One Click Away: A Study on the Prevalence of Non-Suicidal Self Injury, Suicide, and Eating Disorder Content Accessible by Search Engines

PRESENTED BY

NCRI
NETWORK CONTAGION RESEARCH INSTITUTE

COMMISSIONED BY

Ofcom
Making communications work for everyone

AUTHORS

Lee Jussim, PhD, Rutgers University (Principal Investigator)

Danit Finkelstein, Rutgers University

Sonia Yanovsky, Rutgers University

Alex Goldenberg, NCRI (Corresponding Author)

Joel Finkelstein, PhD, NCRI

Edward Selby, PhD, Rutgers University

Richard Benson, NCRI

Sagar Patel, NCRI

Tayler Glover, NCRI

Collin Vasko, NCRI

Ankita Jagdeep, NCRI

Lamia Abdallah, Rutgers University

Sean Harley, Rutgers University

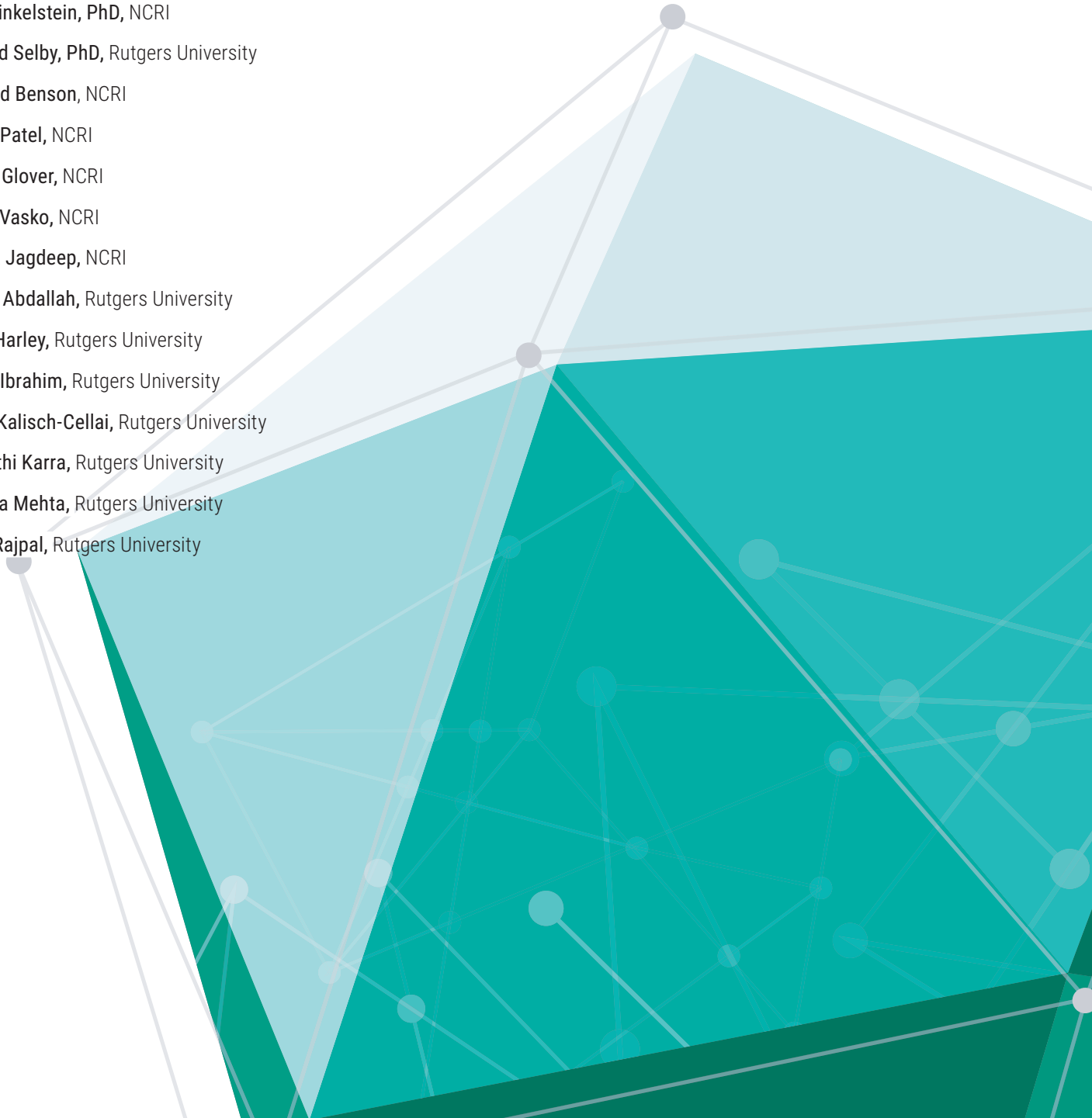
Claire Ibrahim, Rutgers University

Carlo Kalisch-Cellai, Rutgers University

Pranathi Karra, Rutgers University

Ananya Mehta, Rutgers University

Ryan Rajpal, Rutgers University



[CONTENT WARNING]

This report contains discussion and depiction of distressing content related to self-injury, suicide, and eating disorders. The purpose of this content warning is to inform readers about the sensitive nature of the material discussed in the report. We understand that these topics and images can be triggering and emotionally distressing for some individuals. If you feel that you may be affected by such content, we encourage you to consider your well-being and decide whether reading further is appropriate for you at this time.

Please note that the information and images presented in this report are provided for educational and informational purposes only. It is not intended to replace professional advice or treatment. If you or someone you know is in need of immediate assistance, please reach out to the appropriate resources listed on the next page.

IF YOU OR SOMEONE YOU KNOW IS STRUGGLING WITH SELF-HARM, SUICIDE, OR EATING DISORDERS, PLEASE CONSIDER SEEKING HELP FROM THE FOLLOWING RESOURCES:

Samaritans (United Kingdom and Ireland)

Phone: 116 123 (UK and Ireland)
Website: https://www.samaritans.org

PAPYRUS Hopeline UK (Prevention of Young Suicide)

Phone: 0800 068 4141
Website: https://papyrus-uk.org

Mind Infoline

Phone: 0300 123 3393
Website: https://www.mind.org.uk

Beat (Eating Disorders)

Phone: 0808 801 0677
Website: https://www.beateatingdisorders.org.uk

International Association for Suicide Prevention (IASP)

Website: https://www.iasp.info/resources/Crisis_Centres

CALM (Campaign Against Living Miserably)

Phone: 0800 58 58 58
Website: https://www.thecalmzone.net

TABLE OF CONTENTS

Remember, you are not alone, and there are people who can provide support and assistance.

EXECUTIVE SUMMARY

Ofcom, the UK's communications regulator, commissioned the Network Contagion Research Institute (NCRI) to investigate the prevalence of self-injurious content in search platform results.

This report presents our research team's findings on trends and patterns in online content that promote three specific types of self-injurious behaviour: non-suicidal self-injury (NSSI; for example, cutting), suicide and eating disorders (ED).

The team's research methodology combined keyword analysis and surfacing, data collection from search engine platforms and link coding when assessing search results, including content that seeks to prevent self-injurious behaviour or provide support services, as well as content that glorifies, instructs and/or celebrates behaviours online associated with NSSI, suicide or ED. The search engines we used were Google, Microsoft Bing, DuckDuckGo, Yahoo!, and AOL, across three types of search functions (general web search, image, and video).

HEADLINE FINDINGS

- In the NCRI's assessment of 37,647 individual search result links across all five search engine platforms, we found that 22% of links contained content that celebrates, glorifies, or instructs self-injurious behaviour. All this content was available through a single click ('one click away') from the main search engine page results.
- This material is easily accessible: among the first five search results (i.e., those appearing on the very first page), 22% of analysed content celebrated, glorified, or instructed self-injurious behaviour.
- Among the categories of self-injurious behaviour studied, celebration/glorification occurred in 2,993 (25%) of 11,970 results for content related to non-suicidal self-injury (e.g., making cuts to body parts), in 1,580 (12%) of 13,054 results for content related to suicide, and in 3,695 (29%) of 12,623 results for content related to eating disorders (anorexia, bulimia and binge eating).
- In our sample, image searches accounted for 50% of all celebration/glorification content, compared to 28% from web searches and 22% from video searches.
- Three social media platforms accounted for 59% of content in this sample which celebrates or glorifies self-injurious behaviour.
- Users and communities seeking to discuss or even promote self-injurious behaviour do not always use explicit terms. Instead, they use abbreviations and cryptic words to obscure the topic and escape detection by search filters Our results suggest that content containing obscure and cryptic keywords constituted the vast majority of results found to promote self-injurious behaviour.

'one click away'

Throughout this report, we refer to all findings as 'one click away' because all could be reached by a single click of a search result. When our research team coded each individual link, they assessed the link as they found it – even if they were prompted to sign in, go through a paywall, or click through to any other part of the website, they always remained on the same page on which they started. So, each of the links that contained content coded as promoting self-injurious behaviour was also 'one click away'.

This means every link analysed in this report was easily available to any user conducting that search, highlighting the role search engine providers play in enabling access to potentially harmful content. This includes children and individuals who come across such content unintentionally while searching for topics like recovery or educational information.

INTRODUCTION TO THE PROJECT AND SCOPE

Research Questions

1. How easily accessible is content that potentially encourages self-injurious behaviour? How much can be accessed by a single click of a search result? How much appears early in search results?
2. What were the patterns, if any, in the content that potentially encourages self-injurious behaviour?
3. How, if at all, do answers to the above differ between search services (Google, Bing, DuckDuckGo, Yahoo, and AOL)?
4. Where are disproportionate amounts of content promoting self-injurious behaviour found? Are the answers to the above questions different for different search products (main results feed, image search, and video search)? Do certain platforms or sites host disproportionate amounts of content promoting self-injurious behaviour?

Non-Suicidal Self-Injury, Suicide, and Eating Disorders Online

NSSI, suicide, and eating disorder websites, discussion boards, and social media communities have created a rich virtual ecosystem[1,2] where users can discuss their experiences and, at times, enable harmful acts. Pro-anorexia (pro-ana), pro-NSSI and pro-suicide content can proliferate in these spaces as users share, and are exposed to, high volumes of instructive and enabling content. Individuals suffering from mental health disorders that manifest as anorexia, NSSI, or suicidal ideation may find a feeling of belonging and acceptance in these online groups.[3] Search engines may produce results that lead to enabling communities. Exposure to such enabling communities can aggravate existing conditions, and deter people from seeking professional care.[4] Furthermore, the anonymity of the internet allows people to post explicit photos and intimate tales without fear of criticism or punishments, perhaps inspiring others who are prone to these behaviours to do the same.[5]

In light of the accessibility of communities that promote harmful behaviours online,[6] there is a growing need to examine the availability of such content through search engines. To do this, NCRI conducted an extensive quantitative analysis of search content promoting self-injurious behaviours on search engine platforms in the public domain. We sought to determine the prevalence of such content generated among over 37,000 search results across 5 search platforms – Google, Bing, DuckDuckGo, Yahoo, and AOL – in response to queries around terms commonly used to refer to non-suicidal self-injury, suicide and eating disorders.

DEFINITION AND DESCRIPTION OF TERMS AND CODING CATEGORIES

Terms

In this report, we use the term self-injurious behaviour to refer to non-suicidal self-injury (NSSI), suicide, and eating disorders (ED). The term is intended to be descriptive and makes no assumptions about intentions, motivations or other psychological disorders. We use self-injurious behaviour as the umbrella term to refer to NSSI, suicide, and ED together.

Non-suicidal self-injury (NSSI) refers to intentionally inflicting pain or damage on one's own body without suicidal intent.[7] NSSI commonly manifests as cutting slices into body parts (arms, legs, etc.), but any self-inflicted injury (burns, bruises) qualify as non-suicidal self-injury.

1 Goldenberg, A., Farmer, J., Jussim, L., Sutton, L., Finkelstein, D., Ramos, C., Paresky, P., & Finkelstein, J. (2022). Online communities of adolescents and young adults celebrating, glorifying, and encouraging self-Harm and suicide are growing rapidly on Twitter. Network Contagion Research Institute. https://networkcontagion.us/reports/8-29-22-online-communities-of-adolescents-and-young-adults-celebrating-glorifying-and-encouraging-self-harm-and-suicide-are-growing-rapidly-on-twitter/

2 Jarvi, S., Jackson, B., Swenson, L., & Crawford, H. (2013). The impact of social contagion on non-suicidal self-injury: A review of the literature. Archives of Suicide Research, 17(1), 1-19.

3 Boero, N., & Pascoe, C. J. (2012). Pro-anorexia communities and online interaction: Bringing the pro-ana body online. Body & Society, 18(2), 27-57.

4 Lewis, S. P., & Seko, Y. (2016). A double-edged sword: A review of benefits and risks of online nonsuicidal self-injury activities. Journal of Clinical Psychology, 72(3), 249-262.

5 Bell, J. (2014). Harmful or helpful? The role of the internet in self-harming and suicidal behaviour in young people. Mental Health Review Journal, 19(1), 61-71.

6 Boyd, D., Ryan, J., & Leavitt, A. (2011). Pro-self-harm and the visibility of youth-generated problematic content. ISJLP, 7, 1.

7 Selby, E. A., Bender, T. W., Gordon, K. H., Nock, M. K., & Joiner Jr, T. E. (2012). Non-suicidal self-injury (NSSI) disorder: A preliminary study. Personality Disorders: Theory, Research, and Treatment, 3(2), 167.

Suicide refers to taking or attempting to take one's own life.

Eating disorders (ED) included in this report are anorexia (self-starvation), bulimia (starvation alternating with bingeing and purging) and binge eating (consuming unusually large amounts of food and not being able to stop); these are the most common eating disorders diagnosed in the mental health field.[8]

Glorification refers to extolling or admiring self-injurious behaviour. Celebration refers to praising one's self or others for engaging in self-injurious behaviour. Encouragement refers to any online post that invites others to engage in self-injurious behaviour, or provides explicit advice, assistance, or instructions to others to do so. We use the umbrella term 'promotes' (as in 'promotes self-injurious behaviour' or 'promotes eating disorders') to refer to content that glorifies, celebrates, or encourages self-injurious behaviour.

Coding Categories

The categories used to code the content of the links were developed through a systematic process. (See Appendix A for more details.)

Ofcom assisted us in determining how different types of content could be coded for the purposes of the research, including the development of guidelines for the research team on how to categorise the content in search results. To ensure consistency and enhance the reliability and validity of the coding, NCRI research staff conducted a series of initial meetings with the coders. During these meetings, all coders were trained on how to effectively utilise the categorisation system, aiming to establish a shared knowledge and understanding among them, and creating a set of guidelines for coding search result links. This was accomplished by ensuring coders were familiar with the coding categories and their descriptions (presented in Appendix A), understood to refer back to the descriptions as a way to resolve coding uncertainty, and engaging in group practice sessions, wherein coders assigned categories and had to explain and justify why they did so in group discussions with project leaders.

This process identified and eliminated misunderstandings of the system and enhanced the reliability and validity of the subsequent coding of the links analysed for this report.

The final coding categories were:

- Extreme refers to posts that encourage others to engage in self-injurious behaviour. This included treating self-injurious behaviour as a game; explicit invitations to join in self-injurious behaviour activities; and apps or other interactive spaces making it easy for people to engage in self-injurious behaviour.
- Likely to be in scope refers to posts that glorify or celebrate self-injurious behaviour on the part of oneself or others. They are likely to be in scope of promoting self-injury. It is similar to 'extreme' in that both celebrate self-injurious behaviour; but it is distinguishable from extreme in that it does not explicitly encourage others to engage in self-injurious behaviour. A post declaring, 'Come join in!' would be rated as extreme (if referring to some form of self-injurious behaviour) whereas a comment on a post displaying some form of self-injurious behaviour such as, 'That's beautiful,' would be rated as likely to be in scope.
- Prevention refers to posts that focused on getting people help (e.g., mental health services), explicitly argued against engaging in self-injurious behaviour, or provided educational information about the dangers of self-injurious behaviour.
- Ambiguous posts: 1. Were relevant to one or more self-injurious behaviours but it was unclear whether the subject was prevention or promotion; or 2. Included information that was some combination of both prevention and promotion, glorification or celebration.
- Unrelated/not relevant posts were none of the above.

When we discuss search engine content that promotes or is a promotion of self-injurious behaviour, we are referring to content that has been classified as likely to be in scope or extreme, because that content celebrates, glorifies, or instructs self-injurious behaviour.

8 Polivy, J., & Herman, C. P. (2002). Causes of eating disorders. Annual Review of Psychology, 53(1), 187-213.

METHODOLOGY

To answer the research questions, NCRI performed an analysis of the content produced by five major search engines in the United Kingdom – Google, Microsoft Bing, DuckDuckGo, Yahoo!, and AOL. All searches were conducted using a clean browser with no browsing history and exclusively using a Virtual Private Network (VPN) which assigned Manchester, England as the point of origin for all search queries. The study assessed the prevalence of content potentially conducive to self-injurious behaviour, and accessibility of that content in over 37,000 search results. It also assessed differences between search services and products, and patterns in self-injurious behaviour content. NCRI performed searches for three categories of self-injurious behaviour: NSSI, suicide, and eating disorders, and performed separate searches based on the type of content, referred to as search products: web (Safesearch off), image, and video. Ultimately, the final number of search results was 37,647. The search results were collected between March 3rd, 2023 through March 20th, 2023. Categorization of results started within that time period until May 20th, 2023.

All search results reported herein were 'one click away' – that is, when the search result was clicked, the next page that appeared was the page that was coded. To analyse the links, coding protocols were followed (see Appendix A for details). The links were classified into the following categories: unrelated/not relevant, prevention, ambiguous, likely to be in scope, and extreme. Two independent classifiers assessed each link. When the coders agreed on a category, that coding was final. Disagreements were identified and resolved by the coding team (see Appendix A for details on the process by which disagreements were resolved). The coding process was conducted meticulously, adhering to scientific standards for reliability and rigour when coding media.

The agreement rates for the first two coders provide a measure of the reliability of the coding process. The agreement rate for links related to NSSI was 79%, for suicide-related links it was 63%, and for eating disorder-related links it was 65%. The overall agreement rate across all classifications was 69%.[9] Details about how and why disagreements could reasonably arise are discussed in the section on Case Studies. More detailed information about the agreement rates, and on how disagreements were resolved, can be found in Appendix A. Nonetheless, a 69% agreement rate is the same as a 31% disagreement rate. Even though these disagreements were resolved by the tie-breaking processes discussed in Appendix A, this means there is some level of uncertainty regarding the reported results. The reported percentages are, therefore, best thought of as approximations of the prevalence of various types of content among our searches related to self-injury.

Analysis of the coding results focused on assessing:

1. Overall levels of content that was 'one click away' and promoted self-injurious behaviour in 37,647 search results.
2. Similarities and differences between:
    a. Type of self-injurious behaviour (NSSI, Suicide, ED)
    b. Search service (Google, Bing, DuckDuckGo, Yahoo, AOL)
    c. Search product (web, image, video)
3. Exploration of whether content promoting self-injurious behaviour was more common among certain domains.
4. Ease of access to content promoting self-injurious behaviour, measured as content appearing in the first, and top five search results.

Keywords and Surfacing Methods

NCRI's 'keyword surfacing' methods involved a combination of Open-Source Intelligence (OSINT) techniques – essentially, the collection and analysis of publicly available information – and data analytics to identify common words and phrases used online that promote NSSI, suicide, and eating disorders. To effectively assess content related to self-injurious behaviour, our comprehensive approach included both common phrases and terminology specific to online communities dedicated to self-injurious behaviours.

NCRI analysts conducted a survey of different online self-injurious communities to identify the most commonly used niche terminologies within and between communities promoting NSSI, suicide, and eating disorders. Analysts also identified keywords that are popular according to Google Trends and have been previously cited in papers on the topic of self-injurious

9 Macnamara, J. R. (2005). Media content analysis: Its uses, benefits and best practice methodology. Asia Pacific Public Relations Journal, 6(1), 1-34.

behaviour and the internet.[10, 11] This research identified 30 search terms (10 each for NSSI, suicide, and ED) that were then used to conduct web, image, and video searches using Google, Bing, DuckDuckGo, Yahoo, and AOL search engines.

The goal of this methodology was to produce a balanced, thorough and accurate analysis by:

1. Capturing community-specific language: By including niche terminology, the team uncovered and analysed content that may otherwise remain hidden and overlooked due to the use of obscured and cryptic language, slang, and acronyms within these communities.
2. Surfacing risks produced by data voids: Niche terminologies can create data voids due to the limited amount of information available on search engines. This lack of information can create an environment where harmful materials are more likely to surface and dominate search results.
3. Assessing general accessibility: Including common phrases allowed us to evaluate how easily accessible potentially harmful content is for individuals who are not familiar with the cryptic language of these online communities.
4. Ensuring more comprehensive coverage: The combination of niche terminology and phrases in common usage provided a more complete representation of the content landscape than had we focused on phrases in common usage alone.

Understanding Search Engines

For the purposes of this report, it is crucial to understand the basics of how search engines operate. Search engines do not host content; rather, they act as intermediaries, surfacing content in response to users' text or image queries. When a user enters a query, the search engine sifts through its vast index to provide the most relevant results. Algorithms used by search engines consider several factors, including but not limited to a page's popularity or authority, users' past search behaviour, and relevance of page content to the search query.[12, 13] This is why different users can receive slightly different results for the same search query, and why the same user can get different results at different times.

Content ranking is determined by a mix of factors which helps the service to surface content that will be the most relevant and useful for the user in relation to their specific search query. This means that motivated searchers who are seeking harmful content will often be able to find what they are looking for, particularly if the search query is very specific.

When surfacing relevant results, the search engine's algorithm considers a combination of factors including location and popularity to rank content. A user's history of engaging with content, including potentially harmful content can also at times be a factor.

ANALYSIS AND RESULTS

Research Question 1: Ease of Access

Our first set of analyses addressed our first research question: Of the 37,647 searches we conducted, how easily accessible is content that potentially encourages self-injurious behaviour? How much can be accessed by a single click of a search result? How much appears early in search results?

One Click Away

Every search result link we collected, coded, and reported was obtained after a single click of each link returned by the search engines in response to searches for NSSI, ED, and suicide related terms. Therefore, we use the phrase 'one click away' to characterise all search content analysed herein. When our research team coded each individual link, they assessed

10 Barros, J.M., et al. (2019). The validity of Google trends search volumes for behavioural forecasting of national suicide rates in Ireland. International Journal of Environmental Research and Public Health, 16(17), 3201. https://doi.org/10.3390/ijerph16173201

11 Arora, V. S., Stuckler, D., & Mckee, M. (2016). Tracking search engine queries for suicide in the United Kingdom, 2004–2013. Public Health, 137, 147-153.

12 Luh, C., Yang, S., & Huang, T.D. (2016). Estimating Google's search engine ranking function from a search engine optimization perspective. Online Information Review, 40, 239-255.

13 Ham, C. D. (2019). Why is this first? Understanding and analyzing internet search results. Journal of Educational Research and Practice, 9(1), 400-412.

the link as they found it – even if they were prompted to sign in, go through a paywall, or click through to any other part of the website, they always remained on the same page on which they started. The research team never clicked a second time to another page. This means that each of the links that contained content coded as promoting self-injurious behaviour was also 'one click away.' This is important because it means every link analysed in this report was easily available to any user conducting that search. These findings highlighted the role search engine providers play in enabling access to potentially harmful content. This includes children and individuals who come across such content unintentionally while searching for unrelated topics like recovery or educational information.

Ease of Access to Promotion of Self-Injurious Behaviour

Material that appears on the first page of a search is far more accessible and, therefore, more likely to be accessed than material that appears many pages into a search.[14] Therefore, the next analyses assessed the second part of this research question about ease of access: How much content that potentially encourages self-injurious behaviour appears early in search results? We analysed the accessibility of the material in two ways: 1. The frequencies with which the very first result of each search promoted self-injurious behaviour (Figure 1); and 2. The frequencies of self-injurious behaviour promotion among the first five search results (Table 1).

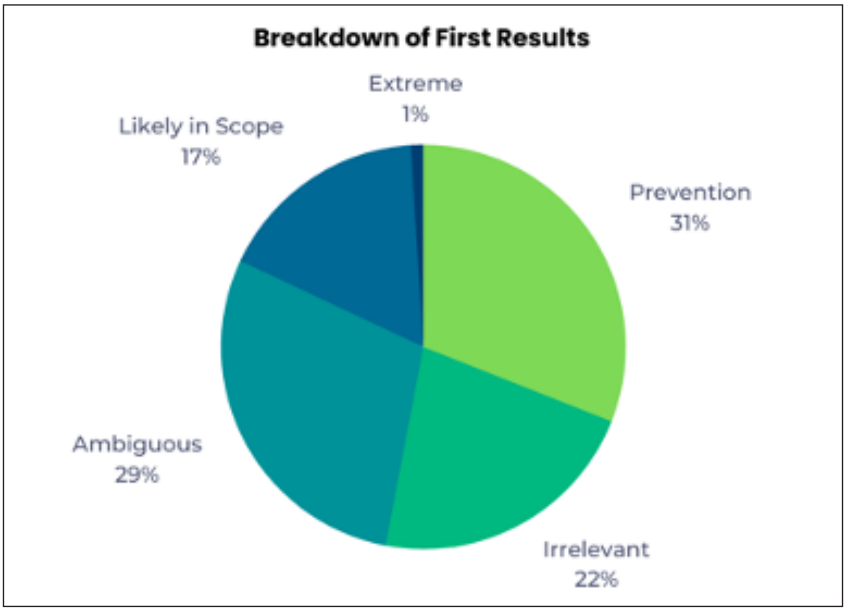


Figure 1: Breakdown of First Results by Classification

There were 438 first search results. Overall, 19% of these were coded as promoting self-injurious behaviour (18% likely to be in scope, 1% extreme), a pattern that did not vary much by search engine (all ranged from 15-21%). Another 29% were coded ambiguous.

Researchers analysed the first five search results because these will almost always show up on the first page of the results of a web search, and are, therefore, highly visible and easily accessible to whoever has conducted the search. Indeed, 70% of clicks are typically to one of the first five results. 22% of the results among the first five were also coded as promoting self-injurious behaviour and another 26% were ambiguous.

14 Beus, J. (2020). Why (almost) everything you knew about Google CTR is no longer valid. Sistrix. https://www.sistrix.com/blog/why-almost-everything-you-knew-about-google-ctr-is-no-longer-valid/

Table 1: Breakdown of First 5 Results by Classification

First 5 Results	Frequencies	Percentages
Prevention	608	28%
Unrelated/Not Relevant	510	23%
Ambiguous	580	27%
Likely to be in Scope	465	21%
Extreme	26	1%
Total	2189	100%

Overall, these findings indicate that content coded as promoting self-injurious behaviour appeared early in results produced by the search engines.

Research Question 2: Prevalence of Self-Injury Content

More than One in Five Search Results Were Coded as Promoting Self-Injurious Behaviours

These analyses addressed our second research question: *Among the 37,647 searches we conducted, how much content that potentially encourages self-injurious behaviour was delivered by search services?*

Across all results, over one in five were coded as promoting self-injurious behaviour (see Table 2 and Figure 2). Specifically, 21% were coded as likely in scope and another 1% were extreme, for a total of 22% that were coded as clearly promoting self-injurious behaviour in some way. An additional 23% of results were coded as ambiguous, meaning that many may have had information potentially interpretable as promoting self-injurious behaviour to at least some degree. Approximately one in five, 22%, were coded as seeking to prevent self-injurious behaviour.

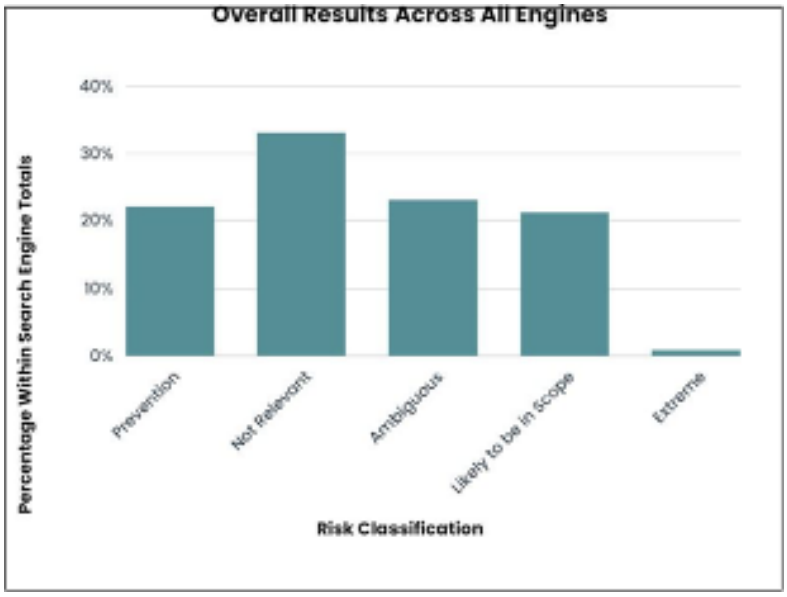


Figure 2: Distribution of Content Categories Across All Search Engines

Table 2: Overall Results

	Frequencies	Percentages
Prevention	8195	22%
Unrelated/Not Relevant	12600	33%
Ambiguous	8583	23%
Likely to be in Scope	7959	21%
Extreme	309	1%
Total	37647	100%

More Content Classified as Likely to Be In Scope/Extreme Pertained to Eating Disorders, Followed Closely by NSSI

The NCRI found some substantial differences between the search results for different types of self-injurious behaviour (see Table 3 and Table 4). The greatest number of search results were coded as promoting self-injurious behaviour occurred for eating disorders (29% likely to be in scope or extreme), followed by 25% likely to be in scope for NSSI and 12% likely to be in scope for suicide. Although the lowest level of promotion of self-injurious behaviour was found for suicide-related search results (12%), a number of results were coded as ambiguous (37%), meaning that preventing or promoting suicide was unclear or ambivalent in that content. Such sites would require a deeper analysis to fully evaluate the extent to which they encourage or seek to prevent self-injurious behaviour.

Table 3: Main Results by Type of Self-Injurious Behaviour, Frequencies and Percentages

	NSSI	Suicide	Eating Disorders
Prevention	2079 (17%)	3302 (25%)	2814 (22%)
Unrelated/Not Relevant	5450 (46%)	3294 (25%)	3856 (31%)
Ambiguous	1448 (12%)	4878 (37%)	2258 (18%)
Likely to be in Scope	2965 (25%)	1553 (12%)	3441 (27%)
Extreme	28 (<0.5%)	27 (<0.5%)	254 (2%)
Totals	11970 (100%)	13054 (100%)	12623 (100%)

Table 4: Overall Promotion Results, Frequencies and Percentages

	NSSI	Suicide	Eating Disorders
Likely to be in Scope + Extreme	2993 (25%)	1580 (12%)	3695 (29%)
Total	11970	13054	12623

Research Question 3: Comparing Search Engines

These Overall Patterns Did Not Vary Substantially By Search Engines

Our next set of analyses addressed the third research question: How, if at all, do answers to the above differ between search services (Google, Bing, DuckDuckGo, Yahoo, and AOL) and between search products (main results feed, image search, and video search)?

Table 5 and Figure 3 display the main results across search engines. Notably, the results appeared remarkably similar across search engines. Extreme results were coded as appearing in 1% or less of the results. Likely to be in scope was coded as appearing in 18-26% of results across search engines (highest for Google at 26% and lowest for AOL at 18%). Ambiguous results were coded as appearing slightly above 20% across the search engines. Prevention, too, was very similar across search engines, with Google highest at 28% and the remainder appearing in 20% of the results.

Despite differences in usage patterns and search algorithms, all the major search engines in our sample consistently: 1. Produced a small proportion of results that were coded as promoting self-injurious behaviour; and 2. Produced similar patterns of results (see Appendix B for detailed results specific to each platform).

Table 5: Main Results by Search Engine, Frequencies and Percentages

	Google	Bing	DuckDuckGo	Yahoo	AOL
Prevention	2487 (28%)	1553 (20%)	1407 (20%)	1281 (20%)	1467 (20%)
Not Relevant	1974 (22%)	3070 (39%)	2578 (37%)	2032 (32%)	2946 (40%)
Ambiguous	2111 (24%)	1735 (22%)	1504 (22%)	1629 (25%)	1604 (22%)
Likely to be in Scope	2275 (26%)	1506 (19%)	1426 (20%)	1413 (22%)	1339 (18%)
Extreme	58 (<1%)	54 (<1%)	65 (<1%)	81 (<1%)	51 (<1%)
Total	8905 (100%)	7918 (100%)	6980 (100%)	6436 (100%)	7408 (100%)

Results within columns sometimes do not add up to 100% due to rounding.

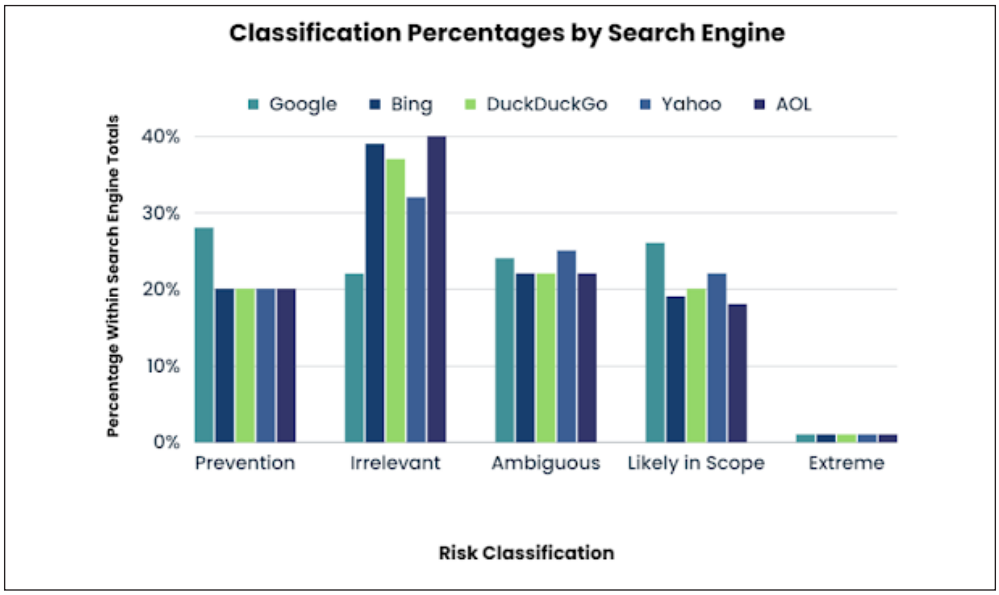


Figure 3: Main Results by Search Engine

Research Question 4: Where are Disproportionate Amounts of Content Promoting Self-Injurious Behaviour Found?

We also performed analyses addressing our fourth research question: What, if any, patterns are there in self-injurious behaviour content that is accessible from search services? These analyses assessed differences in the content produced by different types of searches, and whether certain domains were particularly likely to host content promoting self-injurious behaviour.

Image Search Results Were Disproportionate Sources of Self-Injury Content

The percentages in Table 6 are in relation to the total number of links across all search services categorised as likely to be in scope or extreme, amounting to 8268 links, which served as the denominator in this analysis. Among these links, images constituted 50% of the total. Table 7 presents the full breakdown of all results for web, image and video searches.

Table 6: Proportion of Potentially Harmful Content by Search Function, Frequencies and Percentages

	Web	Images	Videos
Likely to be in Scope + Extreme	2294 (28%)	4157 (50%)	1817 (22%)

Percentages in this table are based on the 8268 results coded as likely to be in scope or extreme.

This research found that image searches across all five search engines were coded as producing a higher proportion of results that were likely in scope or extreme (50%) compared to web results (28%) and video results (22%). This may be due to two reasons. First, while detection algorithms identify some extreme content in images, the complexity of analysing visual data makes it difficult to effectively detect this content in general.[15] Distinguishing images depicting, for example, glorification of intentionally self-inflicted wounds from medically relevant visuals might be a difficult task. The intricate variations in appearance, ranging from genuine medical images to potentially sensitive content, require algorithms to comprehend subtle contextual cues and nuances. Second, image searches may be more likely to contain the most extreme NSSI, suicide, and ED content because images have a powerful visual impact, evoking stronger emotional responses

15 Yang, G., Yang, S., Luo, K., Lan, S., He, L., & Li, Y. (2023). Detection of non-suicidal self-injury based on spatiotemporal features of indoor activities. IET Biometrics, 12(2), 91-101. https://doi.org/10.1049/bme2.12110

and potentially influencing individuals to engage in self-harming behaviours.16 The far higher number of results that were potentially harmful among image searches, as compared to both web and video searches, may indicate heightened risk associated with image searches. As one study on the influence of online images on self harm posits, 'images rather than textual interactions are the primary reason cited for using the Internet for self-harm purposes'.16 Images can evoke a physical reaction, and adolescents and young adults claimed images were more likely than text content to inspire them to engage in acts of self-injury.[16]

Images displaying self-injurious behaviour, or its effects, might evade detection by automated moderation systems more readily than text.[17] Text-based moderation has advanced significantly, but the technology to accurately and consistently identify harmful imagery is still evolving.[18]

Table 7: Main Results by Search Product, Frequencies and Percentages

	Web	Image	Video
Prevention	3774 (33%)	2400 (16%)	2021 (18%)
Not Relevant	2965 (26%)	5027 (34%)	4608 (40%)
Ambiguous	2402 (21%)	3116 (21%)	3066 (27%)
Likely to be in Scope	2169 (19%)	3991 (27%)	1799 (16%)
Extreme	125 (1%)	166 (1%)	18 (<0.5%)
Totals	11435 (100%)	14700 (100%)	11512 (100%)

Results within columns sometimes do not add up to 100% due to rounding.

Domains That Appear Most Often in Search Results

We next examined whether any particular sites or platforms were especially likely to host content celebrating self-injurious behaviour. Results of these analyses were: 59% of all content coded as promoting self-injurious behaviour (8268 sites that were coded as likely to be in scope or extreme) could be found on three major social media sites.

Community Cryptic Language and Data Voids: A Significant Challenge for Content Moderation

The data from our research revealed a pattern pertaining to the use of community-specific words: they generated more search returns that were categorised as likely to be in scope or extreme than search returns using general terms and phrases such as 'how to commit suicide', which were coded more as prevention or not relevant. Community-cryptic words often include abbreviations, insertions of symbols into real words to mask them, or use of homonyms to mask the content. This serves two purposes: It signals to the community of those engaging in self-injury that 'this post is for you', and it serves to mask the content to avoid detection by text filters designed to block them.

To illustrate how cryptic terms are constructed – imagine if one was part of a society in which dancing was stigmatised, but one wished to engage and promote dancing surreptitiously. Such a community might create hashtags and formulate words such as #datwt (dance twitter), d@nse, dants, dbd (dance baby dance), dnzr (dancer) and the like to connect with others who dance and to avoid detection by algorithms looking for 'dance,' 'dancer,' or 'dancing.'

16 Jacob, N., Evans, R., Scourfield, J. (2017). The influence of online images on self-harm: A qualitative study of young people aged 16-24. Journal of Adolescence, 60, 140-147. https://doi.org/10.1016/j.adolescence.2017.08.001.
17 Conti, M., Pajola, L., & Tricomi, P. P. (2023). Turning captchas against humanity: Captcha-based attacks in online social media. Online Social Networks and Media, 36, 100252. https://doi.org/10.1016/j.osnem.2023.100252
18 Farid, H. (2021). An overview of perceptual hashing. Journal of Online Trust and Safety, 1(1). https://doi.org/10.54501/jots.v1i1.24

NCRI's surfacing methods identified such cryptic words used among communities of users to discuss NSSI, suicide and eating disorders. Of the 22% of search results that were coded into the categories likely to be in scope or extreme, 19% were community-cryptic words (like 'd@nse'). Only 3% of these results were generated through terms in general usage (like 'dance'). These findings may point to the existing moderation mechanisms' relative effectiveness at identifying and filtering out harmful content when search terms are common words in general use. However, when terms are cryptic, search engines appear to have a harder time identifying, and, therefore, filtering out harmful content. Continuing with the hypothetical example involving dancing above, it is as if search engines do a much better job at filtering out 'dance' than 'd@nse.'

This discovery also underscores a pervasive and challenging issue of what can be termed 'data voids.' 'Data voids' refer to situations where the search demand for certain keywords is not met with reliable or safe information due to relative obscurity of search terms or phrases. Searches utilising cryptic language may lead to more harmful content, as algorithms aim to provide relevant results, but lack necessary safe and accurate information to fill these voids.[19]

CASE STUDIES ON NSSI, ED, AND SUICIDE CONTENT CLASSIFICATION

Content Warning: This section contains some explicit images and text discussions of distressing content related to self-harm, suicide, and eating disorders, using real-world examples (case studies) of content found during this research.

Throughout the content classification process, researchers encountered cases that varied in their level of difficulty to categorise. Some cases exhibited clear patterns, identifiable characteristics, or distinct outcomes, which made it relatively straightforward to assign them to specific categories. These categorisations were based on observable factors and facilitated relatively easy classification. However, it is important to note that coding certain types of content posed challenges when its meaning or intent was ambiguous and subjective in terms of perceived promotion of self-injurious behaviours. Categorising content of this nature required careful consideration of the context and intended audience to make decisions around user intent or level of risk. We next present case studies of easily-coded results for which there was agreement, and more difficult examples on which there was disagreement.

The following images are examples of categorisations that were agreed upon by all coders. Shown below (Image 1) is a result for a search of eating disorder terms, linking to a social media post, which was coded as likely to be in scope. It displays a variety of near-starvation, low-calorie diets. The intent is plausibly to inform or enlist others to use these diets, which might deserve to be coded as 'extreme.' However, because it does not say 'You should try this!' or 'Who wants to join me?' or something similar that explicitly encourages others to adopt one of these diets, it was coded as likely to be in scope. This type of result was very common for eating disorders, though eating disorder results were also more likely to be coded as extreme than were either NSSI or suicide results.

Image 1: a Twitter post on December 20, 2020 with multipe guides to low-calorie dieting.[20]

19 Chancellor, S., Pater, J. A., Clear, T., Gilbert, E., & De Choudhury, M. (2016, February). # thyghgapp: Instagram content moderation and lexical variation in pro-eating disorder communities. Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing (pp. 1201-1213).

20 https://twitter.com, March 2023. post has since been removed by Twitter(X).

Image 2, which has been omitted from this report due to its graphic depiction of self-harm, is a Bing result for a search of NSSI terms. It reads 'keep calm and cut yourself' with scars from cuts visible just below the word 'CALM.' If this image had shown only scars, it would have been coded as likely to be in scope. However, the message 'keep calm and cut yourself' made it extreme, as it encourages others to engage in self-injurious behaviour.

The following result (Image 3) was generated from a Yahoo search using suicide-related terms. In these search results, videos feature fictional characters who romanticize the dramatic aspects of suicide. These videos belong to users within the suicide community who create and exchange this type of content among themselves. These users frequently act out different scenarios involving self-injurious behaviour and suicide. In the video below, the character on the left has hung themselves. The video in question was coded as likely in scope due to its graphic portrayals of suicide.

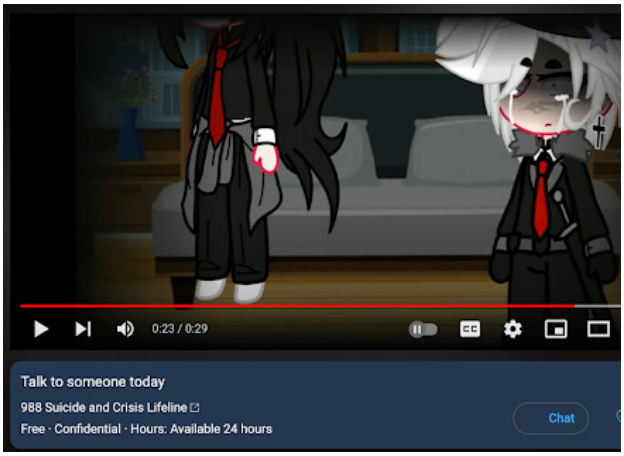


Image 3: an animated YouTube video posted on X which dramatizes suicide[21]

This result (Image 4) is from a DuckDuckGo eating disorders search. It was coded as ambiguous because it is a video from a news site discussing eating disorder trends on social media. Despite the video being educational in content, it is not prevention because it includes sufficiently detailed personal experiences which may be triggering to someone with an ED, but it is not likely to be in scope because those personal experiences do not include promotion or glorification of ED. Even when framed as prevention, detailed personal experiences can be interpreted almost as an instruction manual on how to engage in a type of self-injurious behaviour. Therefore, our 'ambiguous' category specifically included such detailed personal descriptions, and this was coded as such.

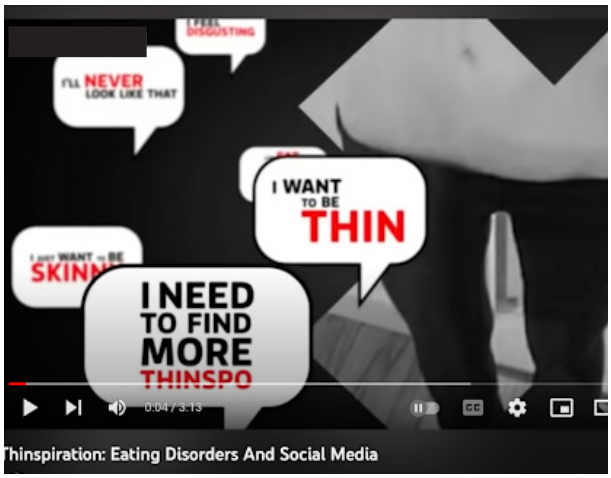


Image 4: a 2015 report from Sky News posted on YouTube discussing ED trends on social media[22]

21 https://www.youtube.com, March 2023.
22 https://www.youtube.com/watch?v=WU-JMYnEgyI, March 2023.

The provided link (Image 5) directs to an AOL search result page related to NSSI terms. This particular result was marked as ambiguous because it leads to a song that has been created and titled with a cryptic term known within the NSSI online community. However, it is important to note that the song falls into the category of art that does not explicitly glorify self-harm. It is ambiguous because it is relevant to self-injury but whether the song is a celebration or attempt at prevention is not clear; indeed, it might be interpretable as either by different people.

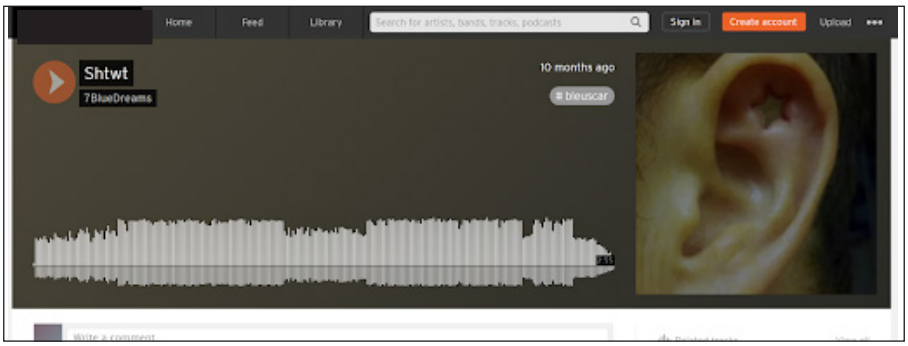


Image 5: a song posted on SoundCloud in March 2023 with implicit references to NSSI[23]

The webpage shown below (Image 6) is a clear example of a prevention search result found across search engines for NSSI terms. The tone of the website is support, not glorification, and 'we're here to listen' was interpretable as an attempt to make some sort of psychological services available to those who might contact them.

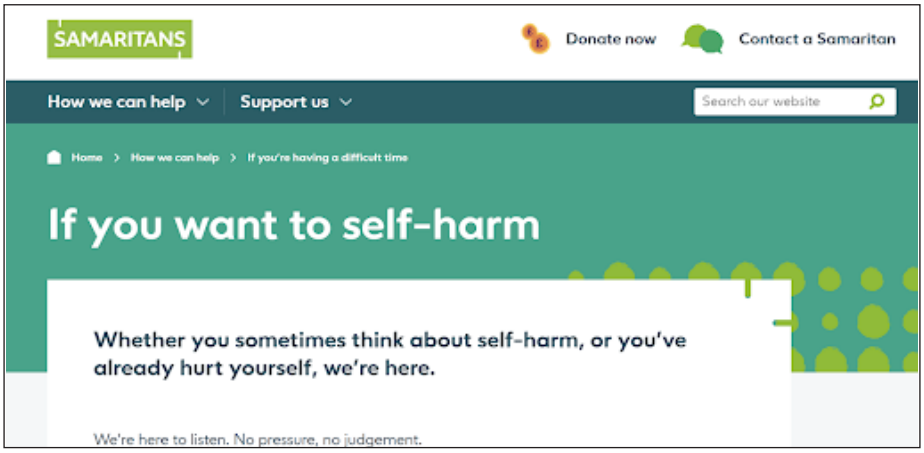


Image 6: samaritans.org prevention search result[24]

However, there were also instances of coder disagreement which required tiebreakers. These cases were more complex and ambiguous, making it challenging to assign them to a specific category without considering additional information or engaging in the type of thorough deliberation that was beyond the scope of this project. When the initial pair of coders disagreed, a third coder independently evaluated the search result (without knowing the first two codings that disagreed) and functioned as a tiebreaker.

For example, a video emerged as a complex case during a Bing search conducted within the suicide category. The video features a young woman sharing her personal experience of working in a psychiatric hospital and helping people experiencing suicidal ideation or who have attempted suicide in the past. Upon initial evaluation, one coder categorised it as ambiguous, while the other coder classified it as prevention. To reach a final decision, a tiebreaker was involved, and ultimately, the video was classified as falling within the prevention category.

23 https://soundcloud.com, March 2023.
24 https://www.samaritans.org/how-we-can-help/if-youre-having-difficult-time/if-you-want-self-harm/, March 2023.

The initial differing classifications of the video may arise due to subjective interpretation or individual perspectives of the coders involved in the evaluation process. It is not always obvious whether web content aimed at 'supporting' people who engage in various forms of self-injurious behaviour intends to get them the help they need to prevent such behaviour in the future or to 'support' them to engage in these behaviours even more. It highlights the inherent challenge of assessing and categorising content that can be open to interpretation, leading to varying initial classifications by different coders.

ADDITIONAL FINDINGS

Gaps in Terms of Services and Community Guidelines

Terms of Service (TOS) and Community Guidelines of online platforms seek to govern user experience and create a safe online environment. Previous research published in 2021 by the Stanford Internet Observatory highlighted concerns regarding the adequacy of self-harm related standards in search engine policies.[25]

The TOS and/or Community Guidelines of search engines like Google,[26] Bing,[27] DuckDuckGo,[28] Yahoo,[29] and AOL[30] primarily focus on regulating user interactions within their platforms, addressing broader issues such as harassment, hate speech, violence, and explicit content. References to NSSI, suicide, and eating disorder content are either not present or very limited among TOS and/or Community Guidelines of the search engines we assessed. These omissions raise questions about the comprehensive nature of these policies as it relates to self-injurious content, although these issues are beyond the scope of our report.

CONCLUSIONS

Ofcom, the UK's communications regulator, commissioned the Network Contagion Research Institute (NCRI) to conduct this report on content accessible via search services related to three types of self-injurious behaviour: non-suicidal self-injury, suicide and eating disorders.

This groundbreaking research has yielded novel findings on trends and patterns in online content that promote these self-injurious behaviours.

We found that, for the terms searched, more than 1 in 5 returns from search engines were coded as promoting likely in scope or extreme content (although extreme was rare, typically about 1%). Although such content could be found in any type of search we conducted, image searches were more likely than web or video searches to produce results promoting self-injurious behaviour.

The research also found that such content is easily accessible, often appearing high up in search results across all major search engines. Notably, the research revealed a particularly pronounced role played by social media networks in populating content coded as extreme and likely in scope in the search engine results we analysed — over half of the returns for such content in this sample linked to just three social media sites.

Furthermore, our findings suggest the use of cryptic language constitutes an especially important dimension in the reach of self-injurious content and these findings highlight key vulnerabilities for search engines specifically. Members of self-injury communities use specific jargon or alternative terms to communicate about their interests and activities while avoiding detection by content moderation systems employed by search engines. Our research revealed that it is six times more likely that a user may encounter extreme or likely in scope content on search platforms if community-specific keywords are used compared to general terms.

In their totality, the findings of this report suggest that search engines pose a potential risk to users suffering with mental health issues or to children, because they serve as an open portal to content that encourages harmful behaviours related to NSSI, suicide, and eating disorders.

25 Perkins, S., Cryst, E., & Grossman, S. (2021). Self-Harm Policies and Internet Platforms. Stanford Internet Observatory Cyber Policy Center. https://stacks.stanford.edu/file/druid:tm443wf7913/20210408-Self-Harm-Policies-And-Internet-Platforms.pdf
26 Google. (2023) Privacy and Terms. https://policies.google.com/terms
27 Microsoft. (2022). Microsoft Services Agreement. https://www.microsoft.com/en-us/servicesagreement
28 DuckDuckGo. (2023) DuckDuckGo Terms of Service. https://duckduckgo.com/terms
29 Yahoo. (2021). Yahoo Community Guidelines. https://legal.yahoo.com/ie/en/yahoo/guidelines/index.html
30 AOL. (2021). Guidelines for Comments on AOL. https://help.aol.com/articles/Guidelines-for-Comments-on-AOL

APPENDICES

Appendix A: System for Coding Search Results

Link Collection - Protocols - Query Results and Recording:

All searches were conducted exclusively using a Virtual Private Network (VPN) which assigned Manchester, England as the point of origin for all search queries. Searches were performed on Google, Bing, DuckDuckGo, Yahoo, and AOL.

Each search term was entered into the five different search engines. The three main types of searches performed were web search, image search, and video search. For each of the three search categories (web search, image search, and video search), the first 100 links were collected. This yielded a possible total of 300 links per search term.

This created 15 categories of searches: five search engines multiplied by three categories of self-injurious behaviour. Then within each of these 15 categories, NCRI performed separate searches based on the type of content: web (Safesearch off), image, and video, producing 45 total categories of searches.

Within each of those 45 categories, NCRI searched for 10 terms, producing 450 searches. NCRI also coded the first 100 links in response to those 450 searches. This would have produced a total of 45,000 searches (450 multiplied by 100), but some search results did not generate 100 links. Any search category that produced fewer than 100 links still had all the available links gathered and included in the analysis. In total, 37,647 links were successfully recorded.

Table A.1: Descriptions of Categories Used to Code Search Results

Category	Percentages
Unrelated/Not Relevant	Unrelated to self-injurious behaviour of any kind
Prevention	Related to NSSI, suicide, or eating disorders. Content is educational in nature from a promoted or verified source in addressing self-injurious behaviour (e.g., a news article, charity helpline) OR provides coping mechanisms or resources related to prevention of or recovery from self-injurious behaviour.
Ambiguous	Related to NSSI, suicide, or eating disorders, but whether preventative or promoting them is unclear, open to interpretation. Also includes content that provides mixed messages, both pro- and anti-self-injurious behaviours. Personal experiences relating to recovery may be 'ambiguous' (instead of 'prevention') if sufficiently detailed so as to pose a risk of harm. Content related to 'safe' self harm where the intent is to protect, not enable further self harm. Definitions of self-injurious behaviour terms (e.g., Wikipedia articles, online dictionaries). Feeds sharing information or discussing recent events, broadcasting of a TV programme, film with related themes or memorial pages for individuals who have died by suicide. Art (including poetry, movies, fictional stories and books, songs, playlists, music videos). Euthanasia/assisted suicide content.

Category	Percentages
Likely to be in Scope	Content that promotes, glamorises, glorifies, celebrates, romanticises, or normalises NSSI, suicide, or eating disorders in a context not clearly linked to prevention or recovery. Graphic and non-graphic depictions/descriptions of NSSI, suicide, or eating disorders (this could include art if it encourages or glamourises).
Extreme	Content that encourages or overtly suggests users take action to commit serious self-injury or provides instructional guidance about how to commit serious self-injury. Overt promotion of engaging in self-injurious behaviour activity such as interactive spaces (apps, forums, chat rooms, games, etc.)

Coding followed the guidelines below to classify each link:

1. Two people were assigned to independently classify each link into one of the five categories without consulting or discussing with each other. They were instructed to spend no more than one minute classifying each link. Sometimes, links were classified as ambiguous because a one-minute examination was not enough to make a determination and a deeper dive would have been required to disambiguate.
2. If the two classifiers agreed on the category, the link is considered classified, and no further action was taken.
3. If the two classifiers disagreed on the category, a third person (the tie-breaker) was brought in to independently classify the link. The tie-breaker assigned the link to one of the five categories.
4. If the tie-breaker agreed with one of the initial classifications, that classification was used and the link was considered classified. If a tie-breaker is a member of the leadership team, their classification was used. For all links that required another tie-breaker, a member of the leadership team reviewed the classifications. If that member agreed with the tie-breaker, that was the classification that was used. If the leadership team member did not agree with the tie-breaker, a second member of the leadership team was brought in, and they discussed it until they reached agreement, which was then the classification that was used.
5. If the tie-breaker disagreed with both initial classifications, one or more members of the leadership team were brought in to determine if there was an error and classify the link. An error was defined as the link clearly belonging to one of the three categories. This member(s) of the leadership team corrected the error, or if there was none, the link was classified as ambiguous.
6. All classification decisions were recorded and kept on file for reference.

Agreement Rates

Methodological standards recommend independent classification by multiple individuals before involving a tie-breaker or a leadership team (see footnote 10). This approach mitigates biases and enhances objectivity in the classification process. The requirement for independent classification before seeking consensus minimises the risk of groupthink or undue influence on the initial judgments.

To ensure the reliability and credibility of the coding agreement process, all classifications were documented and retained for audits, reviews, and comparisons, facilitating quality control, and ensuring compliance with industry standards. Keeping records of the classifications provides a historical reference that can be utilised for future analysis, benchmarking, and improvement of the classification process.

Continuous monitoring and evaluation of the agreement rates were routine to identify areas of improvement and address potential sources of disagreement. Regularly reviewing the agreement rates helps organisations identify patterns, discrepancies, or challenges that may arise during the classification process. This ongoing assessment allowed for the refinement of guidelines, training, and decision-making protocols to enhance reliability and consistency.

Experiences of Professionals on the Frontline

In conducting this study, a team of professional researchers, including psychology professors, clinical psychiatrists, Ph.D. candidates, and trained analysts and data scientists extensively engaged with harmful online content related to NSSI, eating disorders, and suicide. While the NCRI had implemented safeguards to support the team, it is important to acknowledge the disturbing nature of some of the content and the impact it had on the researchers, who were trained professionals. The graphic images, depicting self-harm by teenagers and young adults in traumatic ways, affected team members, causing some distress. The inclusion of professional psychologists as collaborators mitigated these effects in ways that permitted the project to be conducted. Nonetheless, it is crucial to recognise the severity of some of the content and the need for safeguarding measures, especially when considering the potential effect on unequipped internet users, particularly children and young adults. These vulnerable groups, often unaware of the potential harm associated with such content, may lack the necessary knowledge, psychological resilience and supportive resources to avoid or cope with its negative impact. The importance of awareness of these risks and safeguarding against them should not be underestimated.

Appendix B: Detailed Results by Search Engine

Google Results

More Than One in Four Google Search Results Promote Self-Injurious Behaviour

Across all searches, NCRI found that 26% of the results we analysed on Google lead to sources that promote self-injurious behaviour. The key results are shown in Table B.1. The 26% figure combines results likely to be in scope and extreme. However, that figure is conservative. An additional 24% of search results were ambiguous, meaning that some are plausibly interpretable by at least some people as promoting self-injurious behaviour.

Table B.1: Overall Results for Google Searches (NSSI, Suicide, ED)

	Frequencies	Percentages
Prevention	2487	28%
Unrelated/Not Relevant	1974	22%
Ambiguous	2111	24%
Likely to be in Scope	2275	26%
Extreme	58	1%
Total Number of Results	8905	100%

These patterns were consistent across different search products (web pages, images, or videos), with 22-29% of results either likely to be in scope or extreme, and another 23-24% ambiguous.

However, there were stark differences between types of self-injurious behaviour. As Table B.2 shows, 39% of the results for NSSI were coded as likely to be in scope or extreme, whereas this was true for only 27% of eating disorder results and only 13% of suicide results.

Table B.2: Results for Google Searches, Separately for NSSI, Suicide, and Eating Disorder Searches: Frequencies and Percentages

	NSSI	Suicide	Eating Disorders
Prevention	752 (25%)	834 (28%)	901 (31%)
Unrelated/Not Relevant	791 (26%)	572 (19%)	611 (21%)
Ambiguous	285 (10%)	1206 (40%)	620, 21%)
Likely to be in Scope	1168 (39%)	386 (13%)	721 (25%)
Extreme	4 (0.13%)	2 (0.07%)	52 (2%)
Total Number of Results	3000	3000	2905

Left = total count; right = proportion of all results for that type of self-injurious behaviour.

Promotion of Self-Injurious Behaviour Comes up Early in Google Searches

At least 20% of the initial results from the Google searches the NCRI conducted were coded as promoting self-injurious behaviour. Across all of our Google searches, 20% of the very first results produced were coded as likely to be in scope and extreme, and another 31% were coded as ambiguous. This pattern was particularly true for video searches (31% likely in scope and extreme and another 31% ambiguous). A lower, but still high level of harm self-promotion was found for image searches (23% were coded as likely to be in scope and extreme and another 20% ambiguous). Web searches produced the lowest level of results coded as likely to be in scope and extreme (7%), but a large portion (43%) were coded as ambiguous. Overall, therefore, 40-60% of first results possibly included content that was, or was plausibly interpretable as, promoting self-injurious behaviour.

With respect to the first five results, a similar pattern emerged but the percentages were somewhat higher. Overall, 22% of these results among the first five were coded as likely to be in scope and 2% were extreme. Another 27% of these searches were coded as ambiguous, for a grand total of 52% that were or might be promoting self-injurious behaviour (the sum does not always equal the total because of rounding decimals). Again, the highest frequencies were found for videos (31% among the first five were likely to be in scope or extreme and another 29% ambiguous). 23% of image searches produced among the first five were coded as likely to be in scope and another 27% ambiguous. 20% of web searches among the first five were coded as likely to be in scope or extreme, and another 24% were coded as ambiguous.

Bing Results

At Least One in Five Bing Searches Produce Results that Promote Self-Injurious Behaviour

Across all searches conducted in the study, NCRI found that 21% of Bing results lead to sources that were coded as promoting self-injurious behaviour. The key results are shown in Table B.3. The 20% figure combines results likely to be in scope and extreme. However, that figure is conservative. An additional 22% of search results were coded as ambiguous, meaning that some are plausibly interpretable by at least some people as promoting self-injurious behaviour. Thus, over 2 in 5 of all Bing search results that were found in NCRI's searches do or might promote self-injurious behaviour.

Table B.3: Overall Results for Bing Searches (NSSI, Suicide, and ED)

	Frequencies	Percentages
Prevention	1553	20%
Unrelated/Not Relevant	3070	39%
Ambiguous	1735	22%
Likely to be in Scope	1506	19%
Extreme	54	0.68%
Total Number of Results	7918	100%

For Bing searches, there were definite differences between the results across types of searches (web, image, video, see Table B.4). Whereas 30% of image searches were likely in scope or extreme, only 17% and 10% of web and video searches (respectively) were likely in scope or extreme (rounding means total percentages do not always equal the sum of the individual percentages). Including ambiguous results, well over half of all image searches (57%), and about a third (31% of web searches and video searches included material potentially promoting self-injurious behaviour.

Table B.4: Frequencies by Search Product for Bing

	Web	Image	Video
Prevention	872 (36%)	263 (9%)	418 (17%)
Unrelated/Not Relevant	809 (33%)	1031 (34%)	1230 (50%)
Ambiguous	348 (14%)	807 (27%)	580 (24%)
Likely to be in Scope	384 (16%)	885 (30%)	237 (10%)
Extreme	38 (2%)	14 (0.47%)	2 (0.08%)
Total Number of Results	2451	3000	2467

Left = total count; right = proportion of all results for that type of search (web, image or video).

There were few major differences with respect to type of self-injurious behaviour. 20%, 12% and 27% of results were coded as likely to be in scope or extreme for NSSI, suicide, and ED, respectively. Including ambiguous results, findings were 34%, 44% and 45%, respectively, for NSSI, suicide, and ED.

Table B.5: Results for Bing Searches, Separately for NSSI, Suicide, and Eating Disorder Searches: Frequencies and Percentages

	NSSI	Suicide	Eating Disorders
Prevention	315 (12%)	792 (29%)	446 (17%)
Unrelated/Not Relevant	1357 (53%)	699 (26%)	1014 (38%)
Ambiguous	370 (15%)	871 (32%)	494 (18%)
Likely to be in Scope	500 (20%)	315 (12%)	691 (26%)
Extreme	6 (0.24%)	9 (0.34%)	39 (1%)
Total Number of Results	2548	2686	2684

Left = total count; right = proportion of all results for that type of self-injurious behaviour.

Promotion of Self-Injurious Behaviour Comes up Early in Bing Searches

The next set of analyses assessed the frequency with which promotion of self-injurious behaviour appears among the very first result of a search. Across types of searches (web, image, video), 15% of the very first results were coded as either extreme or likely in scope and another 30% were coded as ambiguous. The highest concentration of self-injurious behaviour on Bing occurred for image searches (23% likely in scope and another 30% ambiguous). Web searches produced 23% first results that were coded as either likely in scope or extreme and another 30% that were coded as ambiguous. For video searches, only 4% of the first results were coded as extreme or likely in scope, but 43% were coded as ambiguous.

Overall, therefore, about 40-50% of first results possibly included content that was, or was plausibly interpretable as, promoting self-injurious behaviour.

With respect to how many of the first five results for each search term promoted self-injurious behaviour, a similar pattern emerged but the percentages were somewhat higher. Overall, 19% of these results among the first five were coded as likely in scope and 1% were coded as extreme. Another 28% of these search results were coded as ambiguous, for a grand total of 48% that were or might be promoting self-injurious behaviour. The highest frequencies were found for image searches (27% among the first five were coded as likely in scope or extreme and another 22% ambiguous). 19% of the results for web searches among the first five results were coded as likely in scope, 3% were extreme, and another 27% ambiguous. 9% of video searches produced among the first five were coded as likely in scope, 1% were extreme, and another 31% were ambiguous.

DuckDuckGo Results

At Least One in Five DuckDuckGo Searches Produce Results that Promote Self-Injurious Behaviour

Across all searches, NCRI found that 21% of DuckDuckGo results led to sources that were coded as promoting self-injurious behaviour. The key results are shown in Table B.6. The 21% figure combines results likely to be in scope and extreme. However, that figure is conservative. An additional 22% of search results were coded as ambiguous, meaning that some are plausibly interpretable by at least some people as promoting self-injurious behaviour.

Table B.6: Overall Frequencies for DuckDuckGo Searches (NSSI, Suicide, and ED)

	Frequencies	Percentages
Prevention	1407	20%
Unrelated/Not Relevant	2578	37%
Ambiguous	1504	22%
Likely to be in Scope	1426	20%
Extreme	65	1%
Total Number of Results	6980	100.00%

For DuckDuckGo searches, there were some differences between the results across types of searches (web, image, video, see Table B.7). Whereas 28% of image searches were coded as likely in scope or extreme, 17% and 10% of web and video searches (respectively) were coded as likely in scope or extreme. Including ambiguous results, 44% of web and image searches, and over a third (38%) of video searches included material potentially promoting self-injurious behaviour.

Table B.7: Frequencies by Search Product for DuckDuckGo

	Web	Image	Video
Prevention	444 (24%)	554 (19%)	409 (19%)
Unrelated/Not Relevant	578 (31%)	1044 (36%)	956 (44%)
Ambiguous	389 (21%)	494 (17%)	621 (29%)
Likely to be in Scope	433 (23%)	800 (27%)	193 (9%)
Extreme	27 (1%)	38 (1%)	0 (0%)
Total Number of Results	1871	2930	2179

Left = total count; right = proportion of all results for that type of search (web, image, or video).

For DuckDuckGo searches, 22%, 11% and 31% of results were coded as likely to be in scope or extreme for NSSI, suicide, and ED, respectively. Including ambiguous results, findings were 33%, 50% and 45%, respectively, for NSSI, suicide, and ED (totals are subject to rounding and may not exactly equal the sum of the values displayed; see Table B.8).

Table B.8: Frequencies by Search Product for DuckDuckGo

	NSSI	Suicide	Eating Disorders
Prevention	348 (16%)	529 (22%)	530 (22%)
Unrelated/Not Relevant	1085 (51%)	682 (28%)	811 (33%)
Ambiguous	237 (11%)	916 (38%)	351 (14%)
Likely to be in Scope	457 (21%)	271 (11%)	698 (29%)
Extreme	7 (0.33%)	5 (0.21%)	53 (2%)
Total Number of Results	2134	2403	2443

Left = total count; right = proportion of all results for that type of self-injurious behaviour.

Promotion of Self-Injurious Behaviour Comes up Early in DuckDuckGo Searches

The next set of analyses assessed the frequency with which promotion of self-injurious behaviour appears among the very first result of a search. Across types of searches (web, image, video), 20% of the very first results were coded as either extreme or likely in scope and another 27% were ambiguous. Most instances of promoting self-injurious behaviour on DuckDuckGo occurred for image searches (27% were coded as likely in scope and another 23% ambiguous). Web searches produced 20% first results that were coded as likely in scope and another 27% that were ambiguous. For video searches, 12% of the first results were coded as extreme or likely in scope, but 32% were ambiguous. Overall, therefore, about 40-50% of first results possibly included content that was, or was plausibly interpretable as, promoting self-injurious behaviour.

With respect to how many results among the first five potentially promoted self-injurious behaviour, a similar pattern emerged but, the percentages were somewhat higher. Overall, 25% of these results among the first five were coded as likely in scope and 1% were extreme. Another 24% of these searches produced results that were coded as ambiguous, for a grand total of 50% that were or might be promoting self-injurious behaviour. The highest frequencies were found for web searches, which, among the first five, were coded as likely in scope or extreme 34% of the time, and another 21% were ambiguous. Image searches, among the first five, were coded as likely in scope or extreme 28% of the time, and another 19% ambiguous. 12% of video searches among the first five were coded as likely in scope and another 24% were coded as ambiguous.

Yahoo Results

Over One in Five Yahoo Searches Produce Results that Promote Self-Injurious Behaviour

Across all searches, NCRI found that 23% of Yahoo.co.uk results led to sources that promote self-injurious behaviour. The key results are shown in Table 14. The 23% figure combines results likely to be in scope and extreme. However, that figure is conservative. An additional 25% of search results were ambiguous, meaning that some are plausibly interpretable by at least some people as promoting self-injurious behaviour. Thus, nearly half of all Yahoo search results do or might promote self-injurious behaviour.

Table B.9: Overall Frequencies for Yahoo Searches (NSSI, Suicide, and ED)

	Frequencies	Percentages
Prevention	1281	20%
Unrelated/Not Relevant	2032	32%
Ambiguous	1629	25%
Likely to be in Scope	1413	22%
Extreme	81	1%
Total Number of Results	6436	100.00%

For Yahoo searches, there were some differences between the results across types of searches (web, image, video, see Table B.10). Whereas 33% of image searches were coded as likely in scope or extreme, 15% of web and video searches were coded as likely in scope or extreme. Including ambiguous results, 53% of image searches, 49% of video searches, and 41% of web searches included material potentially promoting self-injurious behaviour.

Table B.10: Frequencies by Search Product for Yahoo

	Web	Image	Video
Prevention	709 (35%)	400 (14%)	172 (11%)
Unrelated/Not Relevant	496 (24%)	908 (32%)	628 (40%)
Ambiguous	523 (26%)	571 (20%)	535 (34%)
Likely to be in Scope	288 (14%)	890 (31%)	235 (15%)
Extreme	19 (1%)	62 (2%)	0 (0%)
Total Number of Results	2035	2831	1570

Left = total count; right = proportion of all results for that type of search (web, image or video).

For Yahoo searches, 25%, 12% and 34% of results were coded as likely to be in scope or extreme for NSSI, suicide, and ED, respectively. Including ambiguous results, findings were 39, 56% and 50%, respectively, for NSSI, suicide, and ED (totals are subject to rounding and may not exactly equal the sum of the values displayed; see Table B.11).

Table B.11: Results for Yahoo Searches, Separately for NSSI, Suicide, and Eating Disorder Searches: Frequencies and Percentages

	NSSI	Suicide	Eating Disorders
Prevention	307 (16%)	466 (20%)	508 (24%)
Unrelated/Not Relevant	908 (46%)	583 (25%)	541 (26%)
Ambiguous	264 (13%)	1050 (44%)	315 (15%)
Likely to be in Scope	492 (25%)	276 (12%)	645 (31%)
Extreme	8 (0.40%)	2 (0.08%)	71 (3%)
Total Number of Results	1979	2377	2080

Left = total count; right = proportion of all results for that type of self-injurious behaviour.

Promotion of Self-Injurious Behaviour Comes up Early in Yahoo Searches

The next set of analyses assessed the frequency with which promotion of self-injurious behaviour appears among the very first result of a search. Across all searches, 21% of the first results were coded as either extreme or likely in scope and another 30% were coded as ambiguous. Primarily self-injurious content on Yahoo occurred in image searches (28% coded as likely in scope and another 24% ambiguous). Web searches produced 20% first results that were coded as either likely in scope or extreme and another 20% that were ambiguous. For video searches, 14% of the first results were coded as extreme or likely in scope, but 46% were ambiguous. Overall, therefore, about 50-60% of first results possibly included content that was, or was plausibly interpretable as, promoting self-injurious behaviour.

With respect to results among the first five, a similar pattern emerged. Overall, 20% of these results among the first five were coded as likely in scope and 1% were coded as extreme. Another 29% of these searches were coded as ambiguous, for a grand total of 50% that were or might be promoting self-injurious behaviour. Among the first five results, 23% of image search results were coded as likely in scope or extreme and another 24% ambiguous). For web searches, among the first five results, 22% were coded as likely in scope, 2% extreme, and another 25% ambiguous. 15% of video searches, among the first five results, were coded as likely in scope, 1% extreme, and another 38% were coded as ambiguous.

AOL Results

Across all searches, NCRI found that 19% of AOL results were coded as leading to sources that promote self-injurious behaviour. The key results are shown in Table B.12. The 19% figure combines results coded as likely to be in scope and extreme. However, that figure is conservative. An additional 22% of search results were coded as ambiguous, meaning that some are plausibly interpretable by at least some people as promoting self-injurious behaviour. Thus, over 2 in 5 of all AOL search results generated by NCRI do or might promote self-injurious behaviour.

Table B.12: Overall Frequencies for AOL Searches (NSSI, Suicide, and ED)

	Frequencies	Percentages
Prevention	1467	20%
Unrelated/Not Relevant	2946	40%
Ambiguous	1605	22%
Likely to be in Scope	1339	18%
Extreme	51	1%
Total Number of Results	7408	100%

For AOL searches, 28% of image searches were coded as likely in scope or extreme, and 12% of web and video searches were coded as likely to be in scope or extreme (see Table B.13). Including ambiguous results, 47% of image searches, 39% of video searches, and 33% of web searches included material potentially promoting self-injurious behaviour (totals are subject to rounding and may not exactly equal the sum of the values displayed).

Table B.13: Frequencies by Search Productfor AOL

	Web	Image	Video
Prevention	838 (40%)	257 (9%)	372 (16%)
Unrelated/Not Relevant	569 (27%)	1299 (44%)	1078 (45%)
Ambiguous	421 (20%)	560 (19%)	624 (26%)
Likely to be in Scope	255 (12%)	796 (27%)	288 (12%)
Extreme	15 (1%)	27 (1%)	9 (0.38%)
Total Number of Results	2098	2939	2371

Left = total count; right = proportion of all results for that type of search (web, image or video).

For AOL searches, 15%, 12% and 29% of results were coded as likely to be in scope or extreme for NSSI, suicide, and ED, respectively. Including ambiguous results, findings were 28, 44% and 48%, respectively, for NSSI, suicide, and ED (totals are subject to rounding and may not exactly equal the sum of the values displayed; see Table B.14).

Table B.14: Results for AOL Searches, Separately for NSSI, Suicide, and Eating Disorder Searches

	NSSI	Suicide	Eating Disorders
Prevention	357 (15%)	681 (26%)	429 (17%)
Unrelated/Not Relevant	1309 (57%)	758 (29%)	879 (35%)
Ambiguous	292 (13%)	835 (32%)	478 (19%)
Likely to be in Scope	348 (15%)	305 (12%)	686 (27%)
Extreme	3 (0.13%)	9 (0.35%)	39 (2%)
Total Number of Results	2309	2588	2511

Left = total count; right = proportion of all results for that type of self-injurious behaviour.

Promotion of Self-Injurious Behaviour Comes up Early in AOL Searches

The next set of analyses assessed the frequency with which promotion of self-injurious behaviour appears among the very first result of a search. Across all searches, 19% of the first results were coded as either extreme or likely in scope and another 26% were coded as ambiguous. The most promotion of self-injurious content in AOL occurred for image searches (30% likely in scope and another 17% ambiguous). Web searches produced 16% first results that were coded as either likely in scope or extreme and another 27% that were ambiguous. For video searches, 19% of the first results were coded as extreme or likely in scope and 26% were ambiguous. Overall, therefore, about 40-50% of first results possibly included content that was, or was plausibly interpretable as, promoting self-injurious behaviour.

With respect to results among the first five, a similar pattern emerged. Overall, 21% of these results among the first five were coded as likely in scope or extreme. Another 25% of these searches were coded as ambiguous, for a grand total of 46% that were or might be promoting self-injurious behaviour.

Among the first five results, 32% of image search results were coded as likely in scope and another 18% ambiguous. For web searches, among the first five results, 18% were coded as likely in scope, 1% as extreme, and another 27% as ambiguous. 11% of video searches, among the first five results, were coded as likely in scope and another 30% as ambiguous.

CITATIONS

1. Goldenberg, A., Farmer, J., Jussim, L., Sutton, L., Finkelstein, D., Ramos, C., Paresky, P., & Finkelstein, J. (2022). Online communities of adolescents and young adults celebrating, glorifying, and encouraging self-Harm and suicide are growing rapidly on Twitter. Network Contagion Research Institute.
2. Jarvi, S., Jackson, B., Swenson, L., & Crawford, H. (2013). The impact of social contagion on non-suicidal self-injury: A review of the literature. Archives of Suicide Research, 17(1), 1-19.
3. Boero, N., & Pascoe, C. J. (2012). Pro-anorexia communities and online interaction: Bringing the pro-ana body online. Body & Society, 18(2), 27-57.
4. Lewis, S. P., & Seko, Y. (2016). A double-edged sword: A review of benefits and risks of online nonsuicidal self-injury activities. Journal of Clinical Psychology, 72(3), 249-262.
5. Bell, J. (2014). Harmful or helpful? The role of the internet in self-harming and suicidal behaviour in young people. Mental Health Review Journal, 19(1), 61-71.
6. Boyd, D., Ryan, J., & Leavitt, A. (2011). Pro-self-harm and the visibility of youth-generated problematic content. ISJLP, 7, 1.
7. Selby, E. A., Bender, T. W., Gordon, K. H., Nock, M. K., & Joiner Jr, T. E. (2012). Non-suicidal self-injury (NSSI) disorder: A preliminary study. Personality Disorders: Theory, Research, and Treatment, 3(2), 167.
8. Polivy, J., & Herman, C. P. (2002). Causes of eating disorders. Annual Review of Psychology, 53(1), 187-213.
9. Macnamara, J. R. (2005). Media content analysis: Its uses, benefits and best practice methodology. Asia Pacific Public Relations Journal, 6(1), 1-34.
10. Barros, J.M., et al. (2019). The validity of Google trends search volumes for behavioural forecasting of national suicide rates in Ireland. International Journal of Environmental Research and Public Health, 16(17), 3201.
11. Arora, V. S., Stuckler, D., & Mckee, M. (2016). Tracking search engine queries for suicide in the United Kingdom, 2004–2013. Public Health, 137, 147-153.
12. Luh, C., Yang, S., & Huang, T.D. (2016). Estimating Google's search engine ranking function from a search engine optimization perspective. Online Information Review, 40, 239-255.
13. Ham, C. D. (2019). Why is this first? Understanding and analyzing internet search results. Journal of Educational Research and Practice, 9(1), 400-412.
14. Beus, J. (2020). Why (almost) everything you knew about Google CTR is no longer valid. Sistrix.
15. Yang, G., Yang, S., Luo, K., Lan, S., He, L., & Li, Y. (2023). Detection of non-suicidal self-injury based on spatiotemporal features of indoor activities. IET Biometrics, 12(2), 91-101.
16. Jacob, N., Evans, R., Scourfield, J. (2017). The influence of online images on self-harm: A qualitative study of young people aged 16-24. Journal of Adolescence, 60, 140-147.
17. Conti, M., Pajola, L., & Tricomi, P. P. (2023). Turning captchas against humanity: Captcha-based attacks in online social media. Online Social Networks and Media, 36, 100252.
18. Farid, H. (2021). An overview of perceptual hashing. Journal of Online Trust and Safety, 1(1).
19. Chancellor, S., Pater, J. A., Clear, T., Gilbert, E., & De Choudhury, M. (2016, February). # thyghgapp: Instagram content moderation and lexical variation in pro-eating disorder communities. Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing (pp. 1201-1213).
20. Twitter.com. (March 2023).
21. YouTube.com. (March 2023).
22. YouTube.com. (March 2023).
23. SoundCloud.com. (March 2023).
24. Samaritans.org. (March 2023).
25. Perkins, S., Cryst, E., & Grossman, S. (2021). Self-Harm Policies and Internet Platforms. Stanford Internet Observatory Cyber Policy Center.
26. Google. (2023) Privacy and Terms.
27. Microsoft. (2022). Microsoft Services Agreement.
28. DuckDuckGo. (2023) DuckDuckGo Terms of Service.
29. Yahoo. (2021). Yahoo Community Guidelines.
30. AOL. (2021). Guidelines for Comments on AOL.