

# Nearly 70% of Premier League footballers are abused on Twitter according to AI used to sift through millions of tweets

August 12 2022, by Bertie Vidgen and Angus Redlarski Williams

---



Credit: Mike from Pexels

As the new Premier League football season gets underway, a few things are certain. There will be goals, drama and excitement, and unfortunately, players will be subjected to vile abuse on social media.

My colleagues at [the Alan Turing Institute](#) and I have published a [report](#), commissioned by [Ofcom](#), in which we found that seven out of ten Premier League footballers face abuse on Twitter. One in 14 receive abuse every day.

These are stark statistics, with huge implications for player welfare. Other analysis has revealed a high rate of online abuse, [particularly racist abuse](#), of footballers that has gone largely [unchallenged](#) by football governing organizations. Mental health is increasingly a [concern in football](#), and there is [plenty of evidence](#) that online abuse can lead to a range of [mental health](#) problems, from depression to suicidal thoughts.

Our report is one of the first to use [artificial intelligence](#) (AI) to systematically detect and track online abuse against footballers at scale. This is almost impossible to do manually because of the sheer size and complexity of [social media](#).

We focused our analysis on Twitter because it is widely used by footballers and fans, and it makes its data freely available to researchers. In total, we collected 2.3 million tweets that mentioned or directly replied to tweets from 618 Premier League footballers during the first half of the 2021-22 season.

At the heart of our analysis is a new machine learning model developed by the Turing's [online safety team](#) as part of our [Online Harms Observatory](#). This model is able to automatically assess whether or not a tweet is abusive by analyzing its language.

To provide a benchmark for our AI model and a more in-depth

breakdown of the tweet content, we also hand-labeled 3,000 tweets, categorizing them as positive, neutral, critical or abusive. Critical tweets were those that criticized a player's actions on or off the pitch, but not in such a way that could be deemed abusive.

We acknowledge that categorizing tweets in this way is [to some degree subjective](#), but we sought to reduce human bias as much as possible by consistently applying the same definitions and guidelines to all tweets.

## What did we find?

Of the 3,000 tweets that we hand labeled, the majority (57%) were positive. Tweets routinely expressed admiration, praise and support for the [players](#), often using emojis, exclamation marks and other indicators of intense positive emotion. A smaller proportion of tweets were labeled as critical (12.5%), neutral (27%) or abusive (3.5%).

Our machine learning model, applied to all 2.3 million tweets, found that 2.6% contained abuse. This might sound like a low percentage, but it represents almost 60,000 abusive tweets over just five months.

Abuse is widespread: 68% of players received at least one abusive [tweet](#) during this period. But players have very different experiences online: just 12 players received half of all abuse. Cristiano Ronaldo, Harry Maguire and Marcus Rashford received the most abusive tweets.

Abuse also varied hugely over the course of the season, with big peaks following key events. For instance, the number of abusive tweets spiked on August 27 2021, when [Manchester United re-signed Cristiano Ronaldo](#), and November 7 2021, when [Harry Maguire sent a tweet](#) apologizing after Manchester United lost to Manchester City.

We found that about 8.5% of abusive tweets (0.2% of all [tweets](#))

attacked players' identities by referencing a protected characteristic such as religion, race, gender or sexuality. This is a surprisingly low proportion given the concerns about racial abuse of footballers online. But we only looked at identity attacks using keywords (whereas we had a full AI solution for identifying abuse), and did not look specifically at the experiences of non-white players.

## Being a good fan online

Addressing online abuse is not an easy task—finding and categorizing abuse is technically difficult and raises fundamental questions around free speech and privacy. But we cannot leave abuse unchallenged.

Some [social media platforms](#), including Twitter, are already taking steps to improve their trust and safety processes, but more can be done. This may include amplifying and promoting content that is not abusive; giving additional, practical support and advice to players (and others) who are receiving large amounts of abuse; and making more use of properly governed machine learning tools to automatically detect and take action against abuse. Ultimately, platforms should shoulder most of the responsibility for cleaning up their services.

In general, [abusive content is underreported](#). Ofcom polled the public about their [experiences of players being targeted online](#), finding that more than a quarter of teens and adults who go online (27%) saw abuse directed at a footballer last season. Among those who came across [abuse](#), more than half (51%) said they found the content extremely offensive and around 30% didn't take any action in response.

There's absolutely nothing wrong with being emotional about football and expressing how you feel online but we should all take care to not cross the line into being abusive and intimidating. And if you see someone else being abusive, be proactive. Report it and show that this

content has no place in football (or anywhere else). Football is a beautiful game, and we can all help to keep it that way.

This article is republished from [The Conversation](#) under a Creative Commons license. Read the [original article](#).

Provided by The Conversation

Citation: Nearly 70% of Premier League footballers are abused on Twitter according to AI used to sift through millions of tweets (2022, August 12) retrieved 26 April 2024 from <https://phys.org/news/2022-08-premier-league-footballers-abused-twitter.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.