a result of his legal identity being connected with that of Violentacrez (Chen 2012b).

Targeting doxing

Targeting doxing reveals specific information about an individual that allows her to be physically located. It reveals physical locatability (rather than communicative locatability, like a telephone number or email address) identity knowledge about the subject.[8] Targeting doxing increases the subject's physical accessibility by removing the obscurity surrounding where a person lives or works. Losing this obscurity makes someone more vulnerable to physical harassment because of whom specifically she is.

Targeting doxing often follows from deanonymization. As Marx's rationales for anonymity suggest, seeking anonymity is frequently adopted to reduce the risk of being targeted. The identity knowledge revealed through deanonymizing doxing makes it easier to uncover further identity knowledge, such as the subject's physical location and workplace.

The forms of harassment made possible by targeting doxing range from irritating pranks to physical assault (or worse). Relatively harmless but annoying pranks can range from calls from car dealers responding to supposed interest in a car to having to cancel unwanted deliveries ordered in the subject's name (Matisse 2015). Even these seemingly

[6] The Violentacrez example may also be interpreted as an instance of delegitimization, given his reputation for deliberately and publicly breaking social norms.

[7] Trolling is a complex phenomenon with nuances and its own cultural norms that I cannot explore here. Whitney Phillips (2015) presents a detailed account that places trolling into a broader cultural context.

[8] Revealing someone's telephone number is deanonymizing doxing, since it reveals a connection between a pseudonym (the phone number) and the subject's legal identity.

minor annoyances can serve as a form of intimidation. Mantilla (2015) mentions a case where a subject received an unordered pizza that had been ordered under the name of an accused murderer known to that individual. Another possible form of harassment is 'swatting', where an attacker makes a hoax call to the police claiming that there is a violent disturbance at the subject's address, prompting an armed police response (Mantilla 2015).[9] The identity knowledge gained through targeting doxing can be used to impersonate the subject, and in extreme cases, has been used to make it appear that the subject herself is encouraging others to attack or sexually assault her (Jouvenal 2013; Citron 2014).

'The Nuremberg Files' website is a notorious example of targeting doxing. This web site began in 1997 and listed the names and personal details of doctors who performed abortions in the US. The site also listed the personal details of the doctors' families (Solove 2007). The Nuremberg Files example illustrates the importance of the context within which identity knowledge is presented.[10] At least some of the information presented there (particularly, the addresses of abortion clinics) would already be publicly accessible. What makes it targeting doxing (beyond the additional identity knowledge about the doctors and their families) is presenting this information in a manner that promotes harassing the subjects.

Delegitimizing doxing

Delegitimizing doxing releases private information with the intention of undermining the subject's credibility, reputation, and/or character. It attempts to shame and humiliate the subject, often by portraying her as a transgressor of an established (or supposed)[11] social norm. Whether the subject herself accepts or promotes the social norm is irrelevant. Revealing the subject to be a hypocrite (by publicly supporting a social norm while privately breaking it) is certainly an attempt at delegitimizing her, but delegitimization goes beyond revealing actual or supposed hypocrisy. It can serve as a tool for maintaining the 'tyranny of the majority' that concerned John Stuart Mill. By drawing attention to how the subject differs from "prevailing opinion and feeling", delegitimizing doxing serves to "fetter the development, and […] prevent the formation of, any individuality not in harmony with its ways" (Mill 1989[1859]: 8).

Reporting information and seeking advice, information, or assistance (rationale 4 in Marx's list) presents the possibility for delegitimization depending on the information sought or reported, and the help requested. The traditional confidentiality of medical records and library borrowings is also motivated by a desire to keep potentially embarrassing or easily misunderstood information secret, so that people can seek medical help or read controversial books without fear of being ostracized for doing so (Rindfleisch 1997; Bowers 2006). A straightforward example is a teenage girl anonymously seeking a pregnancy test or an abortion. If her identity was revealed, she risks being stigmatized and shamed for being sexually active at a young age, especially if unmarried motherhood and/or abortion are unacceptable in her society.

Sexuality is frequently used to delegitimize others. The violent and misogynist language surrounding many instances of delegitimizing doxing implies the objectification of the subject, portraying her as a thing to be used and discarded rather than an autonomous person worthy of respect (Nussbaum 2010). An example is involuntary pornography or so-called 'revenge porn', where intimate or explicit photographs and videos of individuals are posted online without their consent, either by former lovers or by third parties who have somehow acquired them (Mantilla 2015). These images are sometimes accompanied by personal information identifying the person (Citron 2014). As Citron writes, "Harassers post women's nude images because they know it will make them unemployable, undateable, and at risk for sexual assault" (2014:17). While men are also victims of involuntary pornography, the overwhelming majority of victims are women. For example, Reynolds (2016) reports that images of women were involved in 80 % of the 139 cases of involuntary pornography reported in the UK between January and April 2015. Involuntary pornography and other forms of delegitimizing doxing of women based on their sexuality are only the latest instances of the long-lived and surprisingly resilient activity of 'slut-shaming', where women and girls are ridiculed and harassed for their real or imagined sexual activity. Such harassment has a history going back to at least Roman times (Webb 2015). It also reveals a double standard in the social norms associated with sexuality, as male heterosexual activity does not share the same social disapproval (Poole 2013; Citron 2014).

Part of the harm delegitimizing doxing causes is what Franks (2012) calls 'virtual captivity': the abuse directed at someone on the Internet is potentially available to everyone who interacts with her, and so might affect every social

[9] The term 'swatting' is derived from the name of police SWAT (Special Weapons And Training) squads who respond to potentially violent situations involving armed suspects.

[10] Bowman-Grieve (2009) discusses the Nuremberg Files example in more detail and places it into the broader context of violent anti-abortion activism in the US.

[11] A supposed social norm is one held by a minority in society that they believed should hold sway over the majority. The lack of widespread recognition means it is not an established norm (even if it had been historically), but for the group who hold it, they believe that everyone in society should accept it and judge others accordingly.

relation she has. The possibility that everyone the subject interacts with (personally or professionally) has been exposed to the delegitimizing material is enough to cause significant emotional distress and social withdrawal. Martha Nussbaum describes something similar with her concept of subjectivity-violation, where for an abuser, "pleasure is taken in invading and colonizing the person's inner world" (2010: 72).

The context or framing within which delegitimizing doxing occurs is significant, and much of the harm it can cause is a result of taking documentary evidence out of context. An incident described by Boyd (2011) presents a good illustration of this problem. A college admissions officer asked boyd about an apparent contradiction in a prospective student's application: the student claimed to want to leave the 'gang-ridden' community he lived in, but the admissions officer found the student's MySpace page included gang insignia. The officer questioned why the student would lie in his application; Boyd's (2011) response is that adopting gang insignia is a necessity for survival in such a community, and that there is no contradiction in adopting the social norms of a community and secretly desiring to be free of their influence.

Delegitimizing doxing is often accompanied by targeting doxing, and so it might be questioned whether there is a significant difference between them. The difference between is that delegitimizing doxing supplies 'evidence' for targeting the person involved. If targeting doxing supplies the means for harassing the subject, delegitimizing doxing supplies the supposed 'motive' for doing so.

The combination of targeting and delegitimizing doxing is demonstrated by the 'GamerGate' incident, where several high-profile female computer game developers were subjected to prolonged harassment, intimidation, and vilification. The catalyst of this incident was an account posted on the Internet by Eron Gjoni, a former boyfriend of the independent game developer Zoe Quinn, of their failed relationship (Mantilla 2015). Quinn's personal details were released on the Internet and she became the target of prolonged and sustained harassment, intimidation, and vilification (Mantilla 2015). While attempts were made to justify these attacks as attempts to expose wrongdoing in computer games journalism, as one of the men Quinn was alleged to have had a relationship with was a video games journalist (who had not even written about Quinn's game), misogynism is a more convincing explanation (Mantilla 2015). Following the attacks on Quinn, other prominent women associated with computer games, including developer Brianna Wu and critic Anita Sarkessian, were also targets of sustained harassment and intimidation that included targeting and delegitimizing doxing (Mantilla 2015).

Can doxing be justified?

I now discuss whether any instances of doxing are justifiable. I will argue that deanonymizing doxing may be acceptable depending on the rationale for anonymity and if there is a compelling public interest justification for revealing someone's identity. I also claim that delegitimizing doxing may be permissible if it exposes evidence of actual wrongdoing of public interest, and that the information revealed must only be sufficient to establish that such wrongdoing has occurred. I will argue that targeting doxing is unjustifiable, as it deliberately increases the risk of physical harm to the subject. In all cases, however, the burden of proof is on whoever wishes to disclose identity knowledge about the subject to justify why her anonymity or obscurity should be removed.

The motivation behind doxing is significant for deciding whether it is defensible or not. Doxing as a form of intimidation is unacceptable as it attempts to silence the subject and prevent her from participating in social, political, and public activity. All three types of doxing may be used for intimidation. Deanonymization intimidates those who adopt a pseudonym or seek anonymity to express unpopular or controversial views that they are otherwise uncomfortable in expressing. Targeting doxing increases the ease with which someone may be physically harassed or harmed. Delegitimization vilifies the subject, inspiring further harassment and reducing the likelihood that her opinions will be given the public respect that they might otherwise receive.

My arguments place considerable weight on the concept of 'public interest'. My interpretation of the concept is based on two claims by Bok (1989): "[t]he public has a legitimate interest […] in all information about matters that might affect its welfare" (1989:258) and that information reported to the public that only satisfies their curiosity rather than affects their welfare must take into account the privacy of those affected. These two claims reflect the public/private distinction common to liberal political philosophy. Deanonymizing and delegitimizing doxing are acceptable only if they concern matters that affect the welfare of the public. If we accept that individuals should have control over who has access to identity knowledge about herself, the burden of proof should be on whoever attempts to reveal such information to justify why revealing it is in the public interest. If they are to be justified, deanonymizing and delegitimizing doxing cannot be indiscriminate: it must reveal *only* the identity knowledge that is relevant to establish wrongdoing by a specific individual. Such doxing can be considered analogous to whistle blowing that reveals wrongdoing by or within organizations.

Particular instances of doxing could be justified if there are allegations of legal wrongdoing, or if there is a legitimate public interest reason for establishing someone's identity. This would seem to rule out most (if not all) cases of targeting doxing, as these cases are intended to intimidate and promote further harassment of the subject. The legitimacy of doxing depends on the motivations behind publically releasing some of the subject's identity knowledge and the foreseeable risks of harm to the subject from doing so. An instance of deanonymizing doxing, therefore, might be justified on a consequentialist basis if the benefits to the public of exposing wrongdoing or deception outweigh the foreseeable harms to the subject due to her loss of obscurity.

An author using a pen name, for instance, does not seem to be a compelling target for deanonymization on public interest grounds if the pen name is not used to deceive readers and is merely a way for a writer to adopt different personas for different styles and genres of writing. Revealing that Charles Dodgson was Lewis Carroll does not seem to be particularly compelling from a public interest standpoint as there is nothing inherent deceptive in the claims made by the two personas. The rationales of play and promoting experimentation are frequently the motivations for adopting these personas (rationales 12 and 13 from Marx's list), rather than any attempt to deceive. Pen names for authors and stage names for performing artists are often little more than designations of public personas, and offer little in the way of anonymity or obscurity. Revealing that the musician Bono was born Paul Hewson does little (if anything) to affect the meaning of his music or his political and social activism. While deanonymization does not seem particularly troubling in these cases, there also seems to be little reason for doing so other than satisfying curiosity.

So-called 'cross-penning', where an author adopts a pen name of a different gender, is slightly more problematic as it may be intended to mislead the reader. However, here the pseudonym is often adopted to lend the work credibility and allow it to be judged on its own merits rather than unfairly influenced by gender bias. This is the rationale of wishing to keep attention to the work itself rather than to its creator (rationale 3 from Marx's list). George Sand, the pen name of Armandine Dudevant, is just one example (Levmore 1996). Given the disproportionate chances of women receiving misogynistic hate speech for their writings on the Internet, female authors sometimes adopt masculine names as a means of avoiding becoming targets for online abuse (including doxing) (Citron 2014). Cross-penning in these cases is both an attempt to have their work judged fairly and as a means of protection against harassment. However, this does not extend to literary hoaxes where the author falsely claims to have personal experiences or attributes that lend unjustified credibility to her work. Autobiographies that feature elaborate false accounts of the author's circumstances and experiences are one example (Manning 2012). It is more plausible to argue in such cases that the author is being deceptive in these cases, as they are not motivated by a desire to focus attention on the work itself but on how the falsely claimed characteristics of the author lend credibility to the work.

Ghostwriting is an interesting case where the problem is reversed: the actual author is not the attributed author. Ghostwriters are often an 'open secret': it is assumed that many public figures use ghostwriters to produce works published in their name (Goldacre 2012). At best, employing a ghostwriter allows for the attributed 'author' to better express her own ideas. At worst, the 'author' is misrepresenting her abilities to the readers of 'her' work. If ghostwriting is used to obscure the source or interests of the actual author, then there is a public interest justification for revealing this. For example, such justifications exist in the case of medical research, where the information published may be used to decide on medical treatment and to direct future research (Ngai et al. 2005). The names of seemingly independent researchers are sometimes attached (with their permission) to pharmaceutical studies to obscure the fact that were primarily designed and conducted by pharmaceutical company researchers (Goldacre 2012). Revealing the use of ghostwriters to gain unwarranted credibility for published works would be delegitimizing doxing with a public interest justification.

I now return to the two specific instances of deanonymizing doxing described earlier: Satoshi Nakamoto and Violentacrez. Is there a public interest justification for deanonymizing Satoshi Nakamoto? It depends on whether identifying Nakamoto is merely satisfying public curiosity or establishing information that benefits the public. There is certainly historical interest in establishing the identity of the creator (or creators) of such an influential technology. The anonymity of Bitcoin's creator may also raise suspicions about the intent behind creating it. However, given that both the theory behind Bitcoin and the source code of the software implementing it are open to public review and revision, it seems unlikely that there is anything malicious within the design and implementation of Bitcoin itself. The pseudonym is unlikely to have been adopted to deceive others for Nakamoto's benefit, and the adopting the pseudonym offers Nakamoto protection against unwanted interference and outside interest. Discovering Satoshi Nakamoto's identity would certainly be interesting given Bitcoin's influence and technical merit, but there seems little public benefit (in the sense that it would better inform the public in matters that affect it) in revealing this information beyond satisfying this curiosity. There does not appear to be a strong reason for removing

the actual creator (or creators) of Bitcoin from self-imposed obscurity.

The case of Satoshi Nakamoto raises another important point: does doxing in any form have to be accurate to be harmful? I suggest that the credibility that releasing documentary evidence has is important for what makes doxing particularly harmful, as it cannot be simply dismissed as gossip or hearsay. Through no fault of his own, Dorian Satoshi Nakamoto lost his obscurity after the allegations that he was 'Satoshi Nakamoto' were published. Inaccurate or out-of-date personal information released as targeting doxing could lead to unconnected individuals being harassed. Inaccurate or false doxing may not be as harmful to the subject as accurate doxing, but it is still an attempt to remove the subject's anonymity or obscurity, and may harm others who are wrongly identified as the subject individual.

Is there a public interest in revealing the identity of Violentacrez? Again, it is not straightforward that disclosing Violentacrez's identity is in the public's interest or just something to satisfy the public's curiosity. Unlike Dorian Satoshi Nakamoto, Michael Brutsch acknowledged that he was 'Violentacrez', and so his own actions under that pseudonym led to his loss of obscurity. Violentacrez's actions were certainly (and deliberately) offensive to many people, and the deanonymizing doxing forced him to stop. Distinguishing between offensive speech and hate speech offers a potential justification for this deanonymization. Hate speech expresses claims that those with certain characteristics (such as gender, race, or sexual preference) are inferior in moral worth and little more than objects to be used and exploited (Citron 2014). There is a public interest in resisting such expression as it promotes harmful divisions within society. Hate speech damages the perception (and if left unaddressed, eventually the treatment) of such people as moral agents equal to ourselves that we have duties toward and with rights of their own (Waldron 2010). In the case of Violentacrez, many of the sub-reddits that he created or moderated (such as 'chokeabitch' and 'rapebait') may be classified as 'hate speech' or objectification. Deanonymization might be justified as a means of limiting or stopping such hate speech by increasing the speaker's accountability.

A strong objection to this conclusion is that such deanonymizing doxing risks of turning into the private enforcement of public laws and moral standards, and has the potential to further develop into vigilantism. Trottier suggests the term 'digital vigilantism' for "a process where citizens are collectively offended by other citizen activity, and respond through coordinated retaliation on digital media" (2016:2). The identity knowledge revealed through deanonymization makes it considerably easier to perform targeting and delegitimizing doxing of the subject. While it may appear to be 'just desserts' for a hate speaker to be harassed, it should be rejected on the same grounds that the intimidation promoted by targeting doxing is rejected.

Another objection is the concern that the costs and harms of deanonymization to the individual concerned outweigh the social benefits of making her accountability for offensive behavior. Consider a situation where someone uses a pseudonym to express controversial views that could be portrayed as harmful to the public interest, such as seditious comments or questioning strongly held religious or social beliefs in ways that are not hate speech. Those who object to such views might justify deanonymizing this person on public interest grounds. This would return deanonymizing doxing to being a tool for intimidating those with unpopular views rather than as a means of making those who cause harm through their anonymous or pseudonymous actions accountable.

The Violentacrez case is a good illustration of this second objection. The deanonymization was used to shame Michael Brutsch and harm him materially, as he lost his job through being deanonymized.[12] Had he not been protected by his pseudonym, it is likely he would have acted differently. If we accept that at least some of Violentacrez's postings were hate speech and objectification, that there is a public interest in controlling expressions of such speech. However, given the material harms that he would foreseeably suffer as a result of anonymization, it is worth considering what alternatives were available for stopping the hate speech from taking place. In this instance, there is a clear alternative: the site operators should have removed his deliberately offensive postings and sub-forums from Reddit. This alternative may be challenged by claims that it limits freedom of expression, but it may be defended if freedom of expression is not considered to be an absolute right that cannot be limited by other rights (Waldron 2010).

Another response is to emphasize that accountability should go both ways in deanonymizing someone: whoever performs the deanonymization should not be anonymous or pseudonymous herself. In both the 'Nakamoto' and Violentacrez examples is that the persons revealing identity knowledge about the subjects were not themselves anonymous. The journalists involved did not use anonymity as a means of avoiding responsibility for their actions. Journalists also have editors and their own professional judgment about what is in the public interest to reveal in news stories. Revealing information under a legal name makes it easier to be held accountable for doing so.

[12] It should be noted that there is a difference between this shaming and that covered by delegitimizing doxing. Unlike delegitimizing doxing, the material used to delegitimize Brutsch was already revealed: his offensive contributions to various sections of Reddit.

This accountability does not in itself legitimize doxing, as there are other protections that those revealing information may enjoy that are unavailable to the subject. Private individuals are unlikely to be able to afford costly legal disputes with media enterprises that reveal their personal information. Nonetheless, this offers a potential solution for difficult cases where it is not clear whether there is a public interest in revealing someone's identity or not.

Finally, I will briefly consider the possibility of targeting and delegitimizing doxing that is claimed to be the public interest. Exposing corruption is often used to justify doxing or Internet campaigns that feature doxing. An example is the so-called 'Gamergate' controversy, which its defenders claim is an attempt to expose corruption in computer games journalism.[13] Quinn's relationship with the journalist Nathan Grayson had no effect on how her game was reviewed, as Grayson's published work only mentioned her game in passing and did not actually review it (Mantilla 2015). The 'exposing corruption' justification is further weakened by the forms of doxing used on Quinn (particularly targeting doxing) and Gjoni's apparent motivation behind releasing delegitimizing information, which appears to be a desire to punish his former girlfriend (Mantilla 2015). The public interest should not justify any instance of doxing that objectifies the subject, as objectification portrays the target as unworthy of personhood, without legitimate interests of her own that should also be recognized (Nussbaum 2010).

Conclusion

Anonymity and obscurity protect us from the unwanted intrusion of others into our lives, and allow us to express our ideas and ourselves in circumstances where we otherwise could not. They are also useful tools for deception and hiding wrongdoing. Doxing hinders all of these purposes by deliberately removing some of the subject's obscurity. In the three forms I have described here, doxing can be a tool for establishing accountability for wrongdoing, a means of intimidation and incitement to cause harm, and a way of silencing minority or dissenting views. I have argued that only revealing personal information that is in the public interest to disclose and only to the extent necessary to establish wrongdoing is justified. There is much more to say about doxing and the role it plays in public discourse, and I hope this brief account of how acts of doxing can be classified will be useful to those who examine these issues in the future.

[13] By (willfully?) misinterpreting academic studies into video games, some have even made the outlandish claim that it has revealed a 'Communist' and 'feminist' conspiracy to undermine video games (Chess and Shaw 2015).

Acknowledgments I wish to thank the participants of the 2015 Amsterdam Privacy Conference and my colleagues in the philosophy department of the University of Twente for their encouragement and suggestions on earlier drafts on this paper. I also like to thank Kevin Macnish for a stimulating discussion of privacy, Nolen Gertz and the students in the University of Twente's Philolab class for their comments and questions, and the anonymous reviewers for their helpful comments, suggestions, and encouragement.

References

Beaujon, A. (2014). Redditors furious Newsweek 'doxxed' Bitcoin Founder. *Poynter*. http://www.poynter.org/news/mediawire/242348/redditors-furious-newsweek-outed-bitcoin-founder/.

Bok, S. (1989). *Secrets: On the ethics of concealment and revelation*. New York: Vintage Books.

Bowers, S. L. (2006). Privacy and library records. *The Journal of Academic Librarianship, 32*(4), 377–383.

Bowman-Grieve, L. (2009). Anti-abortion extremism online. *First Monday* 14(11). http://firstmonday.org/ojs/index.php/fm/article/view/2679.

Boyd, D. (2011). 'Real Names' policies are an abuse of power. *Apophenia*. http://www.zephoria.org/thoughts/archives/2011/08/04/real-names.html.

Chen, A. (2012a). Unmasking Reddit's Violentacrez, the biggest troll on the web. *Gawker*. http://gawker.com/5950981/unmasking-reddits-violentacrez-the-biggest-troll-on-the-web.

Chen, A. (2012b). Reddit's biggest troll fired from his real-world job; Reddit continues to censor Gawker articles. *Gawker*. http://gawker.com/5951987/reddits-biggest-troll-fired-from-his-real-world-job-reddit-continues-to-censor-gawker-articles.

Chen, A. (2012c). Reddit troll Michael Brutsch defends himself on CNN: 'I treated Reddit as a game.' *Gawker*. http://gawker.com/5953097/reddit-troll-michael-brutsch-defends-himself-on-cnn-i-treated-reddit-as-a-game.

Chess, S., & Shaw, A. (2015). A conspiracy of fishes, or, how we learned to stop worrying about #GamerGate and embrace hegemonic masculinity. *Journal of Broadcasting & Electronic Media, 59*(1), 208–220.

Citron, D. K. (2014). *Hate crimes in cyberspace*. Cambridge, MA: Harvard University Press.

Franks, M. A. (2012). sexual harassment 2.0. *Maryland Law Review, 71*, 655.

Gao, L., & Stanyer, J. (2014). Hunting corrupt officials online: The human flesh search engine and the search for justice in China. *Information, Communication & Society, 17*(7), 814–829.

Gavison, R. (1980). Privacy and the limits of law. *The Yale Law Journal, 89*(3), 421–471.

Goldacre, Ben. (2012). *Bad pharma: How drug companies mislead doctors and harm patients*. London: Fourth Estate.

Goodman, L. M. (2014). The face behind Bitcoin. *Newsweek*. http://www.newsweek.com/2014/03/14/face-behind-bitcoin-247957.html.

Honan, M. (2014). What is doxing? *Wired*. http://www.wired.com/2014/03/doxing/.

Jouvenal, J. (2013). Stalkers use online ads as weapon against victims. *The Washington Post*. http://www.washingtonpost.com/local/i-live-in-fear-of-anyone-coming-to-my-door/2013/07/14/26c11442-e359-11e2-aef3-339619eab080_story.html.

Levmore, S. (1996). The anonymity tool. *University of Pennsylvania Law Review, 144*(5), 2191–2236.

Manning, M. G. (2012). A tale of three hoaxes: When literature offends the law. *Columbia Journal of Law & the Arts, 36*, 127–156.

Mantilla, Karla. (2015). *Gendertrolling: How misogyny went viral*. Santa Barbara, CA: Praeger.

Marx, G. T. (1999). What's in a name? Some reflections on the sociology of anonymity. *The Information Society, 15*(2), 99–112.

Mattise, N. (2015). Anti-doxing strategy—Or, How to Avoid 50 Qurans and $287 of Chick-Fil-A. *Ars Technica*. http://arstechnica.com/security/2015/03/anti-doxing-strategy-or-how-to-avoid-50-qurans-and-287-of-chick-fil-a/.

Mill, J. S. (1989). *On liberty and other writings*. Edited by Stefan Collini. Cambridge: Cambridge University Press [1859].

Nakamoto, S. (n.d.). Bitcoin: A peer-to-peer electronic cash system. *Bitcoin*. https://bitcoin.org/bitcoin.pdf.

Ngai, S., Gold, J. L., Gill, S. S., & Rochon, P. A. (2005). Haunted manuscripts: Ghost authorship in the medical literature. *Accountability in Research, 12*(2), 103–114.

Nussbaum, M. C. (2010). Objectification and internet misogyny. In S. Levmore & M. C. Nussbaum (Eds.), *The offensive internet: Speech, privacy, and reputation* (pp. 68–87). Cambridge, MA: Harvard University Press.

Ockenden, W., & Leslie, T. (2015). 'Scarily accurate': What you found in our reporter's metadata. *ABC News*. http://www.abc.net.au/news/2015-08-24/metadata-what-you-found-will-ockenden/6703626.

Phillips, W. (2012). What an academic who wrote her dissertation on trolls thinks of Violentacrez. *The Atlantic*. http://www.theatlantic.com/technology/archive/2012/10/what-an-academic-who-wrote-her-dissertation-on-trolls-thinks-of-violentacrez/263631/.

Phillips, Whitney. (2015). *This is why we can't have nice things: Mapping the relationship between online trolling and mainstream culture*. Cambridge, MA: The MIT Press.

Poole, E. (2013). Hey girls, did you know: Slut-shaming on the internet needs to stop. *University of San Francisco Law Review, 48*, 221–260.

Reichel, P. L. (1977). Dossier building as a social problem topic. *Teaching Sociology, 4*(3), 293–306.

Reynolds, E. (2016) Revenge porn: There's no 'silver bullet' for ending non-consensual pornography. *Wired UK*. http://www.wired.co.uk/news/archive/2016-03/08/revenge-porn-facebook-social-media.

Rindfleisch, T. C. (1997). Privacy, information technology, and health care. *Communications of the ACM, 40*(8), 92–100.

Solove, D. J. (2007). *The future of reputation: Gossip, rumor, and privacy on the internet*. London: Yale University Press.

Trottier, D. (2016). Digital vigilantism as weaponisation of visibility. *Philosophy & Technology,* 1–18. doi:10.1007/s13347-016-0216-4.

Waldron, J. (2010). Dignity and Defamation: The Visibility of Hate. *Harvard Law Review, 123*(7), 1596–1657.

Webb, L. M. (2015). Shame transfigured: Slut-shaming from Rome to cyberspace. *First Monday, 20*(4). http://firstmonday.org/ojs/index.php/fm/article/view/5464.