
WARNING

The report you are about to view contains explicit examples of discriminatory, violent and threatening content that has been directed at student-athletes, coaches and officials from abusive online users.

The NCAA condemns this behavior and is committed to driving change about this issue.



ONLINE ABUSE IN NCAA CHAMPIONSHIPS
NCAA Pilot Study 2023-24



AI MONITORING: PROTECTING ATHLETES, COACHES & OFFICIALS NCAA PILOT STUDY

CONTENTS

Introduction	4
Project Scope	5
SUMMARY FINDINGS	
Summary Analysis	7
Executive Summary Breakdown	8
Key Figures: All Championships	9
Comparative Analysis	10
Comparative Statistics	11
ISSUE INSIGHTS + EXAMPLES	
Sports Betting & Match Fixing	13
Sexual Content, Racism, & Violence	14
Homophobia & Transphobia, Doping, & Match Officials	15
MENTAL HEALTH	
Mental Health Toll on Victims	17
APPENDIX A: GLOSSARY & DEFINITIONS	
Gravity of Abuse Definitions	19
Terminology & Categories	20



AI MONITORING: PROTECTING ATHLETES, COACHES & OFFICIALS NCAA PILOT STUDY

INTRODUCTION

The NCAA partnered with Signify Group and its artificial intelligence (AI) Threat Matrix service to identify, analyse, and investigate online abuse and threats received by athletes, coaches and officials in a number of championships across 2023 and 2024. This report provides a summary analysis of the pilot study.

Threat Matrix was in operation across open-source data with a focus on X (formerly Twitter), Instagram, Facebook, and Tik Tok. It covered the following criteria:

- Identification of targeted online abuse and threat
- Levels and categorisation of abuse and threat, including gravity of offences, sources and identification of tactics
- Immediate notification to NCAA of any identified (fixated) threat or serious harassment
- Expediting abusive or threatening content take-down/sanctioning of account holders - where platform community guidelines are evidenced to have been broken
- Unmasking of abusers to enable action by NCAA and/or prosecution by law enforcement

The Threat Matrix service covers 39 languages and monitors for abusive and threatening posts that are identified based on text, with layered word, image, emoji and phrase categorisation / detection.

Threat Matrix operates a specialist AI-powered threat detection algorithm that also covered athletes and officials under the umbrella of this project.

Proactive monitoring of athletes' and officials' private direct messages (DMs) was not part of the project. Signify has a proprietary DM service but onboarding is required by each individual due to the involvement of private data. Therefore, the DM service is implemented on an intelligence-led basis – and highlighted to the athletes most at risk.

Signify provided support and remediation where issues arose and where they were reported to the NCAA. This also included assessment and investigation of abuse and threats sent by other digital means such as email.

For the duration of this pilot project, the Threat Matrix service focused on:

- **Monitoring** – of public accounts and mentions for abusive, discriminatory or threatening content.
- **Investigation & Action** – where content was detected and verified, it was actioned by Signify. Captured abusive content was reported to platforms, and evidence of content that passed the threshold of criminal activity / hateful communications were shared with law enforcement (via the NCAA) to support real-world action.
- **Wellbeing of individuals** – the central goal of the pilot study was to enable the NCAA to better understand how online abuse and threats affect its student-athletes, coaches and officials; while also putting in place a mechanism whereby perpetrators can be held to account.

During the pilot study 16 different categorisations of online abuse and threat were identified across the different championships (see [Filter categories](#) for more detail).

All detected and verified abusive content is reviewed and tiered in order to assess any action required (see [Gravity of abuse definitions](#) for more detail).



AI MONITORING: PROTECTING ATHLETES, COACHES & OFFICIALS NCAA PILOT STUDY

PROJECT SCOPE + COVERAGE

All participants and selected stages of competitions in the following championships were included in the pilot study:



DI Women's Volleyball Championship
(Regionals + Championships)



FCS / FBS Football Championships
(Championship Games)



DI Men's and Women's March Madness
(Second Round onwards)



DI Gymnastics
(Regionals + Championships)



DI Women's Softball Championships
(Super Regionals + Championships)



DI Men's Baseball Championships
(Super Regionals + Championships)

ACCOUNTS MONITORED

3,164

College Athletes

489

Team Coaches

197

Event Officials

165

Team Accounts

12

NCAA Official Channels

Coverage incorporates the relevant social media accounts across X/Twitter, Instagram, Facebook, and Tik Tok.

A BESPOKE PROCESS



Build a proactive net around participants: Scanning for abuse/threats of targeting social media accounts and individuals across social media platforms.



Proactively identify accounts used by abusers at scale - a combination of AI tech and the nuance of security and intelligence experts.



Unmask abusers using specialist Open-Source Intelligence (OSINT) forensic tools and real-world intelligence to **de-anonymise and identify abusers to an evidential standard.**



Report abusive accounts to social media platforms, pressuring the **removal of abusive accounts every month.** Removing the expectation on participants to report abuse themselves.



NCAA PILOT STUDY 2023-24

PILOT SUMMARY FINDINGS

AI MONITORING: PROTECTING ATHLETES, COACHES & OFFICIALS NCAA PILOT STUDY

SUMMARY ANALYSIS

During the pilot study, 16 different types of online abuse and threat were identified across the different championships.

Beyond General abuse (which made up 17% of all detected abuse across this study), Sexual abuse was the most prevalent form of abuse with 18% of all detected and verified abusive comments. This was followed by Sexist abuse (14%).

Sports Betting & Match-Fixing related content made up 12% of the total levels across all the championships covered during the pilot, with 743 instances of abuse triggered by this issue. The March Madness Championships contributed 73% of the detected Sports Betting & Match-Fixing related content across all covered Championships.

There was also a concerningly high number of violent posts / comments – 380 instances, making up more than 6% of the overall total across all events.

In terms of championships, March Madness participants (Men and Women athletes and officials) received the overwhelming majority of abuse and threatening content – 3,915 verified abusive / threatening messages. 80% of all verified malicious content across the entire pilot study was identified during the tournament, which was covered from the Second Round onwards.

It should be highlighted that these are very significant numbers of verified abuse/threat. For comparison, a three-week Tennis Grand Slam (Men and Women combined) may expect to see 1,000-2,000 verified abusive / threatening messages.

Although direct comparisons between the abuse / threat targeting Men and Women athletes are difficult because they were not equal in participant numbers, some observations can be made particularly in the same sport. For example, during March Madness, Women's tournament participants received almost 3x the abuse compared to the Men's.

In Baseball and Softball, the ratios of abuse types were fairly consistent between the Men's and Women's competitions. The main difference was in Sports Betting & Match-Fixing with the Women's competition receiving almost 100% more than the Men's; however, this was linked to cheating and corruption allegations, and not sports betting.

In some instances, an event and its athletes attracted the unwanted attention of bots. 92% of the content detected under the Sexual category in the Gymnastics championships was sent from bot accounts that encouraged online users to access sexual content. After Signify reported this content to the platform (X), all accounts detected as sending this content were suspended and comments removed from team and gymnasts' channels.

The pilot study also included monitoring for abuse and threats towards match/event officials and NCAA Officials, such as those on Selection Committees. Much of this content was either linked to sports betting or accusations of tournament rigging to favour certain teams and players. It was also identified that when head coaches criticised officials in the press, this negatively impacted on the perception of the match officials' decision-making and crossed over into the online space.

The pilot study focused on open-source data and did not cover abuse and threats sent via private direct messages (DMs) to athletes and officials. However, comments in the media indicated Men's and Women's athletes were heavily targeted via this private channel of communication. The NCAA and Signify also received reports directly from officials of abuse and threats; these were investigated and relevant action taken where necessary.



AI MONITORING: PROTECTING ATHLETES, COACHES & OFFICIALS NCAA PILOT STUDY

EXECUTIVE SUMMARY BREAKDOWN

WHAT THE DATA ILLUSTRATES – TOP 10 FINDINGS

1. Sexual abuse was the most prevalent type of content used to target athletes (18% of all detected abuse).
2. Sports Betting related abuse and match-fixing allegations made up 12% of all abuse identified – 743 instances (More than the level of Racist abuse targeted at athletes). This demonstrates that other college sports (outside of the pilot scope) will already be receiving volumes of abuse.
3. Abusive content is sent privately as well as publicly. College athletes received abuse and threats over DMs that they have to deal with on their own.
4. Violent content was also concerningly high with 380 instances (6% of the overall study).
5. Female athletes received 3x the amount of abuse of men in a championship like March Madness.
6. March Madness athletes need protecting from social media abuse. The extra attention that fell upon this year's event drew both viewers and abusive content in equal measure – 80% of the detected abuse in this study came from that event.
7. Threat and risk is present in amongst the content captured. This study marked out 96 accounts for further investigation in this regard.
8. Risks do not always come from the big events – in some instances, Volleyball and Gymnastics flagged more concerning instances of concern than March Madness or Football Championships.
9. The mental health of athletes will be impacted by not just the type of content, but also the volume. One athlete received over 1,400 abusive messages in the space of 2 weeks.
10. It is not just players – Coaches, Officials and NCAA Committee Members require support on this issue.

AI MONITORING: PROTECTING ATHLETES, COACHES & OFFICIALS
NCAA PILOT STUDY

KEY FIGURES: ALL CHAMPIONSHIPS

	TOTAL MESSAGES (Mentioning target accounts)	TOTAL FLAGGED BY AI (Reviewed by analysts)	TOTAL VERIFIED ABUSIVE (Reported to platform)
 VOLLEYBALL	75,808	2,497	85
 FOOTBALL	58,146	1,834	219
 BASKETBALL	848,096	54,309	3,915
 GYMNASTICS	49,215	773	55
 SOFTBALL	83,462	2,554	142
 BASEBALL	198,738	10,445	604
TOTALS	1,313,465 Mentioning target accounts	72,412 Reviewed by analysts	5,020 Reported to platform

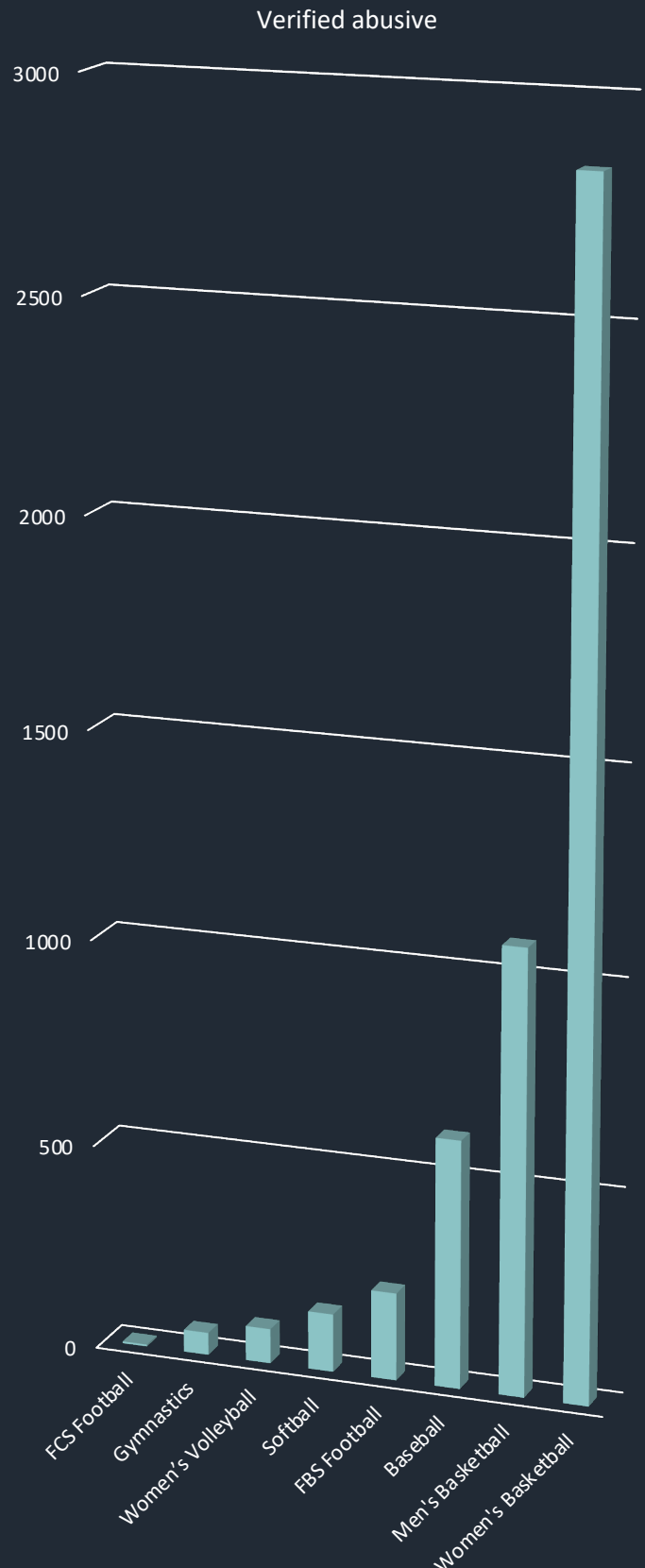
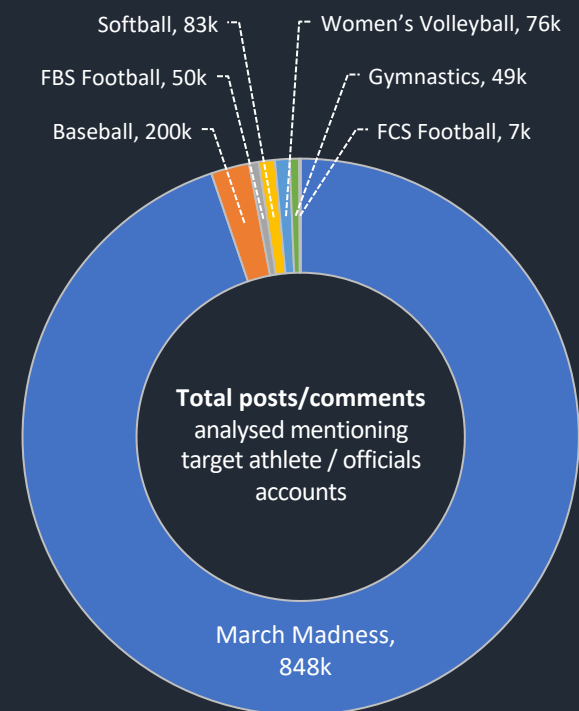
AI MONITORING: PROTECTING ATHLETES, COACHES & OFFICIALS NCAA PILOT STUDY

COMPARATIVE ANALYSIS

A total of 5,020 posts and comments were reported to the social media platforms for action. Of these, a number were also shared with the NCAA and law enforcement for further assessment, and this is detailed in the Investigations section of this report.

80% of all verified abusive or threatening content was identified during March Madness. This reflects the profile and popularity of the tournament as well as the sports betting interest in it. On top of this the Women's competition increased in profile again this year, including that of key athletes, adding to the volume of abuse as a result.

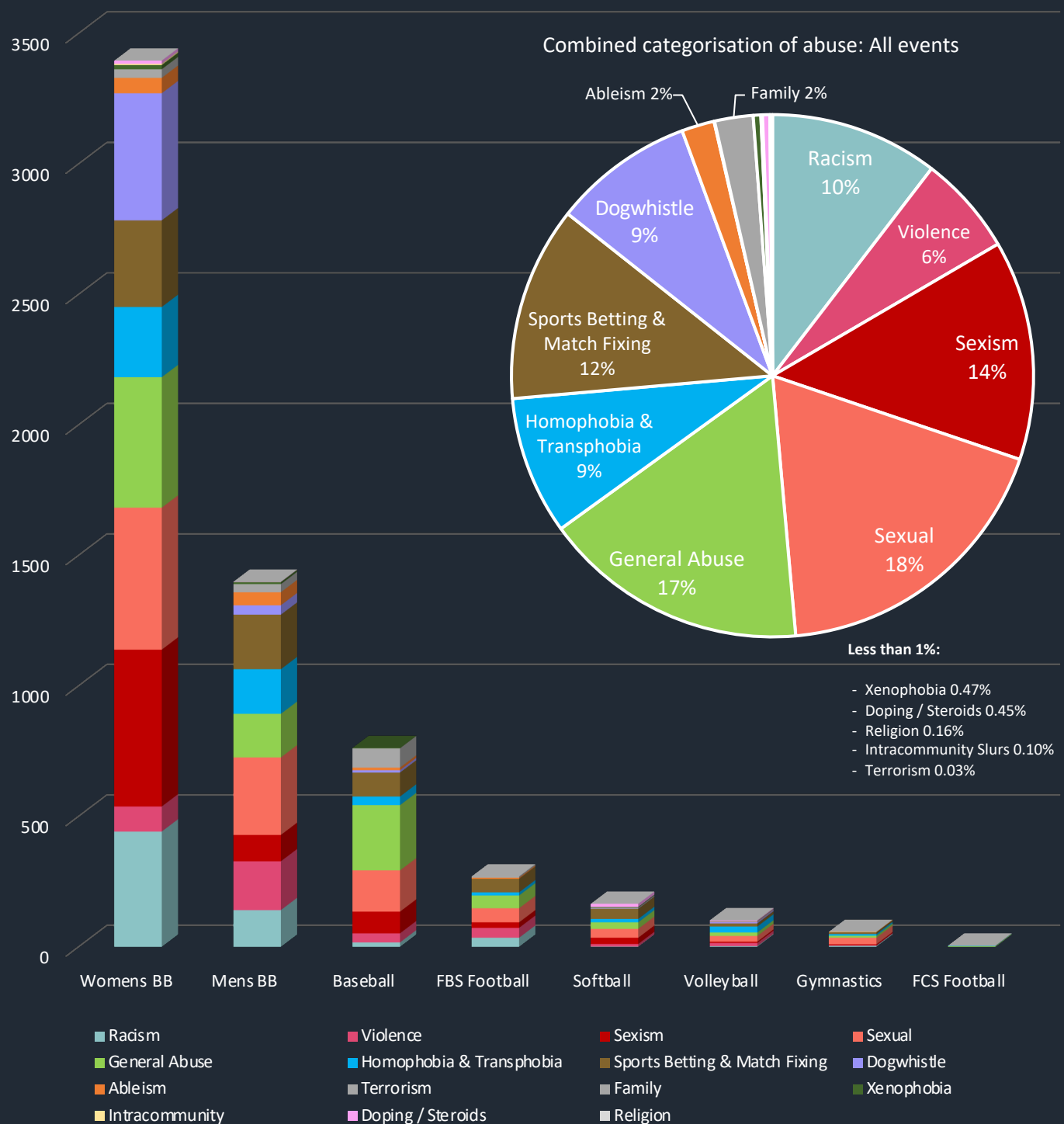
The pie chart (below) demonstrates the dominance of March Madness in terms of data covered across the pilot study. The chart illustrates the total number of messages mentioning player handles in association to each of the events. March Madness generated 850k posts (all of which went through the Threat Matrix system for review) vs other events in the low hundreds of thousands or tens of thousands.



AI MONITORING: PROTECTING ATHLETES, COACHES & OFFICIALS NCAA PILOT STUDY

COMPARATIVE STATISTICS

The chart below illustrates the difference in volume between each sport / event, together with a breakdown of the categorisations of verified abuse.





NCAA PILOT STUDY 2023-24

ISSUE INSIGHTS + EXAMPLES

WARNING: THE FOLLOWING PAGES CONTAIN EXPLICIT EXAMPLES OF DISCRIMINATORY, VIOLENT AND THREATENING CONTENT

AI MONITORING: PROTECTING ATHLETES, COACHES & OFFICIALS NCAA PILOT STUDY – ISSUES FOCUS

SPORTS BETTING & MATCH-FIXING

Sports Betting & Match-Fixing related content made up 12% of the total levels across all the championships covered during the NCAA pilot study, with 743 instances of abuse triggered by this issue.

Within the malicious content targeting participants in the public domain there were overt references to sports betting as well as content assessed as being motivated by gambling.

For this study, “Angry Sports Bettors” are defined as individuals who engage in problematic and intrusive communication due to match events and results contradicting bettors’ predictions and thus fuelling abusive messages towards student-athletes, coaches and officials.

The March Madness Championships contributed 73% of the detected Sports Betting & Match-Fixing related content across all covered Championships. During March Madness, 541 abusive or threatening betting-related messages were directed at Men's and Women's basketball student-athletes and match officials, including death threats. Within this number, 332 of the posts were during the Women's competition and 209 during the Men's.

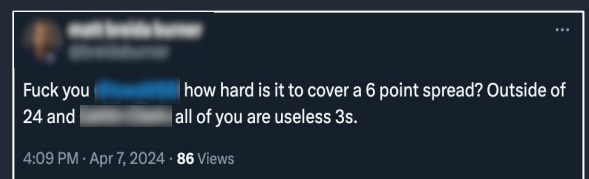
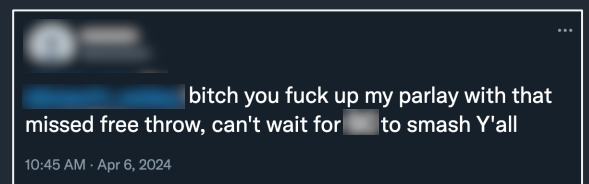
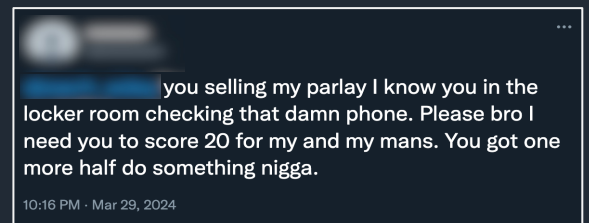
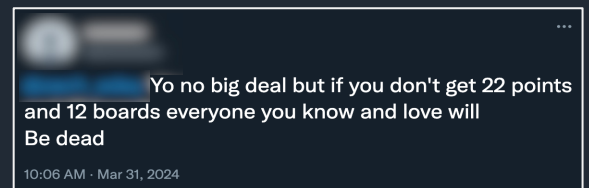
March Madness was not the only event where this issue was apparent. Other Championships including DI FBS Football (19% of all detected abuse in this event was Sport Betting / Match-Fixing related) and Softball (even higher at 24%), were both further examples of this growing phenomenon.

This is becoming a familiar and growing pattern across sports. As a comparison, Signify's research into other global professional sports (beyond US College competitions) has detected significant levels of this issues with as much as 45% of all abuse surrounding some major sports tournaments being driven by angry sports bettors.

Allegations of match-fixing and rigging were a consistent feature across the championships during the pilot study. Where certain sports did not have a high betting interest (e.g. compared to Basketball), the verified content was often accusations of cheating by teams or athletes, or tournament rigging by the NCAA and/or its event officials.

It should be noted that the numbers in this study are for open-source data and do not include abuse or threats participants received via their DMs / Private messages, where the issue is likely to be even worse in many cases.

743 abusive / threatening content containing references to sports betting / match-fixing



AI MONITORING: PROTECTING ATHLETES, COACHES & OFFICIALS NCAA PILOT STUDY – ISSUES FOCUS

SEXUAL CONTENT

Sexual abuse was the most prominent content targeting players in the study.

Both male and female athletes were targeted with this kind of content, and it was identified across multiple sports / events.

The examples displayed here were selected to illustrate this, although please note that these are the more moderate examples detected in the study.

GO HOME DILDO BITCHES AND NEVEE COME BACK!!! FUKIN DISGRACE

1:24 AM · Mar 29, 2024 · 4 Views

You are a disgusting waste of a vagina. If you are not proud of my country 🇺🇸🇺🇸🇺🇸 GTFO!!

RACISM

The ratio of Racism during the pilot study was 10% of all verified abusive and threatening content.

Every championship in the pilot study, apart from the FCS championship game, observed a level of online Racist abuse or threat.

In terms of volumes, March Madness as a whole and the Men’s and Women’s competitions individually, had the highest levels. These were then followed by the FBS championship game.

Cry you gorilla

5:40 AM · Apr 6, 2024 · 21 Views

Women basketball players were subjected to racist abuse during the tournament. Individual players, and most teams and referees endured this type of mistreatment.

The level of Dogwhistle content during the Women’s championships should be highlighted as well in connection with racism.

Throughout the NCAA Men’s Basketball tournament, there were disturbing incidents of online users directing racist remarks at players. Toxic online fans resorted to racist mockery, comparing players to monkeys and labelling them as thugs.

One quarter left and they start sh!tying the bed and playing like a bunch of retarded monkeys

7:48 PM · Mar 25, 2024 · 12 Views

VIOLENCE

The ratio of Violence during the pilot study was 6% of all verified abusive and threatening content.

FBS Football (14%), Men’s March Madness (13.3%), and Volleyball (12%) all received high proportions of violent abusive or threatening content. This was directed at athletes, officials, and coaches.

Violent abusive and threatening content was identified by Threat Matrix AI and then risk assessed by in-house analysts and experts. When content was assessed as Tier 3, an expedited alert was sent to the NCAA.

OSINT investigations were conducted into the senders of Tier 3 content with evidence shared with the NCAA and law enforcement for further assessment and action.

i hope your family dies

5:09 AM · Jan 9, 2024 · 19 Views

You bitches can go suck big hairy moose balls. Fuck you and the high horse you think you rode in on. I hope NONE of you live to see your graduation

11:20 AM · Apr 2, 2024 · 20 Views

I am in Cleveland and Im gonna wait for you outside the arena to beat the fuck out of your knees with a spiked baseball bat

12:00 PM · Apr 6, 2024

AI MONITORING: PROTECTING ATHLETES, COACHES & OFFICIALS NCAA PILOT STUDY – ISSUES FOCUS

HOMOPHOBIA + TRANSPHOBIA

Homophobia and Transphobia were consistent issues across the championships ranging from 4.33% (Baseball) to 22% (Volleyball).

While these issues have been mostly grouped together in the monitoring and analysis, the pilot study demonstrated that they need to be addressed individually given the range of issues and terms connected to them.

In the Volleyball championships the NCAA Volleyball account received 47% of the total volume of abuse with a large proportion of this being Transphobic.

Trans issues are a lightning-rod for online abuse in sport in general at present, and this is also the case in college sports. It is amplified by certain social media commentators and in college sports also demonstrates the crossover between online abuse and real-world actions such as demonstrations by protesters.

Some tweets quickly gained viral attention, some amassing over 5 million views and numerous transphobic replies.

DOPING + STEROID ACCUSATIONS

From a sport integrity perspective, allegations with regards to Doping was not a significant issue during the pilot study.

There were 28 instances of Doping related content during the pilot (less than 1% of all detected messages).

However, there was a limited spike during the Softball championships where these accusations made up 7% of all detected abusive messages at the event.

yup just like ██████████ softball forcing steroids down those poor girls throats, it's disgusting to witness
5:33 PM · Jun 3, 2024 · 58 Views

Steroids, they feed these ██████████ softball players roids. Doesn't take a rocket scientist to figure it out.
1:40 AM · Jun 6, 2024 · 187 Views

MATCH OFFICIALS

The pilot study included monitoring for abuse and threats towards match/event officials and NCAA Officials, such as those on Selection Committees.

Much of the malicious content targeting these individuals was either linked to sports betting or accusations of tournament rigging to favour certain teams and players. Match/event officials and Selection Committee members received abuse or threats connected to allegations of rigging championships both in terms of fixtures and team selections, as well event results (with the latter allegations targeting match/event officials only).

It was also identified that when head coaches criticised officials in the press, this negatively impacted on the perception of the match officials' decision-making and crossed over into the online space.

.@NCAA @MarchMadnessMBB PLEASE KILL THESE REFS
1:20 AM · Mar 24, 2024 · 73 Views

This home plate umpire really don't know how to call the right calls. THIS SHIT IS RIGGED!!!! 🤬 #Fuck█████████NCAASoftball
3:39 AM · Jun 7, 2024 · 80 Views

Above are examples of open-source threats to match officials which were detected in real-time by Signify and assessed further for risk. In many instances, Signify identified abuse but it was not targeted at the match officials directly on their own channels. Instead, it was mentions of match officials in general or the individual's name without their social media handle.

Such public abuse notwithstanding, a number of match officials did receive abuse and threats over social media through private Direct Messages (DMs) that they reported to the NCAA, and which were investigated by Signify with appropriate action taken.

ONLINE ABUSE IN US COLLEGE SPORTS
PILOT STUDY 2023-24

MENTAL HEALTH



AI MONITORING: PROTECTING ATHLETES, COACHES & OFFICIALS NCAA PILOT STUDY

THE MENTAL HEALTH TOLL ON VICTIMS

The risks and mental health challenges associated with being a victim of online abuse or threats are real and have a direct and immediate effect on athletes, coaches, officials, and their families. This can impact them on both a personal and professional level, and ultimately affect their wellbeing and ability to perform at their best.

The abhorrent abuse that Signify identifies and investigates across social media and all online platforms demonstrates how athletes, coaches, officials, and even their families can be exposed to serious risks.

Being a victim of online abuse and continuous hatred online can lead to a range of real-world consequences and symptoms, including (but not limited to) instances of anxiety, depression, loss of confidence, panic attacks and in extreme cases, can be a contributory factor in suicidal tendencies / actions.

Intervention, recognition and outreach can provide a victim with a real sense of support and protection – at the very least, recognition of the issues.

During the pilot study, and March Madness in particular, this was especially pertinent with regards to the abuse and threats directed at several athletes (both Women and Men) as well as some of the officials (selection committee members, coaches, referees), and the support they may have needed as a result. In one serious case of online threat identified with regards to the FBS Championship game, the NCAA reached out to the athlete, his academic institution and law enforcement.

The effect online abuse and threats can have on athletes and other participants should not be underestimated. Even when an individual who is targeted with malicious online content states they are 'fine' when consulted, this should not be assumed to be the case. The targeting of them may continue after a welfare check, which is why action needs to be taken to protect them long-term.

By putting the Threat Matrix service in place, the NCAA can demonstrate to participants that they are being protected and that proactive action is being taken on their behalf to hold abusers to account. Participants in sport are regularly demanding action be taken against those who send them abuse and threats (as opposed to simply moderating/hiding hateful content).

Threat Matrix provides this proactive support to athletes, coaches, and officials giving them the assurance that perpetrators are being flagged, investigated, and reported to the appropriate authorities.

ONLINE ABUSE IN NCAA CHAMPIONSHIPS
PILOT STUDY 2023-24

ANNEXE: GLOSSARY / DEFINITIONS

AI MONITORING: PROTECTING ATHLETES, COACHES & OFFICIALS NCAA PILOT STUDY

GRAVITY OF ABUSE DEFINITIONS

Tier 0 - there was concerning communication*, but not of a level where we would expect platform action – client to review with a view to informing safeguarding / athlete welfare teams in cases of prolific volumes of abuse.

Tier 1 - the concerning communication* may be sufficient to warrant intervention from the platform for breach of guidelines – client to review with a view to informing safeguarding / athlete welfare teams in cases of prolific volumes of abuse.

Tier 2 - egregious abuse or threat warranting intervention from platforms for breach of guidelines and additional examination by the client – client to prioritise and alert safeguarding / athlete welfare teams and event / competition teams (cf. entry to events) in the case of verified accounts.

Signify's in-house OSINT investigations activated within this Tier.

Tier 3 - egregious abuse or threat reaching a criminal threshold or case of imminent threat – intervention for breach of platform guidelines, preparation of evidence pack and referral to law enforcement – client to prioritise and alert safeguarding / athlete welfare teams and event / competition teams (cf. entry to events) in the case of verified accounts.

Signify's in-house OSINT investigations activated within this Tier.

**Concerning communication = any abuse, threat, intrusive or problematic communication either posed publicly or sent directly, which may cause an emotional or physical risk to the recipient.*

AI MONITORING: PROTECTING ATHLETES, COACHES & OFFICIALS NCAA PILOT STUDY

TERMINOLOGY + GLOSSARY

- **Abuse / Abusive posts and accounts:** refers to content that includes discriminatory, egregious and aggravated terminology.
- **Discriminatory flags:** Posts flagged for content that may include racist, homophobic, sexist etc. terms.
- **Dogwhistle:** An abusive message clearly exploiting a racist trope without explicitly using directly racist language.
- **Flagged posts:** Posts flagged for content that may include personally abusive or discriminatory content. Personally abusive content can include calling someone a c**t, etc.
- **GRT:** Gypsy Roma and Traveller communities.
- **Inclusion criteria:** a post will have mentioned one of the monitored player handles.
- **Special terms:** Mentions of particular terms (as an example - Antiziganism).

FILTER CATEGORIES

Identified posts are based on text, with layered word, emoji and phrase categorisation and an AI-empowered threat detection algorithm.

Our flagging system is based on a series of issue specific keywords, from the following example categories (this is not an exhaustive list):

- Ableist
- Child abuse
- Dogwhistle
- Doping allegations
- Family members
- General abuse & allegations
- Homophobia
- Intra-communal abuse
- Misogyny (incl: Sexist & Sexual)
- Sports Betting & Integrity issues
- Racism
- Religion (incl: Antisemitism & Islamophobia)
- Special terms