

## **[6352] Libel Inc: An Analysis of the Libel Site Ecosystem**

PRELIMINARY DECISION: accept

### Summary of Reviews

- Review 1: 2 (4)
- Review 2: 1 (4)
- Review 3: 2 (3)
- Review 4: -1 (3)
- Review 5: 0 (5)

### Reviews

#### **Review 1**

**TOTAL SCORE: 2**

**Overall evaluation:** 2 (accept)

**Reviewer's confidence:** 4 (high)

This is a good paper providing a clear and systematic understanding of the libel site ecosystem. Some room for improvement but acceptable as is.

#### Ethics

-----

Good consideration of the ethics here. This study should have had an external ethics review and you should ask your IRB to fix their policies so that important and sensitive work like this can get appropriate and proportionate review. Impact on the authors of reading through all the nasty posts studied is not mentioned and would be a key consideration in this research. "Quotations of graphic depictions involving child predators are omitted here" I am glad they are omitted here and pity the poor researchers who had to read them. I hope you have been appropriately supported and have access to counselling etc. as necessary.

You should consider what data you would make available to other researchers on request to enable future work or cross comparison. For example, the unredacted list of sites used or even the raw post data. Controlled sharing rather than unrestricted sharing as described in:

@inproceedings{Thomas2017a,

author = {Thomas, Daniel R. and Pastrana, Sergio and Hutchings, Alice and Clayton, Richard and Beresford, Alastair R.},

booktitle = {Internet Measurement Conference},

date = {2017-11},

doi = {10.1145/3131365.3131389},

isbn = {978-1-4503-5118-8},

keywords = {Cybercrime, Data of illicit origin, Ethics, Found data, Law, Leaked data, Menlo report, Unintentionally public data},

location = {London, UK},

publisher = {ACM},

rights = {All rights reserved},

```
series = {IMC},
title = {Ethical issues in research using datasets of illicit origin},
url = {https://www.repository.cam.ac.uk/handle/1810/267728}
}
```

(you don't need to cite this, just including it here for context).

Other comments

-----

4.2.2: Creating a new phone number is not difficult or expensive and would have enabled the forms to be used. Using a different name and email for each site might have been better in case the different sites are run by the same entity and so can spot the repeated requests. Would be interesting to see if the fee varied by libel site or by other factors.

4.2.3: Would be helpful to try and quantify whether any of these non-status code removal options actually happened.

Have you attempted to contact the search engine companies with the details of the sites to see what their response is?

6.3: A little more evidence as to the plausibility of this would help.

Good set of Recommendations. Probably don't need "and suggestions"

Check the references, some capitalisation etc. issues e.g. [4].

## **Review 2**

**TOTAL SCORE: 1**

**Overall evaluation: 1** (weak accept)

**Reviewer's confidence: 4** (high)

This is a very interesting paper on a new topic that is highly interesting for several stakeholders. I enjoyed reading this paper. However, the structure is a bit confusing. The paper is not following the general structure of most academic papers. I think that this should be approved. I advise the authors to look at some example papers with similar data/methods. There should be a clear structure of: introduction, literature review, methods, results, conclusion/discussion. I specifically noticed:

- The paper already introduces the results in the introduction. That is not common and there is even some repetition of the results in the 2nd and 3rd paragraphs. As results are also mentioned in the abstract, I would recommend to delete these from the introduction. I would advise to provide a bit more details on the methods and goals of the paper instead.
- The main research question was not clearly stated, which also makes it a bit difficult to see how the authors are trying to answer this question. There should be a clear link between the research question, methods and results. For example, the authors provide a taxonomy of libel posts, but the introduction did

not mention that this was the goal of the paper. There is also limited information on how this taxonomy was developed and these results were not really used in the rest of the paper or the conclusion/discussion.

- It was not clear to me if section 2 is based on literature or if these are already results from this study (what is the source?). Similarly it is unclear what the goal of section 2.1 is. Shouldn't this be part of the methods section? Right now, previous literature is only mentioned in section 7 at the end of the paper. This should be discussed in the beginning and these previous results should also be used in the story line of the paper. Currently the authors only mention some studies, but they do not really show how these relate to their paper (in the literature review) and their results (in the discussion).
- Information about how the data was analyzed should be in the methods section, but is now in the results section.
- Additional information on the limitations of the data and analyses and how this may have an impact on the interpretation of the results should be added in the methods or discussion section.

Additional comments:

I noticed some typos.

Method

Different types of websites are discussed (also gossip and blogs), but it was not clear to me which of these were included in the data and why.

Results

Some of the post types in the taxonomy still seem to include more than one type. For example number 3 contains both sex worker posts and LGBTQ+ posts and number 5 includes scammer and racist posts. These seem to be different categories. Information on how and why these were categorized in the same type is missing.

### **Review 3**

**TOTAL SCORE: 2**

**Overall evaluation: 2 (accept)**

**Reviewer's confidence: 3 (medium)**

The article details investigations into an interesting and under-studied online extortion ecosystem. "Libel sites" publish anonymous unverified reports with allegations of child predation, racism, fraud or, most commonly, promiscuity and infidelity. The authors explore how these sites appear to be generating revenue via 'reputation management' services advertised on the basis of removing posts about a subject from these sites. The authors' investigations seem to show that removals are infrequently attempted, and incomplete even when they are carried out. The authors contacted several services and established that prices for post removal range from 2k - 6k USD.

I enjoyed much of this paper, and appreciate the multifaceted investigation of the ecosystem, including the impact of indexing and aggregators in amplifying accusations.

I have some comments regarding the recommendations the authors make:

The recommendations for internet search engines to allow reports for exploitative removal practices are perhaps an important step, but have some potential limitations. First, as I understand it the libel sites themselves are distanced from the extortion, as they offer no explicit removal procedure at all, there 'just so happens' to be an advert for a service which claims to be able to help subjects. While we can see this as being no accident in this case, many sites might justifiably balk at being considered for delisting based on the practices of services that end up advertised on their pages.

Secondly, I am concerned about misuse of these reporting schemes by malicious actors. Many anti-fraud organisations collect reports to warn internet users of scams and/or specific scammers, and handing fraudsters tools to de-list warnings about their activities could be quite a negative development. Of course, a legitimate anti-fraud organisation would not engage in exploitative removals, but the accusation could easily be levelled based on the similarity of these sites, and the investigations or appeals could be costly to search engines or fraud organisations.

Neither of the above means this approach is necessarily unworkable, but the policies involved here might have to be carefully trialled.

I would suggest separating the advice for libel post writers from that for the intended audience of libel sites, as they are quite different parties and the advice is quite different. For libel post writers, you may also wish to suggest caution on grounds of truthfulness (do they really know what they think is true?) and legal liability (depending on content and jurisdictions libel can lead to a civil suit). The differing motivations of libel post authors might also be explored and channelled more productively (e.g., if the intent is to vent emotion about a cheating spouse, this could be handled without publishing identifiable details). For the audience, it may be worth discussing mechanisms for automated online background checks to exclude any kind of 'rumour' sites that make unverified claims.

Minor Notes:

Abstract: "that often including"

**Review 4****TOTAL SCORE:** -1**Overall evaluation:** -1 (weak reject)**Reviewer's confidence:** 3 (medium)

The authors present an extended analysis of the ecosystem of libel sites, covering libel sites, reputation management sites, aggregator websites, and relevant online social media sites. The sites are crawled using a variety of methods, depending on the structure and provided functionalities of the site. The crawling is repeated after a 2 week and 4 week interval to evaluate the removal of libel posts.

Although a large dataset of libel sites was collected, the paper mainly reports on a qualitative analysis of the different libel sites. Although the reported findings are interesting, the paper could have benefited from more in-depth analyses on the collected data, currently leaving many questions unanswered, e.g.: to what extent do libel sites share the same content? how frequently are new posts created? are there any infrastructural similarities (hosting company, domain name registrar, ...) between libel sites and reputation management sites that would indicate a shared ownership?

The website collection method leveraging the "snowballing method" seems prone to limit the findings to certain bubbles, i.e., associated sites. It is unclear how many sites were found from the search engine and how many were discovered through this snowballing effect. Furthermore, it should be mentioned that by leveraging a search engine for site discovery, this may affect the results regarding their appearance in search engines (section 4.3).

The proposed recommendation and suggestions seem rather handwavy; it would be interesting to discuss more in depth what the feasibility and consequences of the various suggestions would be.

Using just the status code to verify the disappearance of a post seems inadequate; analyzing the content of the response would be more accurate.

**Review 5****TOTAL SCORE:** 0**Overall evaluation:** 0 (borderline paper)**Reviewer's confidence:** 5 (expert)

Comment from the PC and the Chairs: the PC and the Chairs agree that, given the multidisciplinary nature of the event and the discipline-specific nature of writing conventions, the structure of the paper can remain as-is.

Besides this note, we ask the authors to carefully review and consider all reviewer comments for the camera ready version of the paper.