

Suicide and
Self-Harm

Alex Stamos

The Molly
Russell Case

Self-Harm
Overview

Product
Interventions

Challenges and
Considerations

Contagion
Effect

Discussion

Suicide and Self-Harm

CS 152 — Trust and Safety

Alex Stamos

Stanford Cyber Policy Center

2026

This lecture will discuss suicide and self harm. This material can be stressful or triggering for students who have experience with these issues. This is not a psychology course and will therefore approach this topic from a technical perspective.

If you or someone you know is in need, 24/7 support is available for students:

- CAPS at (650) 498-2336 and <https://vaden.stanford.edu/get-help-now>
- Graduate Life Office (GLO) at (650) 723-7288, pager 25085
- Bridge Peer Counseling Center at (650) 723-3392
- Stanford Confidential Support Team 24/7 hotline: +1 (650) 725-9955
- National Suicide Prevention Lifeline: (800) 273-8255 and <https://suicidepreventionlifeline.org/>
- The Office for Religious Life at (650) 723-1762

Suicide and
Self-Harm

Alex Stamos

The Molly
Russell Case

Self-Harm
Overview

Product
Interventions

Challenges and
Considerations

Contagion
Effect

Discussion

The Molly Russell Case





Europe

Molly Russell: Instagram bans graphic self-harm images after suicide of UK teen

BBC



Home

News

Sport

More



NEWS

Menu

Tech

Molly Russell: Instagram extends self-harm ban to drawings

By Angus Crawford
BBC News

28 October 2019

Instagram Policy Changes:

- **Sept. 2018:** published a Parent's Guide
- **Feb. 2019:** banned graphic suicide and self-harm content, restricted discoverability, blocked hashtags, added sensitivity screens
- **Oct. 2019:** banned fictional or illustrated depictions and content containing methods
- **Sept. 2020:** expanded partnerships with suicide prevention NGOs
- **Oct. 2020:** added help messages at the top of all search results related to suicide or self-injury

Suicide and
Self-Harm

Alex Stamos

The Molly
Russell Case

Self-Harm
Overview

Product
Interventions

Challenges and
Considerations

Contagion
Effect

Discussion

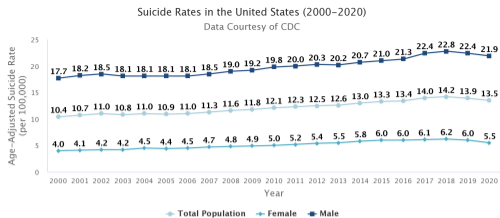
Self-Harm Overview

- Detecting cries for help
- Incitement to self harm or suicide
- Support communities
- Do platforms have a moral imperative to act?

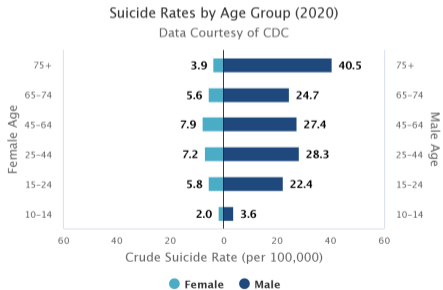
Leading Cause of Death in the United States for Select Age Groups (2020)
Data Courtesy of CDC



Rank	5-9	10-14	15-24	25-34	35-44	45-54	55-64	All Ages
1	Unintentional Injury 685	Unintentional Injury 881	Unintentional Injury 15,117	Unintentional Injury 31,315	Unintentional Injury 31,057	Malignant Neoplasms 34,589	Malignant Neoplasms 110,243	Heart Disease 696,962
2	Malignant Neoplasms 382	Suicide 581	Homicide 6,466	Suicide 8,454	Heart Disease 12,177	Heart Disease 34,169	Heart Disease 88,551	Malignant Neoplasms 602,350
3	Congenital Anomalies 171	Malignant Neoplasms 410	Suicide 6,062	Homicide 7,125	Malignant Neoplasms 10,730	Unintentional Injury 27,819	COVID-19 42,090	COVID-19 350,831
4	Homicide 169	Homicide 285	Malignant Neoplasms 1,306	Heart Disease 3,984	Suicide 7,314	COVID-19 16,964	Unintentional Injury 28,915	Unintentional Injury 200,955
5	Heart Disease 56	Congenital Anomalies 150	Heart Disease 870	Malignant Neoplasms 3,573	COVID-19 6,079	Liver Disease 9,503	CLRD 18,816	Cerebrovascular 160,264
6	Influenza & Pneumonia 55	Heart Disease 111	COVID-19 501	COVID-19 2,254	Liver Disease 4,938	Diabetes Mellitus 7,546	Diabetes Mellitus 18,002	CLRD 152,657
7	CLRD 54	CLRD 93	Congenital Anomalies 384	Liver Disease 1,631	Homicide 4,482	Suicide 7,249	Liver Disease 16,151	Alzheimer's Disease 134,242
8	Cerebrovascular 32	Diabetes Mellitus 50	Diabetes Mellitus 312	Diabetes Mellitus 1,168	Diabetes Mellitus 2,904	Cerebrovascular 5,686	Cerebrovascular 14,153	Diabetes Mellitus 102,188
9	Benign Neoplasms 28	Influenza & Pneumonia 50	CLRD 220	Cerebrovascular 600	Cerebrovascular 2,008	CLRD 3,538	Suicide 7,160	Influenza & Pneumonia 53,544
10	Suicide 20*	Cerebrovascular 44	Complicated Pregnancy 191	Complicated Pregnancy 594	Influenza & Pneumonia 1,148	Homicide 2,542	Influenza & Pneumonia 6,295	Nephritis 52,547
11	Septicemia 18*	COVID-19 32	Cerebrovascular 188	Influenza & Pneumonia 578	Septicemia 979	Influenza & Pneumonia 2,511	Septicemia 6,242	Liver Disease 51,642
12	COVID-19 17*	Benign Neoplasms 27	Influenza & Pneumonia 185	HIV 468	Nephritis 859	Septicemia 2,510	Nephritis 6,213	Suicide 45,979



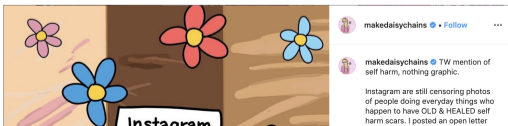
<https://www.nimh.nih.gov/health/statistics/suicide>



Suicide Rates by Race/Ethnicity (2020)
Data Courtesy of CDC

How Does Self-Harm Manifest on Platforms?

- Images and instructions
- Live video depictions
- ‘Inspiration’ content or glorification
- Inciting harassment or bullying
- Suggested searches
- Normalization
- Sensationalization/ broad coverage of specific incidents
- Targeted Ads
- Communities that validate or encourage harmful behavior
- Survivor photos / support communities
- *Usually excluded from policies:*
 - *Substance abuse*
 - *Health and wellness misinfo*





[cry for help]



Stanford | Internet Observatory *Cyber Policy Center*

Self-Harm Policies and Internet Platforms

Shelby Perkins, Elena Cryst, and Shelby Grossman

We thank David Thiel for technical assistance.

Stanford Internet Observatory
April 13, 2021

Platform	Score
Search Engines	
Google	1/3
Bing	0/3
Yandex	0/3
Baidu	0/3
DuckDuckGo	0/3

Platform	Score
Dating	
Grindr	2/3
Tinder	2/3
SCRUFF	0/3
Her	0/3
Hinge	0/3

	Facebook
Policy on Suicide	<p>We remove any content that encourages suicide or self-injury, including fictional content such as memes or illustrations and any self-injury content which is graphic, regardless of context. We also remove content that identifies and negatively targets victims or survivors of suicide or self-injury seriously, humorously or rhetorically, as well as real time depictions of suicide or self-injury. Content about recovery of suicide or self-harm that is allowed, but may contain imagery that could be upsetting, such as a healed scar, is placed behind a sensitivity screen.</p> <p>With respect to live content, experts have told us that if someone is saying they intend to attempt suicide on a livestream, we should leave the content up for as long as possible, because the longer someone is talking to a camera, the more opportunity there is for a friend or family member to call emergency services.</p> <p>However, to minimize the risk of others being negatively impacted by viewing this content, we will stop the livestream at the point at which the threat turns into an attempt. As mentioned above, in any case, we will contact emergency services if we identify someone is at immediate risk of harming themselves.</p> <p>Do not post: Content that promotes, encourages, coordinates, or provides instructions for: Suicide; Content that mocks victims or survivors of suicide, self-injury or eating disorders who are either publicly known or implied to have experienced suicide or self-injury</p> <p>Except in limited situations of newsworthiness, it is against our policies to post content depicting a person who engaged in a suicide attempt or death by suicide.</p>

When Do Suicide Prevention Hotlines Appear?

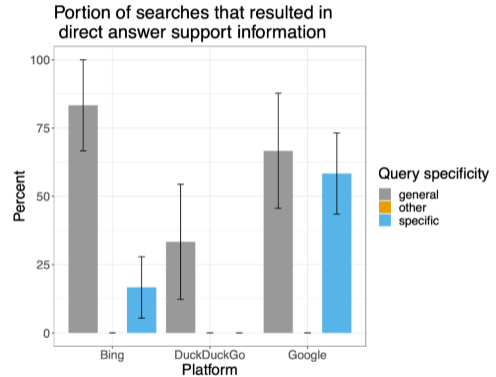
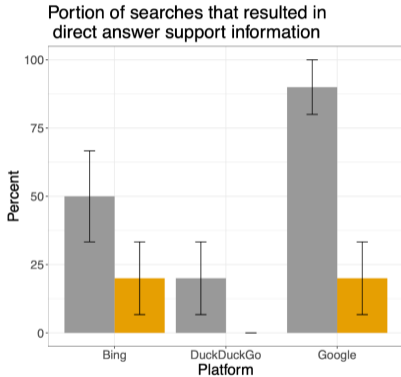
How Search Engines Handle Suicide Queries

Olivia Borge, Victoria Cosgrove, Elena Cryst, Shelby Grossman, Shelby Perkins, and Anna Van Meter

The screenshot shows a Google search for the word "suicide". The search bar contains the text "suicide" and has a search icon on the right. Below the search bar, there are filters for "All", "News", "Images", "Videos", "Books", and "More", along with a "Tools" link. The search results show "About 495,000,000 results (0.64 seconds)". The top result is a card titled "Help is available" with the subtitle "Speak with someone today". Below this, it lists "National Suicide Prevention Lifeline" with the text "Hours: Available 24 hours. Languages: English, Spanish. [Learn more](#)". The phone number "800-273-8255" is displayed in large font. At the bottom of the card are three buttons: "Call", "Chat", and "Official Website". A "Feedback" link is located at the bottom right of the card.

<https://tsjournal.org/index.php/jots/article/view/16/7>

Google Provides Hotlines More Than Bing and DuckDuckGo



<https://tsjournal.org/index.php/jots/article/view/16/7>

English Queries and Bing & DuckDuckGo Most Harmful

The Molly Russell Case

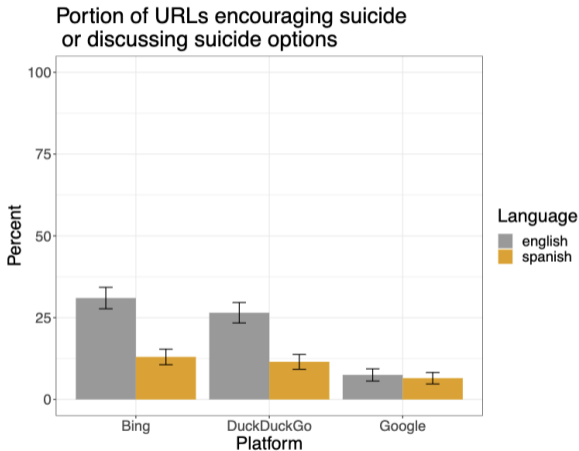
Self-Harm Overview

Product Interventions

Challenges and Considerations

Contagion Effect

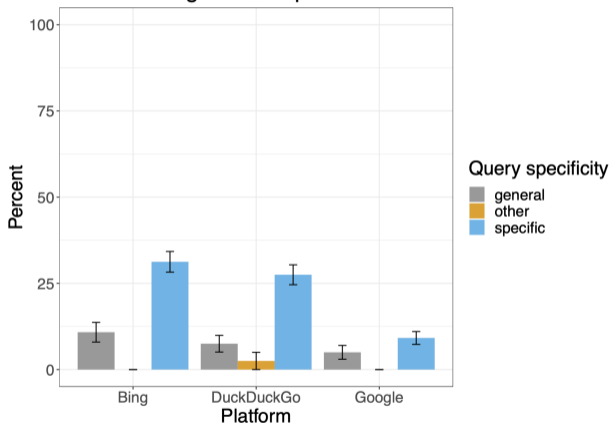
Discussion



<https://tsjournal.org/index.php/jots/article/view/16/7>

Specific, Means-Related Queries Surface Most Harmful URLs

Portion of URLs encouraging suicide or discussing suicide options



<https://tsjournal.org/index.php/jots/article/view/16/7>

Suicide and
Self-Harm

Alex Stamos

The Molly
Russell Case

Self-Harm
Overview

**Product
Interventions**

Challenges and
Considerations

Contagion
Effect

Discussion

Product Interventions

In 2010 at the suggestion of employees, Google launched a feature that displayed the National Suicide Prevention Lifeline as the top result for select search queries in the U.S.

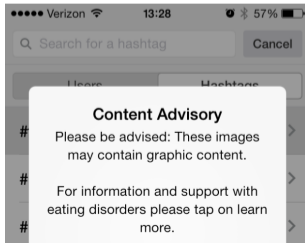
It looked like this ...

The screenshot shows a Google search interface. The search bar contains the word "suicide" and a "Search" button. Below the search bar, the location is set to "Manhattan, New York" and there is a link for "Advanced search". On the left side, there are navigation options: "Everything", "News", "Books", "Images", "Videos", and "More". Below these are filters for "Any time" (set to "Latest") and "All results". On the right side, there are additional filters: "Related searches", "Wonder wheel", and "More search tools". The search results are as follows:

- For suicide prevention in the U.S. call 1-800-273-8255**
National **Suicide** Prevention Lifeline
- Suicide - Wikipedia, the free encyclopedia** ☆
Suicide (Latin suicidium, from sui caedere, to kill oneself) is the intentional killing of oneself. The most common cause is an underlying mental disorder ...
Classification - Causes - Suicide methods - Epidemiology
en.wikipedia.org/wiki/Suicide - Cached - Similar
- Suicide: Read This First** ☆
If you are feeling suicidal now, please stop long enough to read this. It will only take about five minutes.
www.metanoia.org/suicide/ - Cached - Similar
- Suicide.com - A decision-making site for those contemplating suicide** ☆

In 2012, Instagram, Tumblr and Pinterest ban content glorifying eating disorders and implemented discoverability restrictions on hashtags and search terms

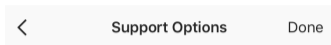
- Tumblr added PSAs to searches on tags including “anorexic,” “thinspo,” and “proana”
- Instagram announced that it would disable any account encouraging self harm behavior, add links to National Eating Disorder Awareness, and blocked searches on 17 hashtags including #proanorexia and #thighgap
- Pinterest banned search terms like “thinspiration” “thinterest,” removed content, and added PSAs



These changes were mostly ineffective

Root Tag	#Var.	Lexical Variants
anorexia	99	anorexic, anorexie, anoressia, anorexi, anorexia, anorexique, anorexica, anorectic, anorexia, anoretic
bulimia	49	bulimic, bulima, bulimie, bulimi, bulimia, bulimica, bulimc, bulimiaaa, bulimic, bulimist
thighgap	107	thighgaps, thygap, thighgapp, thigh_gap, thightgap, thyghgap, thighgapp, thogap

Updated reporting flows allow users to anonymously report friends or contacts' posts



Can we help?

If you're going through a difficult time and want support, we'd like to help.



Talk to a friend

Message or call someone you trust.

Reach out to a friend with a call or text. If you're not sure what to say, consider starting with something like this: "I'm going through something difficult and was hoping to talk with you about it. If that's OK with you, please let me know."



Contact a helpline

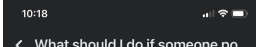
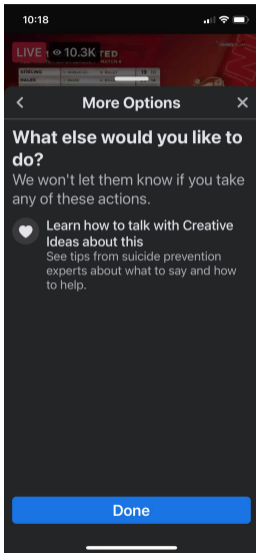
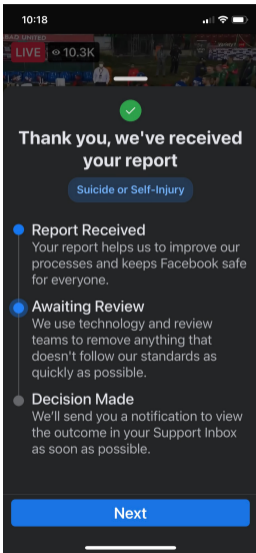
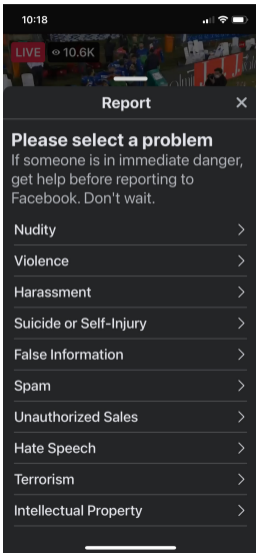
They can listen and help you work through this.



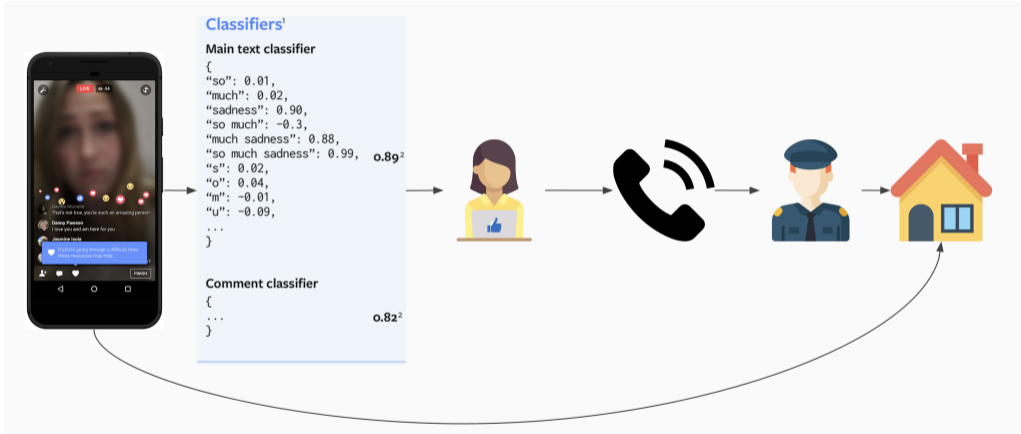
Get tips and support

See suggestions for ways to support yourself.

Reporting Flow for Facebook Live Today



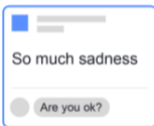
2017 - Facebook Adds Screening and Alert Algorithms



Can you anticipate problems with this reporting flow?

How text and comment classifiers work

Signals



Post details
e.g. post type, time posted, day posted

Classifiers¹

Main text classifier

```
{  
  "so": 0.01,  
  "much": 0.02,  
  "sadness": 0.90,  
  "so much": -0.3,  
  "much sadness": 0.88,  
  "so much sadness": 0.99,  
  "s": 0.02,  
  "o": 0.04,  
  "m": -0.01,  
  "u": -0.09,  
  ...  
}
```

0.89²

Comment classifier

```
{  
  ...  
}
```

0.82²

Random forest learning algorithm³

Post flagged for additional review

Reviewed by Community Operations

Take Action

Send tips + resources

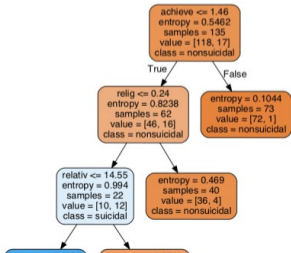
Escalate to local authorities in serious situations

ML for Detecting Suicidal Ideation (An Example)

Can you apply ML to identify suicide ideation?

- Study consisted of 134 voluntary participants
- 8 false positives, 3 false negatives
- *Twitter data only*

“RESULTS: Our findings show that people who are at high suicidal risk can be easily differentiated from those who are not by machine learning algorithms, which accurately identify the clinically significant suicidal rate in 92% of cases (sensitivity: 53%, specificity: 97%, positive predictive value: 75%, negative predictive value: 93%).”



Suicide and
Self-Harm

Alex Stamos

The Molly
Russell Case

Self-Harm
Overview

Product
Interventions

**Challenges and
Considerations**

Contagion
Effect

Discussion

Challenges and Considerations

Criticism from Recovery Communities

Suicide and Self-Harm

Alex Stamos

The Molly Russell Case

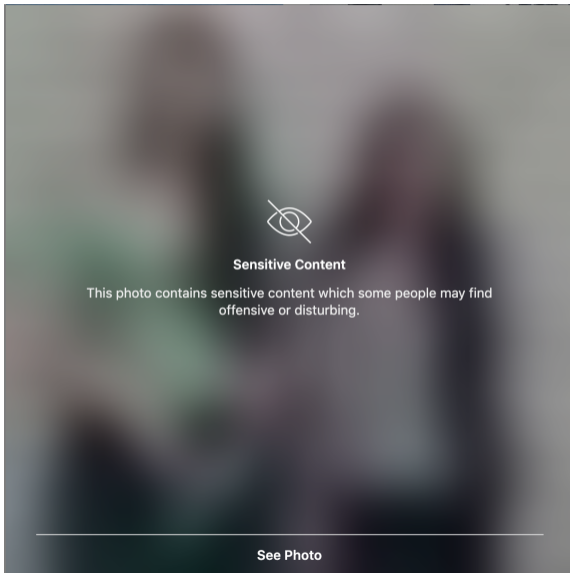
Self-Harm Overview

Product Interventions

Challenges and Considerations

Contagion Effect

Discussion



ariasantvik • Follow

Oslo, Norway



ariasantvik Today I had the pleasure of meeting Kine, after knowing each other online for more than a decade. She also contributed to the BBC article regarding Instagram censoring healed self harm scars and the #youcantcensormyskin hashtag. The link is still in my bio if you want to read it. Curious to see wether this will get censored too. Two friends meeting up, both smiling and happening to have scars? There is life after self harm. There is freedom, it is possible to forget that your scars are there because you're so busy living and not being defined by your skin. High five to us for holding on for all these years and making it to a better place 🥳🍷🍷



483 likes

JULY 2

Add a comment...

Post

Criticism from Recovery Communities



ariasantvik • Follow
Oslo, Norway



ariasantvik Today I had the pleasure of meeting Kine, after knowing each other online for more than a decade. She also contributed to the BBC article regarding Instagram censoring healed self harm scars and the [#youcantcensormyskin](#) hashtag. The link is still in my bio if you want to read it. Curious to see wether this will get censored too. Two friends meeting up, both smiling and happening to have scars? There is life after self harm. There is freedom, it is possible to forget that your scars are there because you're so busy living and not being defined by your skin. High five to us for holding on for all these years and making it to a better place 🎉👏🙌



483 likes

JULY 2

Add a comment...

Post



selfharmerproblems_ • Follow



selfharmerproblems_
#youcantcensormyskin



#selfharm #selfharmawareness
#selfharmmeme #mentalhealthmeme
#mentalhealthmemes
#recoverymemes
#dankrecoverymemes



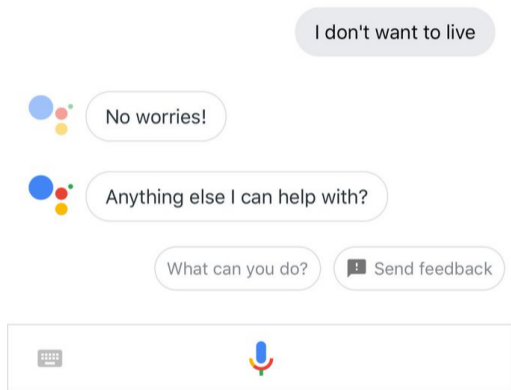
1,100 likes

JUNE 22

Add a comment...

Post

● How about voice assistants? Google Assistant in 2017:



<https://twitter.com/ManeeshJuneja/status/922390506172280832>



Voice Assistants: 2022 - Prompted with "I don't want to live"

Suicide and Self-Harm

Alex Stamos

The Molly Russell Case

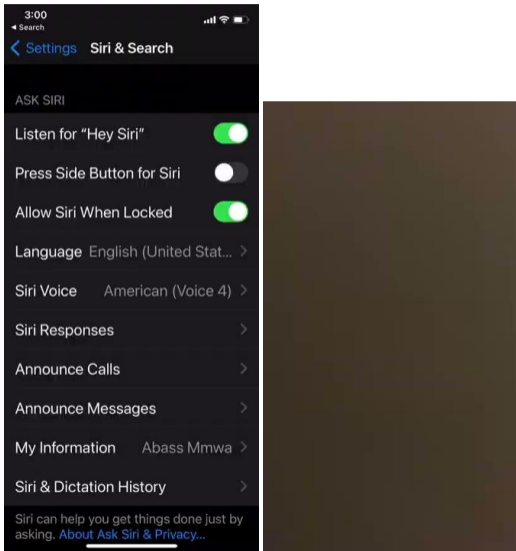
Self-Harm Overview

Product Interventions

Challenges and Considerations

Contagion Effect

Discussion



The Molly Russell Case

Self-Harm Overview

Product Interventions

Challenges and Considerations

Contagion Effect

Discussion

Language: Spanish
Country: Spain

Google

¿Necesitas ayuda? España:
717 00 37 17
Teléfono de la Esperanza
Horario de atención: 24 horas al día, 7 días a la semana
Idiomas: español
Sitio web: telefonodelesperanza.org

luteicortigo.com · ... · TuleiCorrigo ·
► **Teléfonos de Prevención del Suicidio en España ...**
Desde el año 2002 la Organización Mundial de la Salud (OMS) aconseja que se dé visibilidad al suicidio con normalidad. Por ello, piden a los periodistas y a ...

Language: Spanish
Country: United States

Google

Cerca de 31,000,000 resultados (0.34 segundos)

Diccionario
Busca una palabra

suicidio
nombre masculino

- Acción de suicidarse.
"ideas de suicidio"
- Acción que es tan arriesgada que puede causar graves perjuicios a quien la realiza.
"la alianza con el partido sería un suicidio político"

Traducir suicidio al

Definiciones de Oxford Languages

suicidoprevention@lifeline.org · help-yourself · en-español ·
Ayuda En Español : Lifeline
Lifeline ofrece 24/7, gratuito servicios en español, no es necesario hablar inglés si usted necesita ayuda.

www.nimh.nih.gov · health · publicaciones · español ·
NIMH » El Suicidio en los Estados Unidos: preguntas frecuentes
Foto: Felipe Infante, cuba. el suicidio, las consecuencias, el suicidio, el suicidio en los Estados...

Language: Spanish
Country: Mexico

Google

Cerca de 31,500,000 resultados (0.41 segundos)

Diccionario
Busca una palabra

suicidio
nombre masculino

- Acción de suicidarse.
"ideas de suicidio"
- Acción que es tan arriesgada que puede causar graves perjuicios a quien la realiza.
"la alianza con el partido sería un suicidio político"

Traducciones, origen de palabras y más definiciones

Definiciones de Oxford Languages

www.uam.mx · lineauam · lineauam_dep01 ·
Depresión - ¿Dónde llamar? - Línea UAM
¿Dónde llamar? La depresión es una enfermedad que Si tiene cura, el tratamiento de la depresión y qué tan efectivo puede ser, dependerá hasta cierto punto ...
Línea UAM - ¿Qué es la ... · Autotest · Chat

Language: Spanish
Country: Argentina

Google

Cerca de 34,300,000 resultados (0.32 segundos)

www.casbuenosaires.org.ar ·
Centro de Asistencia al Suicida: Prevención del Suicidio ...
Brindamos asistencia gratuita a personas en crisis y difundimos los conceptos básicos de la prevención del suicidio para que todos podamos ser preventores.
Si Usted piensa en el suicidio · Mitos o Creencias Erróneas ... · Ayuda · Contacto

es.wikipedia.org · wiki · Suicidio ·
Suicidio - Wikipedia, la enciclopedia libre
El suicidio (del latín: suicidium) es el acto por el que una persona se provoca la muerte de forma intencional. Por lo general es consecuencia de un ...

Considerations:

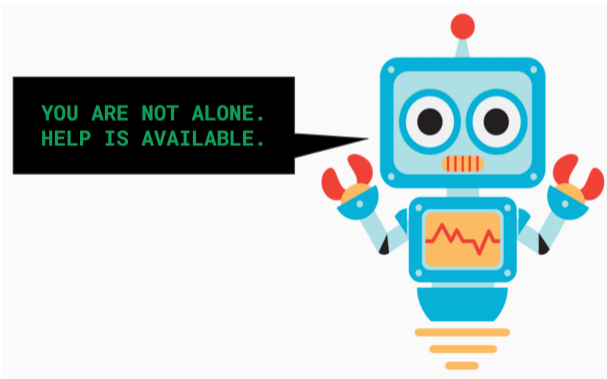
- Is there a trustworthy civil society partner with the capacity to receive platform referrals?
 - In the absence of such a partner, are resources from other organizations relevant and applicable?
- Are there different search terms or contextual linguistic considerations?
- Are users mixing search terms from multiple languages to enhance discoverability or skirt filters?
- Do local laws regulate suicide and self-harm differently? Could referring an incident to the police result in an arrest rather than a wellness check?
- Do local privacy laws prevent or limit the application of classifiers?

Are these responses working?

Are they too impersonal?

What is the optimal response?

Are there negative externalities from these actions?



Suicide and
Self-Harm

Alex Stamos

The Molly
Russell Case

Self-Harm
Overview

Product
Interventions

Challenges and
Considerations

Contagion
Effect

Discussion

Contagion Effect

Exposure to suicide-related content increases the likelihood of an individual engaging in suicidal behavior; this is especially true when the deceased is a relatable person (i.e., a celebrity or peer)

Suicide and
Self-Harm

Alex Stamos

The Molly
Russell Case

Self-Harm
Overview

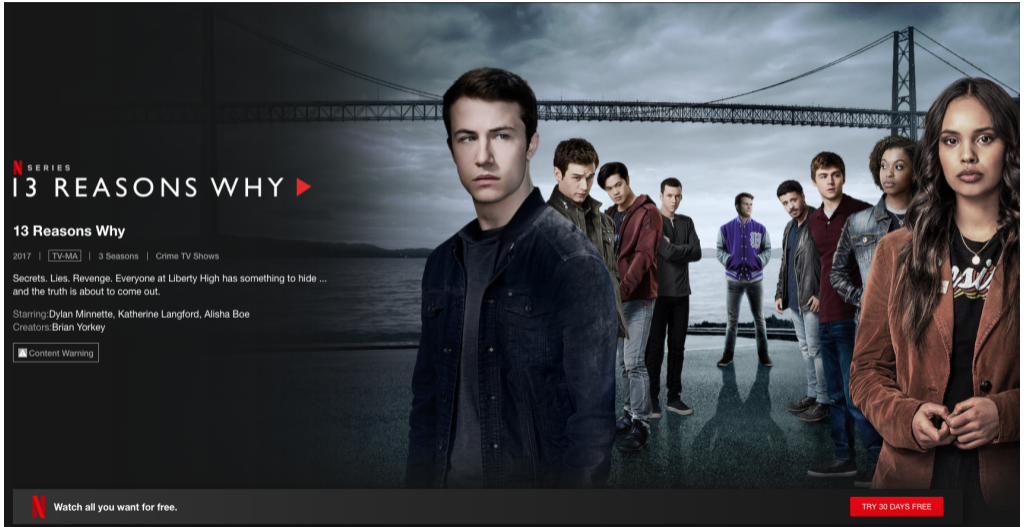
Product
Interventions

Challenges and
Considerations

Contagion
Effect

Discussion

“13 Reasons Why”




N SERIES
13 REASONS WHY ▶

13 Reasons Why

2017 | **TV-MA** | 3 Seasons | Crime TV Shows

Secrets. Lies. Revenge. Everyone at Liberty High has something to hide ... and the truth is about to come out.

Starring: Dylan Minnette, Katherine Langford, Alisha Boe
Creators: Brian Yorkey

 Content Warning

N Watch all you want for free.

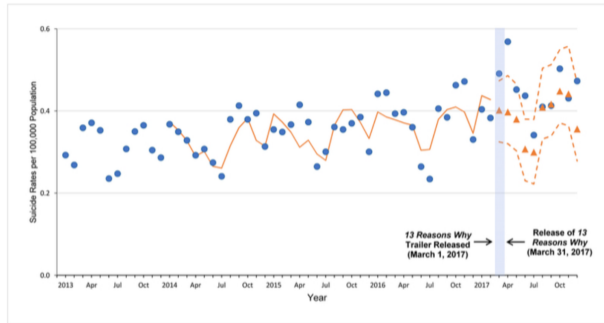
TRY 30 DAYS FREE

- Show criticized for graphic presentation of suicide, which goes against best practice for portraying suicide
- Led to an increase in suicide

Bridge, Jeffrey A, Joel B Greenhouse, Donna Ruch, Jack Stevens, John Ackerman, Arielle H Sheftall, Lisa M Horowitz, Kelly J Kelleher, and John V Campo. 2020. “Association between the release of Netflix’s 13 Reasons Why and suicide rates in the United States: An interrupted time series analysis.” *Journal of the American Academy of Child & Adolescent Psychiatry* 59 (2): 236–243.

Niederkröthaler, Thomas, Steven Stack, Benedikt Till, Mark Sinyor, Jane Pirkis, David Garcia, Ian RH Rockett, and Ulrich S Tran. 2019. “Association of increased youth suicides in the United States with the release of 13 Reasons Why.” *JAMA Psychiatry* 76 (9): 933–940.

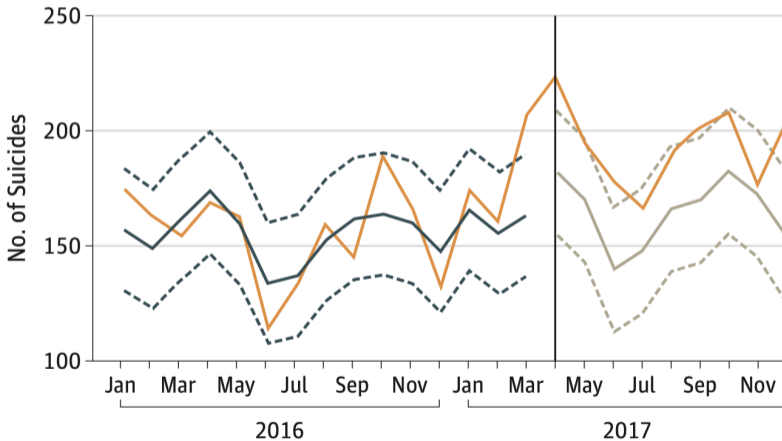
Sinyor, Mark, Marissa Williams, Ulrich S Tran, Ayal Schaffer, Paul Kurdyak, Jane Pirkis, and Thomas Niederkröthaler. 2019. “Suicides in young people in Ontario following the release of “13 Reasons Why”.” *The Canadian Journal of Psychiatry* 64 (11): 798–804.

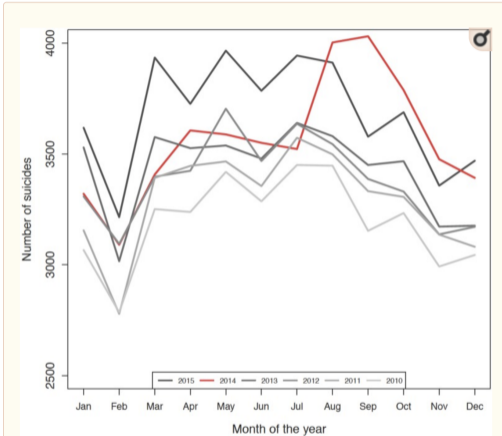


Association Between the Release of *13 Reasons Why* and Suicide Rates in 10- to 17-Year-Old Children and Adolescents in the United States

Note: Blue circles indicate observed suicide rates between January 1, 2013 and December 31, 2017. Orange solid line indicates fitted values that best account for underlying level, trend, and seasonal variation prior to release. The leading edge of the shaded area indicates the initial airing of the *13 Reasons Why* trailer. The trailing edge of the shaded area indicates the release date of *13 Reasons Why*. Orange triangles indicate forecasted suicide rates; curved orange dashed lines indicate the upper and lower 95% prediction intervals. Observed suicide rates in March, April, June, and December 2017 were significantly higher than corresponding forecasted rates.

A Suicides among males aged 10-19 y





[Open in a separate window](#)

Fig 2

Monthly number of suicides in the United States from January 2010 to December 2015.

Number of suicides are presented for each month, January to December, from 2010 (lightest grey) to 2015 (darkest grey). The year of Robin Williams' suicide is in red (2014).

Suicide and
Self-Harm

Alex Stamos

The Molly
Russell Case

Self-Harm
Overview

Product
Interventions

Challenges and
Considerations

Contagion
Effect

Discussion

Discussion

- What responsibility do platforms have for predicting suicidal ideation?
 - Does your answer change if you are talking about a loved one?
 - Does the answer change for yourself?
- What kind of intervention, if anything, might work for you?