

The background of the entire page is a repeating pattern of blue smartphones. Each smartphone screen displays a white icon of a hand holding a heart, which is a common symbol for social media likes or comments. The smartphones are arranged in a staggered grid, creating a dense, textured effect.

Like, Comment, Discriminate Racist Discourse on Dutch Instagram



The **Digital Us**

Content

About The Digital Us	2
Introduction	3
1. Method	4
2. Instagram Posts	8
2.2 Actors	9
2.3 Themes by Actor	11
2.4 Interactions by Actor	11
3. Instagram Comments	13
3.1 Volume per type of racism	13
3.2 Racist Slurs	15
3.3 Dutchness and Whiteness	16
3.4 Worthiness for Work or Asylum	18
3.5 Dog Whistles	19
4. Conclusion	21

About The Digital Us

'The Digital Us' was set up by the [Sharing Perspectives Foundation](#) and [Build Up](#) in 2023. It provides a learning and action space for young people keen to become upstanders instead of staying bystanders in racist online debates in the Netherlands.

The project invites young people to take action to make social media into a more accessible and inclusive space where Dutch youth who experience racism online are able to express themselves safely and participate fully in debates. On social media, we see growing hostility and hardening discussions. The responsibility for combating racism online still lies too much with the people who fall victim to it. This project aims to change that by bringing together people who experience racism online with those who do not, but want to learn to be effective allies to make online interactions more inclusive and safe.

We do this through three core activities: Dialogue, training and implementation. Accompanied by experienced facilitators and trainers, space is provided for open, honest and empathetic conversations. Participants learn from each other, and with each other, about which strategies work to have a positive impact on social media. Together with others, they apply these strategies to make social media more inclusive.

The present analysis provides the evidence base informing where and how participants should intervene in online debates to be effective. The authors are Claudia Meier, Andrew Sutjahjo, Bart van der Velden and Sander van Haperen. The report was finalised in January 2024.

'The Digital Us' is made possible by financial support from [Fonds 21](#), [Stichting Democratie & Media](#), [VSBfonds](#) and [vfonds](#) (Nationaal Fonds voor Vrede, Vrijheid en Veteranenzorg).

FONDS 21

 **Stichting
Democratie &
Media**

VSBfonds,
iedereen doet mee

 **fonds**

Introduction

We encounter racism online every day, to the extent that social media and hate speech sometimes seem inseparable. Positions appear to be increasingly extreme, and differences in opinion no longer support healthy debate. Instead, people reinforce their own beliefs and use extreme language against those who hold different opinions or identities. In the Netherlands, this tension is most evident in discussions about topics like migration and is frequently linked to racism.¹

Social media, with its algorithms promoting extreme views, plays a significant role in exacerbating this issue.² For instance, the use of racist language on Dutch Twitter (X) has skyrocketed since 2018.³ This is making social media increasingly toxic, for young people in general, and for people of colour in particular. The problem is not restricted to the online space – it affects offline life in the Netherlands as well. Racist comments are becoming more and more accepted in Dutch politics.⁴

We would like to see a world, also online, where everyone is free to express themselves without worrying about racist attacks. The Digital Us is joining a growing movement of people and organisations working from different angles to make this world a reality.⁵ We train 75 young people to intervene in racist debates on social media. This intervention can take the form of commenting in racist threads in a way that encourages those posting racist content to stop, or by posting alternative, unifying narratives that provide a visible counterbalance to racist, divisive language. The goal is to ensure that racist expressions do not remain the unchallenged, dominant narrative. Such interventions signal that racism is not acceptable and help prevent a slippery slope toward the normalisation of increasingly extreme racist content.

In order to do this effectively we will first need to answer a simple question: where are racist comments online most prevalent? Of course, there are many online spaces dedicated explicitly to racist ideology, such as Stormfront. However, most people posting in these fora are firmly committed to their racist ideologies, and it is unlikely that a challenge from an outsider will change their behaviour. Instead, we ask where racist comments appear online where they influence the mainstream public debate. This will inform where and how participants should intervene in order to be effective. The analysis also provides us with the baseline data allowing us to monitor whether interventions made a difference in specific debates.

In what follows, we first explain our methodology in more detail. Second, we present findings about the Instagram posts for themes where racist comments are *likely*. Thirdly, we examine the actual substance of racist comments. Finally, in the conclusion we reflect on key findings and how these can inform the interventions.

¹ Universiteit van Amsterdam. 2021. Polarisatie in Nederland: hoe verdeeld zijn we? <http://tinyurl.com/4npuxay8>

² Ashoka. 2023. Why Social Media Amplifies Extreme Views – And How To Stop It. <http://tinyurl.com/yxp3jcx>

³ Movisie. 2020. #Hobbyhaters. De rol van sociale media bij polarisatie. <http://tinyurl.com/yrdhzkwu>

⁴ ECRI. 2019. ECRI-Rapport over Nederland. <http://tinyurl.com/3ywrnnu6>

⁵ See Sosha Duysker's recent call at <http://tinyurl.com/2spb43ph>; Meld. Online Discriminatie at <http://tinyurl.com/3r2db4kt>; the Movisie's online training #ThatMeenJeNiet! at <http://tinyurl.com/2p85fbs7> or De Goede Zaak: <http://tinyurl.com/4xchedws>.

1. Method

Why this report? This social media analysis reveals the specific types of Instagram accounts and themes associated with the use of racist language in the Netherlands. These insights will provide the evidence base for The Digital Us, to help decide where and how participants should intervene in online debates to be effective.

What do we mean by racist language? We focus on expressions that are negative towards individuals on the basis of their skin colour, their descent, or national or ethnic origin.⁶ We do not only focus on overt racism and forbidden hate speech, but also include implicit expressions that are covered by free speech. We opted for this broad understanding because these expressions can be just as hurtful to people who are targeted by racism, and contribute to a toxic digital environment.

Why did we focus on Instagram? We first did a preliminary analysis of racist terms on a variety of platforms, including Telegram, Instagram, Twitter, and Facebook. We found that racist comments from young people were commonplace in Instagram comments. Focusing on Instagram further offered the advantage of publicly accessible data with relatively few anonymised accounts. By contrast, while Telegram has racist content, it is mostly not public, and public Facebook pages are filled to the brim with vile and hateful language, however the groups consist mostly of an older audience. X, formerly Twitter, was ruled out for practical reasons, as it recently restricted API access making research into the posts and threads prohibitively expensive. This means that this report in the end focuses on Instagram, as it has many public profiles and pages, contains quite some racism, is scrapable, and most importantly, is used by young Dutch people, our target audience.

Why did we focus on young people? The purpose of this analysis is to inform the design of a training for young people to intervene against racism in online spaces. We think such interventions are particularly important for young people who rely to a large degree on social media for news gathering and opinion formation. These interventions are designed to leverage norm-setting among peer groups. While young people are particularly at risk of adopting extreme positions, they are also susceptible to norm-setting through social processes. Consequently, in this analysis we focus on young people when identifying racism online, in order to tailor further training and interventions to the experience of young people online.

How did we identify racist content? In order to identify racist online content we worked with an advisory group. This group of eight people consists of participants who are experts on social media, politics, and racism through their profession, study, or life experience. Each expert individually provided sample Instagram content in which they recognised racist expressions. This resulted in a collection of content exemplary of online racism. Semi-structured interviews with the advisory group then helped to identify further content, and distil keywords and racist expressions from that. These interviews have led us to refine the focus of our research, specifically:

- We decided to focus on Instagram comments rather than Instagram posts, since the comments are where the racist expressions appear more frequently and have a bigger impact on the debate.
- We defined nine themes for Instagram posts where racist comments are *likely* to be expressed.⁷ The themes we defined are groupings of specific keywords and expressions. (See Table 1

⁶ We use a simplified adaptation of the racial discrimination definition of the European Union: <https://shorturl.at/xJN25>.

⁷ We initially included posts related to Palestine and Israel, but then discarded it because a manual analysis showed there were many false positives for the terms we had defined.

below). The research team identified the initial set of keywords and themes based on bilateral discussions with individual members of the advisory group. We then identified further keywords and cut unnecessary or incorrect keywords by exploring examples with those keywords. The keywords themselves went through basic natural language processing of stemming and manual acceptance for each stemmed word. The advisory group then came together to validate and revise the list of keywords and themes, leading to the final version used.

- In our preliminary research we used keywords that we considered to have a high chance of being present in racist posts and comments to identify Instagram accounts. We narrowed down this selection based on the amount of followers, eliminating very small accounts, the amount of interactions, eliminating those accounts that had very limited interactions, and a preliminary analysis of the comments on the 10 most recent posts of the account to determine if there was racism within the comments. Through this preliminary exploration we identified 98 Instagram accounts⁸ where racist comments were most prevalent to include in further analysis. We then classified these accounts in different types of actors to support our analysis.
- We grouped racist keywords into four racism classes, following the same iterative process as for the post themes. The classes we defined are: Racist Slurs, terms questioning someone's "Dutchness" and Whiteness, judgements of people's Worthiness for Work or Asylum, and Dog Whistles. "Dog Whistles" are seemingly innocuous language that covertly alludes to discriminatory or racist ideologies. We then analysed the comments according to these four classes. (See Table 2).

How exactly did we do the analysis? We scraped the data for the accounts identified and classified them into the post topics, by actor type (such as journalist, influencer, etc.), and the comment themes, such as discussed above. The data was processed and classified using stemming and basic Natural Language processing methodology. To visualise the data for the analysis, we used the Phoenix process developed by Build Up and datavaluepeople.⁹ From this we derived patterns for Instagram posts and comments under these posts. Our data includes post interactions. Interactions are the measure Instagram tracks to understand how many people actually engaged with content, either by "liking" a post, or by being exposed to the content for a sufficient time rather than scrolling past.

What is the time frame? We included posts and comments from 1 September to 11 December 2023. Because the goal of the project is to design interventions in current debates, a recent time frame was selected to include relevant debates. We restricted the analysis to three months because our purpose is to develop an approach that is sufficiently nimble for intervening within a brief timeframe, and in light of quickly changing online debates.

What were the limitations of this method? The dataset only includes public posts and comments from public Instagram accounts. There is no possibility to analyse accounts set to "private" or direct Instagram messages. This privacy limitation is ethical, but means that the analysis is likely missing a big part of the picture. People who face racism online report that they frequently receive direct messages containing racist language. Rather than attempting to cover all instances of racism online, we opted for precision over recall. Nevertheless, what we found here is likely the tip of the racist iceberg. Another limitation is the scope of the project, which is limited to a semi-supervised content analysis rather than an in depth qualitative inquiry. This means a limited number of false positives are included: particular keywords led to the inclusion of comments that are not racist on further

⁸ The account list can be accessed here: <https://shorturl.at/fjl25>.

⁹ You can find more details and technical specifics on the Phoenix process here: <https://shorturl.at/bmrGT>.

inspection. False negatives are also of concern as our process of identifying racist keywords was not exhaustive and in all likelihood many comments that are in actuality racist were not included. Where applicable, we describe these instances in the individual sections. In addition, consultations with our advisory group have led us to believe that there are more extreme racist messages sent through direct messages than in comments.

How do we deal with the reproduction of racist language? This report reproduces racist language in order to discuss and combat it. It is with great reluctance that we include examples, but are convinced this is necessary in the interest of transparency and deeper understanding. These keywords are instrumental to our methodology that uses such language to identify racist material. We restricted the inclusion of such language to Table 2 and the examples in Section 3, which can be avoided.

Table 1: Instagram post themes where racism is likely to occur in comments (with their keywords)

Immigratie	Protest	Seksualiteit en gender	Politici met migratieachtergrond	
(minderjarige) asielzoekers	ME	Transgender	Sylvana Simons	Tunahan Kuzu
Arbeidsmigratie/ -nt	Actievoerders	Homoseksueel	Kauthar Bouchallikht	Senna Maatoug
Kennismigrant	Demonstratie	Lesbisch	Kati Piri	Mohammed Mohandis
Statushouder	Betogers	Biseksueel	Bahreddine Belhaj	Farid Azarkan
Asiel	Rellen	Queer	Zohair El Yassini	Mikal Tseggai
Vluchteling	Waterkanon	Trans personen	Mustafa Amhaouch	Ahmed Marcouch
Integratie	Activisten	Transgenderwet	Songül Mutluer	Gidi Markuszower
Massa-immigratie		Gendertwijfel	Habtamu de Hoop	Aboutaleb
IND / AZC		Pride	Stephan van Baarle	Dilan Yeşilgöz-Zegerius
Ter apel		Non-binair	Songul Mutluer	
Verblijfsvergunning		Homoseksuele		
niet-westerse Nederlanders		lgbti		
Spreidingswet		Non-binaire identiteit		
massa immigratie				

Nationale verkiezingen	Criminaliteit	Islam	Zwarte Piet	Discriminatie/Racisme
Tweede Kamerverkiezing	Overval	Islam	Sinterklaas	Discriminatie
Verkiezing tweede kamer	Gewelddadig	Islamitische	Intocht	Racisme
kabinetsformatie	Politie	Sharia	Zwarte piet	kwetsende spreekkoren
fractieiders	Misdaad	Jizya	Roet()piet	Lichtgetinte
Kamerzetels	Verdachte	Mekka	Roetvee()gpiet	Slavernijverleden
	Onderwereld	Moskee		Racistisch
	Liquidatie	Koran		Structureel racisme
	Moord	Taliban		Gediscrimineerd
	Doodslag	Allah		Inclusiviteit
	Marechaussee	Islamitisch Nederland		Niet-westerse Nederlanders

Table 2: Racism classes applied to comments (with their keywords)

Racist Slurs	Dutchness and Whiteness	Dog Whistles	Worthiness for work or asylum
kloepoek	Cultuur en Traditie	Waar ze vandaan kwamen	Gelukzoekers
Exotic Wat	Weinig nederlands aan	vandaan kwamen	Asielzoeker
Slaaf	Niks nederlands aan	Religie van vrede	Kansenparels
Slaven	Echte nederlanders	vanaf hier	Hardwerkend(e Nederlander)
Bonobo	Witte mensen	finnen	Echte vluchtelingen
Kkr zwarte	Eigen land	woke	Kennismigratie
back to Africa	Traditie	snowflake	Hoogopgeliède
Vieze mensen	Buitenlanders	je verkracht wordt	Kansloze vreemdelingen
Moslimtuig	Witte Mensen	dutse mense	vol is vol
Kutmarrokaan / -nen	Omvolkings propaganda	aap	Nederland is vol
Kutmarokkanen	Anti wit	bontkraagjes	profiteur
Mecca / Mekka	Eigen mensen	SJW	laagopgeleid
Arabieren	Wat spreek je goed nederlands	Social justice warrior	waardeloze
China virus	Waar komt je nou echt vandaan	Omvolkings propaganda	nietsnut
Spleet oog	zwarte piet	Onze samenleving	waardeloos
Poepchinese	kinderfeest	Ons samenleving	hulp met vertrek
Kut Chinees	minder minder	theedoek	werkeloos / werkeloze
Sambal Bij	NSB'er	kut volk	WW trekker
Pinda	verrader	kankervolk	goede burger
Moslim Tuig	eigen volk eerst	dat volk	geen belasting
Geitenneuker	naar je eigen land		democratie
n*ger	de taal leren		gelukseisers
kk zwarte	Minder Mocros		wij zijn vol
theedoek	leer Nederlands		terug sturen naar het land
hodvod			binnen blijven halen
spook			grenzen dicht

2. Instagram Posts

This section explores the patterns of Instagram posts that are likely to attract racist expressions in comments. It doesn't analyse the racist content directly. Rather, it serves as a preliminary step to understand the themes that provoke such comments. These themes include topics like immigration (where comments often contain slurs against immigrants) and posts about politicians with a migration background (where the focus of comments is more on their identity than the content of their posts). This initial step has aided in narrowing down the comments containing racist content, which will be discussed in the next section.

Our advisory group of experts informed the selection of themes where racist comments are most prevalent, as well as accounts posting on these themes. From these accounts we scraped all posts in the studied timeframe (early September to early December 2023). This corpus of posts was then narrowed based on the themes where racism was most prevalent in comments (table 1). In short, data selection was determined through a preliminary qualitative content analysis of Instagram comments and accounts.

2.1 Themes

We identified 1,580 Instagram posts related to themes where we anticipated racist comments. Figures 1-3 below visualise these findings. The most common theme was crime, with 891 posts, followed by discrimination and racism (183 posts), protest (146 posts), Islam (129 posts), and immigration (125 posts).

Examining the number of comments these posts received, the order changes. Crime, with 135k¹⁰ comments, receives by far the most comments, but Islam (46k), protest (30k), and immigration (26k) attract more comments than posts about discrimination and racism (20k). Although Islam ranks fourth in terms of the number of posts, it receives the second-highest number of comments. Posts about the national election also attract a disproportionately high number of comments. Interactions categorised by theme reveal a similar pattern to that of the comments. There are fewer posts about politicians with a migration biography.

Figure 1: Posts per theme

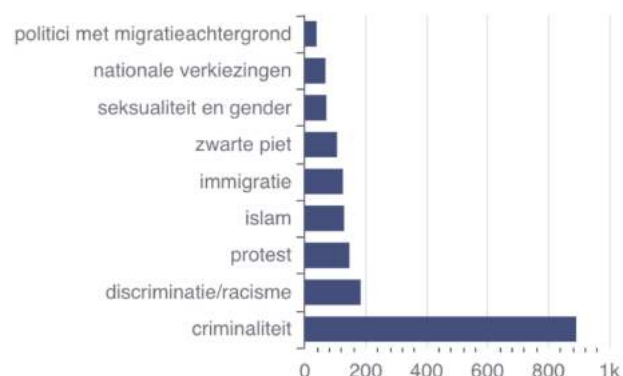


Figure 2: Comments per theme



¹⁰ In the interest of readability, we use 135k to mean 135.000; and 12M to mean 12 million.

Figure 3: Interactions per theme



Source for all graphs in this report: *The Digital Us*

2.2 Actors

To more accurately determine where our interventions can be most effective, we identified accounts where racist comments are most prevalent. We have classified these accounts by actor types to better understand the patterns of Instagram posts, depending on who is posting. Actor types include, for instance, influencers, journalists, or politicians.

Figure 5 shows the number of posts sorted by the type of actor. Most Instagram posts linked to the themes we researched are shared by national news organisations (1030 posts). Next are regional news outlets with 174 posts. A likely reason for the high visibility of news outlets is their frequent posting: on average, these accounts post about the identified themes nine times a day. In contrast, politicians typically post just once a day on average. Individual politicians share posts about our relevant themes nearly twice as often (76 posts) compared to political parties (42 posts).

Figure 5: Posts by actor

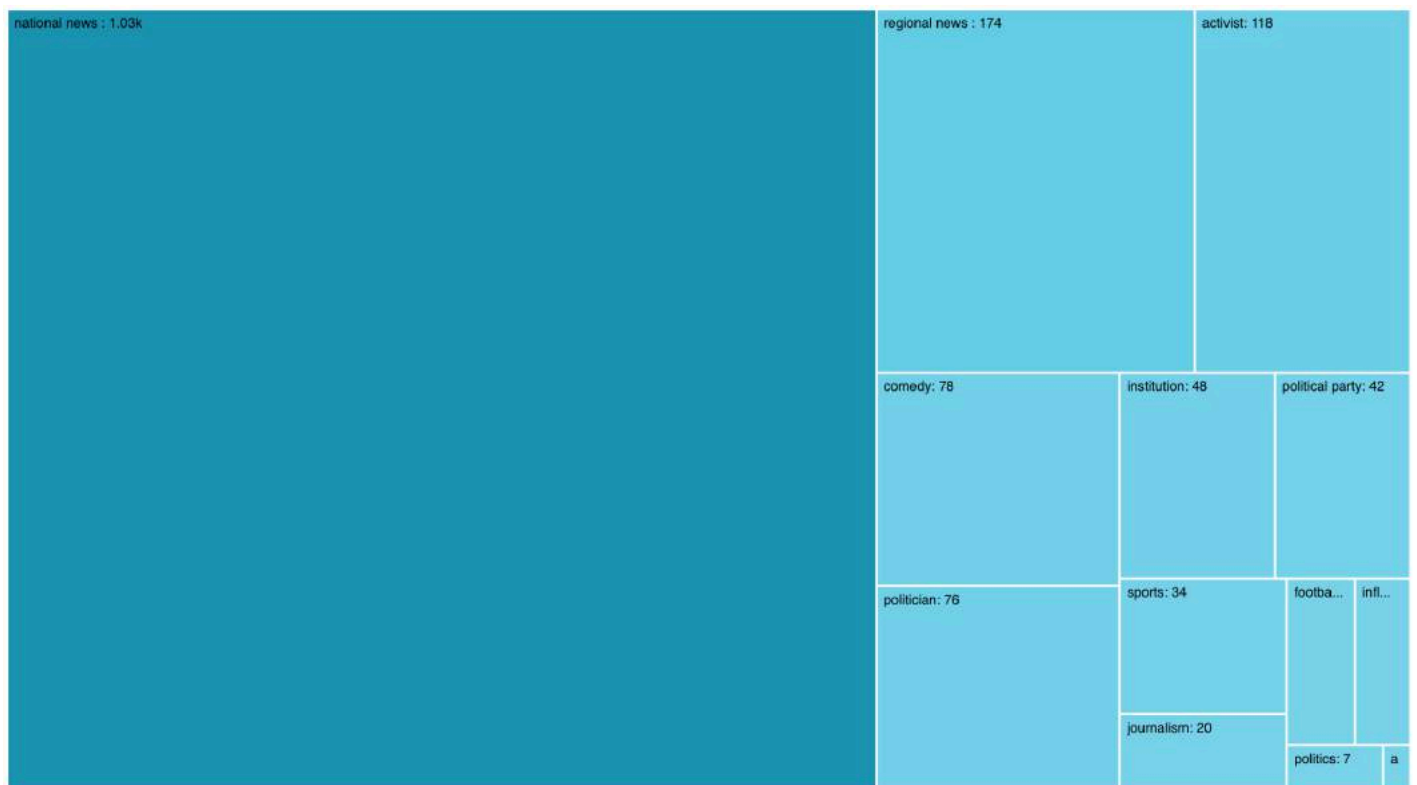


Figure 6 below shows the number of interactions each type of actor receives. Posts by politicians (113k) and political parties (99k) get relatively few interactions. National news posts attract the most

attention, with 9.28 million interactions, while sports (172k) and regional news (163k) lag far behind. Members of the advisory group observed that news outlets are where racist comments are most frequently found. This might be because these platforms have more interactions, which, due to the algorithms, makes these posts more prominent.

Figure 6: Number of interactions received per type of actor



When we examine individual accounts based on interactions, CESTMOCRO stands out significantly from the rest (refer to figures 7 and 8 below). Originally established as a meme account in 2017, CESTMOCRO now shares news from various sources and is notorious for attracting hateful comments and misinformation.¹¹ With almost 5 million interactions on posts relevant to our themes of interest, it far surpasses the next account, RapNieuwsTV, which has 690,000 interactions. CESTMOCRO mainly posts about crime (78 posts), Islam (25 posts), immigration (13 posts), and protests (13 posts).

Fig 7: Top 10 accounts by interaction

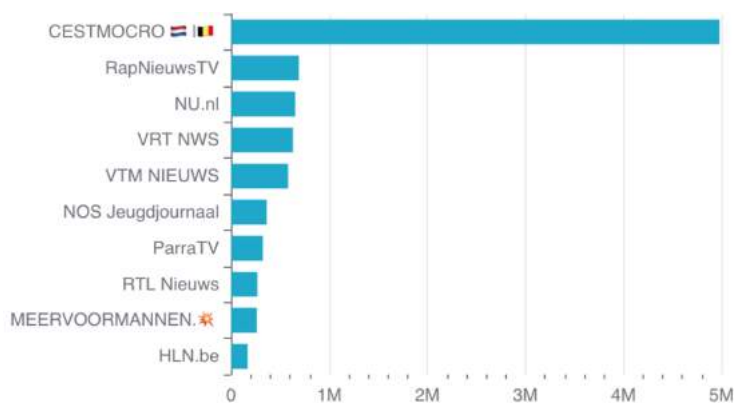
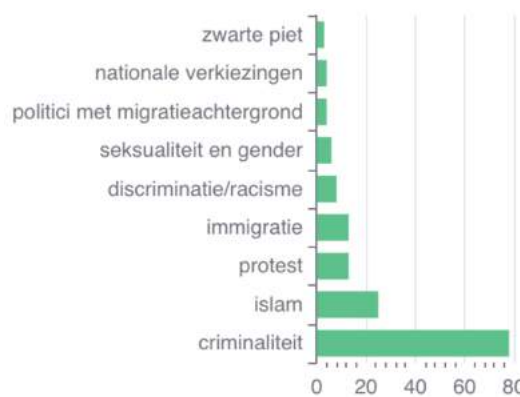


Fig 8: Posts by themes for CESTMOCRO



¹¹ See NOS News: <https://shorturl.at/lpvGJ>

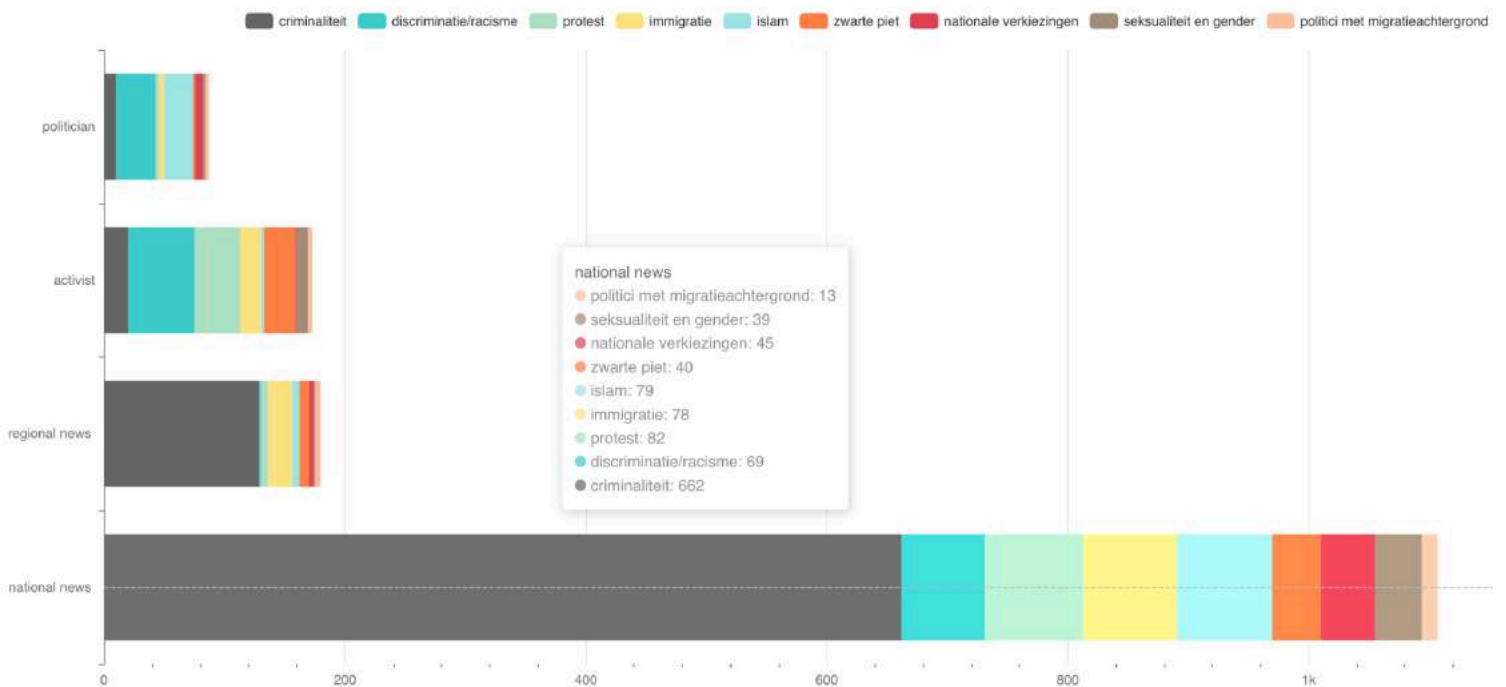
Racist comments under posts published by CESTMOCRO speak to Worthiness for Work or Asylum (40% or 3290 comments), Dutchness and Whiteness (39%), Racist Slurs (16%) and Dog Whistles (5%).

2.3 Themes by Actor

We identified which themes are discussed by type of actor. This helps to understand the focus of different actors and in turn informs where and on what themes interventions can be directed.

Figure 9 below shows the frequency of posts for each theme for the four most prolific actor types: national news, regional news, politicians, and activists.

Fig 9: Themes by actor



Each of these four actors concentrates on different topics. National news accounts mainly report on crime, protest, Islam, and immigration. Regional news outlets focus mostly on crime, which makes up over 70% of their posts, followed by immigration. They seldom cover other themes. In contrast, both activists and politicians mainly discuss discrimination and racism. For instance, activists often talk about protests and Black Pete rather than crime, while politicians speak about Islam more than crime. This pattern could be due to our specific choice of politicians and activists, who are especially concerned with issues of discrimination and racism

2.4 Interactions by Actor

Different actor types attract considerable interactions for specific themes. Knowing how many comments and interactions each actor's posts generate assists in pinpointing where participants should concentrate their engagement for the greatest impact. Three key patterns stand out among national news, sports figures, and politicians.

National news accounts post more about protests than Islam, yet posts about Islam draw far more comments and nearly double the interactions compared to those about protests. Posts on

immigration receive comparatively fewer interactions. Meanwhile, sexuality and gender, although less frequently posted about, garner more interactions in national news posts (see figures 10-12).

Fig 10: Posts (national news)

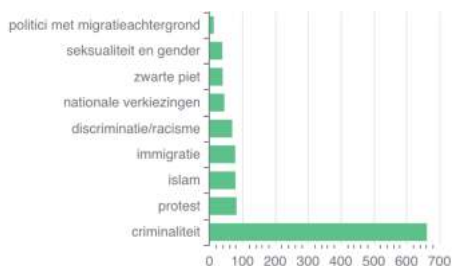
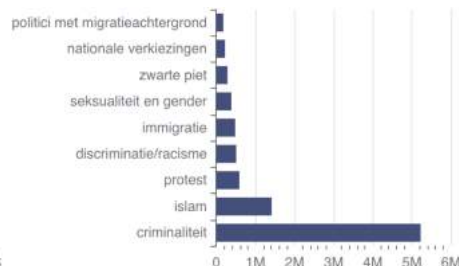


Fig 11: Comments (national news)



Fig 12: Interactions (national news)



Figures 13 to 15 show that sports-related Instagram accounts have made only a few posts on specific topics: just three posts about Black Pete and four concerning discrimination and racism. However, these few posts have prompted an unusually high number of comments and interactions. In comparison, the three posts about protests attracted no comments and very few interactions. Likewise, posts discussing sexuality and gender issues have received very few comments and even less overall engagement.

Fig 13: Posts (sport actors)

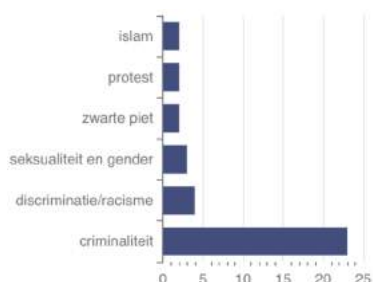


Fig 14: Comments (sports actors)

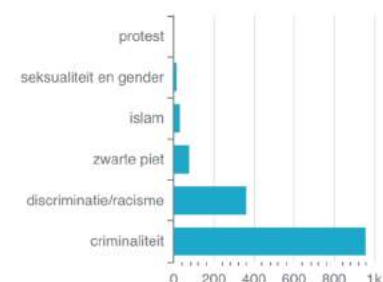
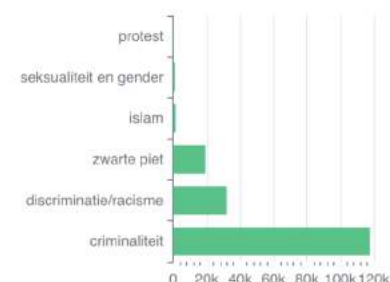


Fig 15: Interactions (sports actors)



Figures 16-18 show the patterns for politicians. It is noteworthy that the 24 posts about Islam generate an enormous amount of 2000 comments. This figure is nearly four times higher than the number of comments received for posts about discrimination and racism, despite there being more posts on the latter topic. Furthermore, the posts about Islam get around 60,000 impressions, which is three times the number achieved by posts on discrimination and racism.

Fig 16: Posts (politicians)

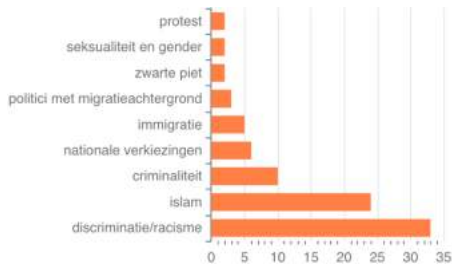


Fig 17: Comments (politicians)

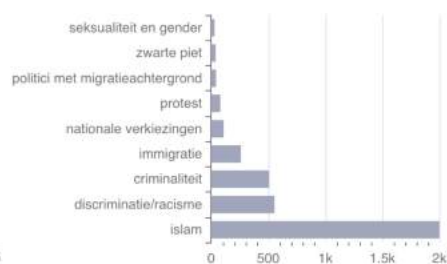
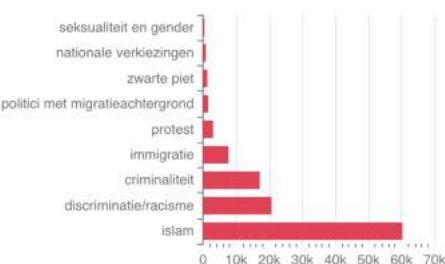


Fig 18: Interactions politicians



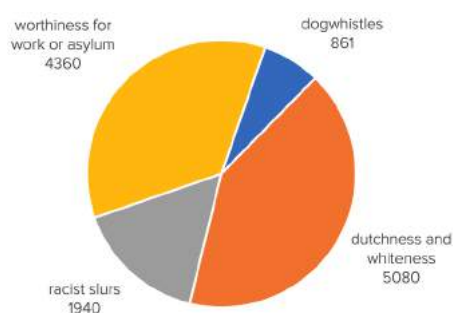
3. Instagram Comments

To better understand patterns of racism on Instagram, we categorised different types of racist remarks. Having pinpointed potential locations for these comments, we now focus on their content. We classified the comments under specific thematic posts based on racist keywords. As detailed in the methodology section, we divided these keywords into four categories: Dog Whistles, Dutchness and Whiteness, Racist Slurs, and Worthiness for Work or Asylum.

3.1 Volume per type of racism

Out of the 309,083 comments on Instagram posts related to themes in our dataset, 12,241 contained one or more racist keywords. In other words: every 25th comment was potentially racist. Even accounting for false positives and acknowledging that some comments may use these terms in a non-racist manner, this significant proportion of comments highlights the prevalence of racism on Instagram and underscores the urgent need for intervention. Figure 19 illustrates the breakdown of comments by type of racism.

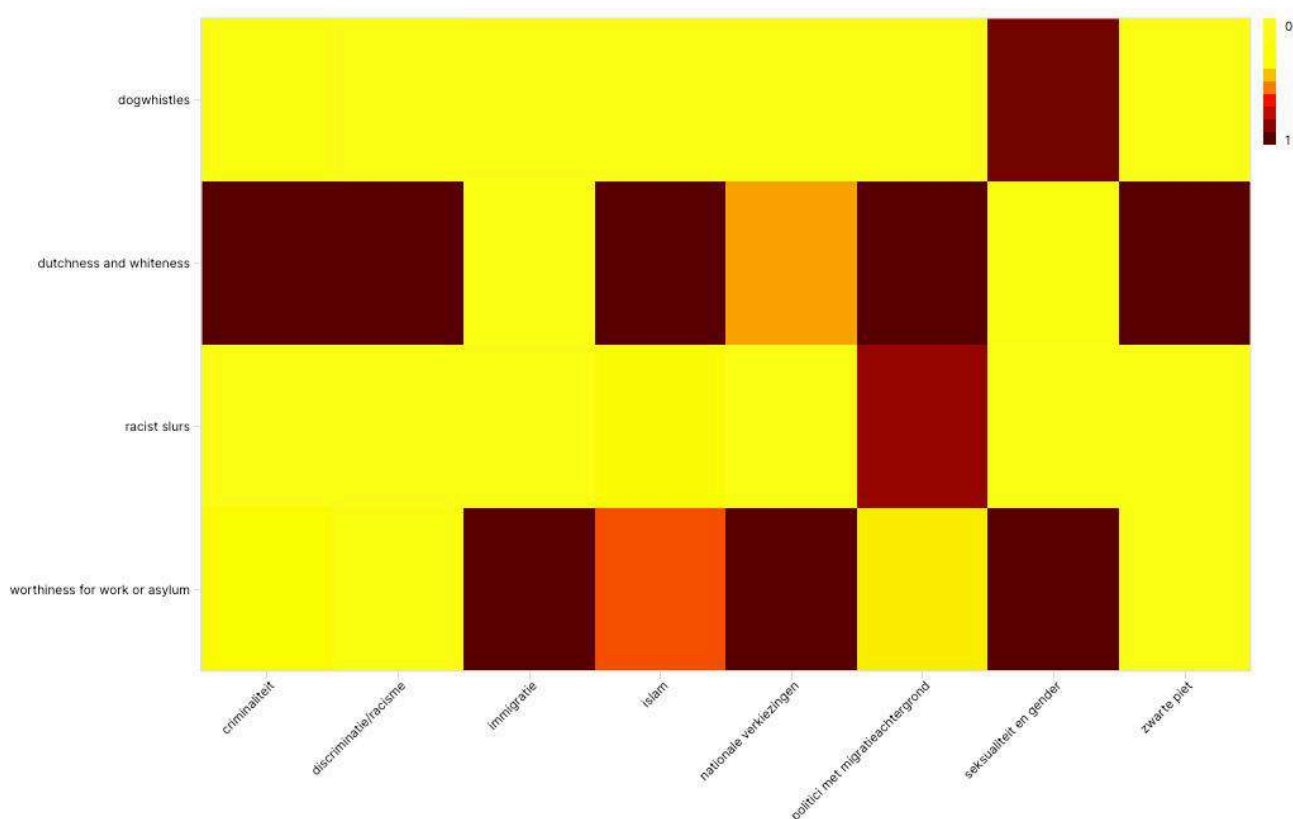
Figure 19: Number of comments by type of racism



41.5% of racist comments relate to Dutchness and Whiteness, 35.5% to Worthiness for Work or Asylum, 16% to Racist Slurs, and 7% to Dog Whistles. As detailed further in section 4.5, the Dog Whistle category shrinks to 571 comments after excluding the keyword “wokeness”, identified as a false positive upon closer examination. We discuss these four categories in greater depth in subsequent sections.

Overt Racist Slurs are often removed through content moderation, which explains their lower frequency. Additionally, our detailed analysis uncovered many false positives within Racist Slurs. Nonetheless, the finding of nearly 1,940 comments containing such language suggests that Instagram's content moderation is not completely effective. Figure 20 displays the occurrence of the four racism classes across different post themes.

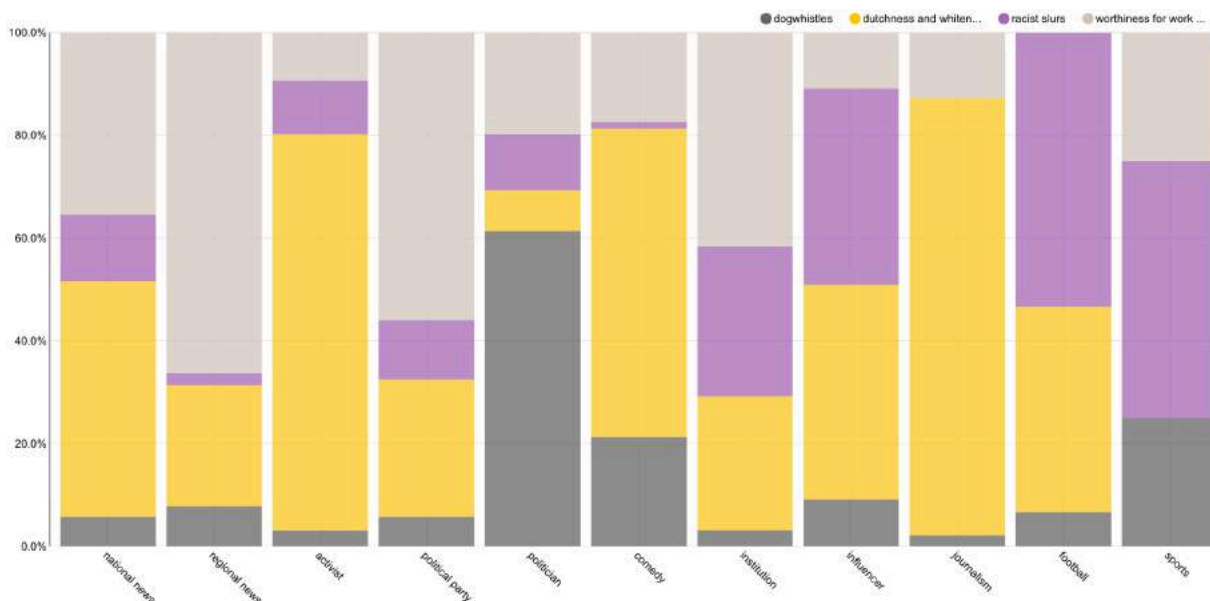
Figure 20: Racism class by post theme



Most of the potentially racist comments on posts relating to immigration fall in the class Worthiness for Work or Asylum. Dutchness and Whiteness is most prevalent for posts about Black Pete - somewhat unsurprisingly - and for posts about racism and discrimination, as well as crime. Posts about Islam attract more comments about Dutchness and Whiteness than the broader theme of immigration. Sexuality and gender get most Dog Whistles, which can be explained by the fact that most Dog Whistle comments are targeting any progressive posts - as they contain the word “woke”, more on this in section 3.5 - and are not necessarily used for racist expression (see 4.5).

A look at comment classes by actor type also reveals interesting patterns (Figure 21). This helps to understand what types of racism are directed to the comment sections of different types of actor.

Fig 21: Comment class by actor



Dog Whistles feature in 61% of likely racist comments for Instagram posts by politicians, and are also more prevalent for sports (25%) and comedy actors (21%) than for others. 85% of racist comments under posts for journalism and 77% for activists are about Dutchness and Whiteness. None of the comments under sports outlet posts are about that, and only comparatively few for politicians (8%). Racist Slurs are most prominent for football (53%), sports (50%), influencers (42%) and other institutions (42%). For Worthiness for Work or Asylum, regional news take the lead (66%), followed by political parties (56%) institutions (42%) and national news (36%). Here again, the difference between comments that political parties and politicians attract on their individual accounts is striking. Individual politicians attract relatively more dog whistles than the accounts of political parties, where comments about Dutchness are relatively more prevalent, this appears to mostly due to individual politicians being called woke. Similarly for sports, it depends on the type - whereas Dutchness and Whiteness comments are more prevalent for football (likely due to racism against Black Dutch soccer players), they are not at all for sports overall.

3.2 Racist Slurs

Closer examination of the comments labelled as Racist Slurs reveals the challenges inherent in classification. This dataset set contains 1940 comments, however, there are many false positives present here. Identifying racist slurs precisely is complicated as the meaning is heavily dependent upon context. This is made more difficult by automated content moderation on Instagram, which is in place to address the most egregious forms of slurs, and thus provides incentives for people to use “alternative slurs”.

One such example would be the term "slave" which is used in a derogatory manner, while some may use the term metaphorically, as observed in expressions like "he's a slave to big oil." This adds a layer of complexity in accurately discerning the intent behind the language, contributing to a substantial number of false positives within this category.

Figure 21 shows that Racist Slurs are primarily concentrated within comment sections of football, sport, and influencer accounts. There are fewer Racist Slurs in comments of national and regional news accounts. This discrepancy may be attributed to the interpersonal nature of Racist Slurs, often targeting specific individuals, such as football players, with a direct intention to insult. This stands in

contrast to other forms of derogatory expressions, discussed below, which are employed for making political statements. We can see this play out in the examples given below. The racist slurs there are used mostly to directly insult a person, and not to make a broader political statement¹².

The comments below are typical examples of the use of Racist Slurs. This section (and the following) will delve into the painful and potentially triggering nature of these comments. Readers are strongly advised to exercise caution and discretion when proceeding through this section.

Original	Translated
<ul style="list-style-type: none"> ● <i>“gA werken n*ger, anders krijg je met de zweep.”</i> 	<ul style="list-style-type: none"> ● <i>“Get a job n*gger, otherwise you’ll get the whip.”</i>
<ul style="list-style-type: none"> ● <i>“volgende keren weer goede n*ger slechte n*ger doen dan?”</i> 	<ul style="list-style-type: none"> ● <i>“Next time we’ll do good n*gger bad n*gger again?”</i>
<ul style="list-style-type: none"> ● <i>“Tering zwarte n*ger”</i> 	<ul style="list-style-type: none"> ● <i>“Fucking black n*gger”</i>
<ul style="list-style-type: none"> ● <i>“Het zijn altijd die kkr zwarte vies vuil aangebrand spul”</i> 	<ul style="list-style-type: none"> ● <i>“It’s always those fucking blacks dirty disgusting burnt stuff”</i>
<ul style="list-style-type: none"> ● <i>“Die n*ger moet zich niet zo aanstellen”</i> 	<ul style="list-style-type: none"> ● <i>“That n*gger shouldn’t be so easily offended”</i>
<ul style="list-style-type: none"> ● <i>“jij bent zeker ook zo’n kutmarrokaan”</i> 	<ul style="list-style-type: none"> ● <i>“You’re probably one of those fucking Moroccans”</i>
<ul style="list-style-type: none"> ● <i>“Kanker n*ger terug naar congo hoerenzoon ziel verkocht voor dollar taboend jimás kk hond”</i> 	<ul style="list-style-type: none"> ● <i>“Fucking n*gger go back to the Congo son of a whore sold your soul for a dollar [taboend jimás?] fucking dog”</i>
<ul style="list-style-type: none"> ● <i>“vieze spleetoog ga jij maar vleermuis eten”</i> 	<ul style="list-style-type: none"> ● <i>“Dirty chink go eat bats”</i>

Most derogatory language specifically targets individuals of African descent. The term "n*ger" (n*gger) is notably frequent, appearing in 65 comments. The overtly malicious nature of these expressions is designed to shock and demean.

By comparison, "kut marokkaan" (fucking Moroccan) and "spleetoog" (chink) both appear only twice throughout the entire dataset. This contrast appears somewhat surprising to anyone following Dutch politics, where most of the ire seems to be aimed at people of Middle Eastern descent.

Two observations emerge from this examination of racial slurs within the dataset. Firstly, the derogatory terms are predominantly directed towards individuals with an intent to demean, the comments generally do not appear to communicate any point beyond this. Secondly, there is a discernible focus on targeting individuals of African descent, as opposed to other minority groups.

3.3 Dutchness and Whiteness

5080 comments in the dataset address themes associated with Dutchness and Whiteness. While there are still false positives within the set, it is noteworthy that their incidence is comparatively lower

¹² A full list of the comments using racist slurs is available upon request. Email bart@sharingperspectivesfoundation.com for more information.

than for Racist Slurs. The distribution of comments on Dutchness and Whiteness across diverse post categories is shown in figure 21. We see that comments in this theme are present across all categories with the exception of sports-related discussions.

That this theme is more present than racial slurs is perhaps unsurprising. First, racial slurs are often deleted by content moderation algorithms. The examples below- typical of what we found in the dataset - show statements that enjoy a wider consensus among the general Dutch public and may be deemed more socially acceptable. They are mostly not direct attacks on individuals based on skin colour, allowing the authors to hide behind the argument that any offence is the responsibility of the offended party being too sensitive. However, we would argue that a discerning audience, guided by reasonableness, would categorise these statements as racist.

Original	Translated
<ul style="list-style-type: none"> • <i>“Door toedoen van die achterlijke Afrikaan, was Jerry Afriyie maar nooit naar Nederland gekomen dan hadden we onze traditie behouden”</i> 	<ul style="list-style-type: none"> • <i>“Because of that retarded African, if only Jerry Afriyie had never come to the Netherland, we would still have our tradition”</i>
<ul style="list-style-type: none"> • <i>“Echte Nederlanders weer.....”</i> 	<ul style="list-style-type: none"> • <i>“Real Dutch people again...”</i>
<ul style="list-style-type: none"> • <i>“Ga terug naar je eigen land als t je niet bevalt hier”</i> 	<ul style="list-style-type: none"> • <i>“Go back to your own country if you don’t like it here”</i>
<ul style="list-style-type: none"> • <i>“Kerk hoort bij Nederland. Moskeeën mogen terug naar hun eigen land”</i> 	<ul style="list-style-type: none"> • <i>“The church is part of the Netherlands. Mosques should return to their own countries”</i>
<ul style="list-style-type: none"> • <i>“Hoe zo Nederlanders”</i> 	<ul style="list-style-type: none"> • <i>“What do you mean Dutch people”</i>
<ul style="list-style-type: none"> • <i>“En nu.... Moet het anders in Nederland? Nederland moet weer Nederland worden eigen volk eerst”</i> 	<ul style="list-style-type: none"> • <i>“And now ... Everything has to change in the Netherlands. The netherlands should become the Netherlands again. Own people first”</i>
<ul style="list-style-type: none"> • <i>“Toen die Marokkanen in Nederland de boel aan het afbreken waren hadden ze zo tegen ze moeten optreden. Nee, met eigen mensen weten ze shit te doen.”</i> 	<ul style="list-style-type: none"> • <i>“When those Moroccans were destroying everything they should have acted like this to them. No, against their own people they know how to do shit.”</i>

There are two types of arguments in these examples. The first revolves around nativism and the criteria defining Dutch identity. Statements such as *"Real Dutch people again..."* assert a categorical exclusion of individuals from Dutchness. This sentiment is reiterated in comments like *"What do you mean Dutch people"* and *"When those Moroccans were destroying everything they should have acted like this to them. No, against their own people they know how to do shit"* The latter example, although not explicitly challenging the Dutchness of individuals, unequivocally posits *"those Moroccans in the Netherlands,"* implying an inherent incompatibility between being of Moroccan descent and having the Dutch identity. Moreover, expressions like *"Go back to your own country if you don't like it here"* explicitly deny the Dutchness of the recipient, insinuating an alternative "land" (country) to which they ostensibly belong.

The second argument centres on perceived foreign elements that supposedly erode essential aspects of Dutch identity. Comments such as *"Because of that retarded African, if only Jerry Afriyie had never come to the Netherland, we would still have our tradition"* are not merely a lamentation over the loss of tradition, but an assimilation of the preservation of said tradition with Dutch identity

("our tradition"). Blame for the tradition's demise is squarely assigned to "that [...] African," emblematic of some discernible external force. The comment "*And now ... Everything has to change in the Netherlands. The Netherlands should become the Netherlands again. Own people first*" conveys a sentiment that a perceived core of Dutchness has been compromised, attributing this dilution to an alternate "volk" (peoples). The prescription for reclaiming Dutch identity involves prioritising "own people first".

3.4 Worthiness for Work or Asylum

4360 comments within the dataset pertain to Worthiness for Work or Asylum. While there are also false positives here, it is noteworthy that these are comparatively lower than observed in the category of Racist Slurs. As we can see in the chart in section 3.1, the theme is present across various post topics, with a pronounced prevalence on posts related to news and politics.

Original

- *"Asielzoekers? Gelukzoekers over onze rug! 🤔"*
- *"Bijna niemand is vluchteling. De meesten zijn illegale migranten die helemaal van niks vluchten en gelukzoekers zijn."*
- *"Haal er nog maar meer naar ons landje. Zijn zulke vriendelijke mensen die asielzoekers. Als je zo met onze gastvrijheid omgaat blijf dan a.u.b weg en blijf in je geboorte land."*
- *"Ik ga lekker stoppen met werken, verdien ik meer als me kapot werken om mijn geld af te staan voor iedereen die geen zin heeft om te werken of voor die kansloze asielzoekers die uit zijn op gratis geld. Eigen volk eerst!!!!"*
- *"Dit krijg je ervan met die linkse rakkers. Ongelooflijk dat dit gebeurt terwijl er zoveel Nederlanders in nood zitten en de echte hulpbehoevende vluchteling nu de dúpe is van dit fiasco. Komt ook nog eens bij dat 80% van die asielzoekers hier niets te zoeken heeft."*
- *"Grenzen dicht gooien weg met dat tuig"*

Translated

- *"Asylum seekers? Fortune seekers at our expense! 🤔"*
- *"Almost no one is a refugee. Most are illegal migrants who run from nothing and are fortune seekers."*
- *"Get more of them to our little nation. Such friendly people those asylum seekers. If this is how you use our hospitality then please stay away and stay in the country of your birth"*
- *"I'm going to stop working, I deserve more than breaking my back and handing my money over to everyone who doesn't want to work or those hopeless asylum seekers who are after free money. Own people first!!!!"*
- *"This is what you get with those leftists. Unbelievable that this happens while so many Dutch people are in distress and the real needy refugee is the victim in this fiasco. On top of that 80% of those asylum seekers have no business here."*
- *"Close the borders and kick out those thugs"*

Two patterns are discernible in these comments. Firstly, a prevailing notion asserts that people arriving in the Netherlands are motivated by motives deemed inappropriate. Characterised as "fortune seekers", these individuals are portrayed as lacking genuine asylum-seeking reasons and, in some instances, are considered detrimental to those deemed authentically in need of refuge.

Commenters use expressions such as "tuig" or thugs are employed in some comments to categorise this group, suggesting a perception of them as undesirable elements.

The second thread revolves around financial considerations, for example: "*Asylum seekers? Fortune seekers at our expense!*". This comment explicitly articulates concern over the financial burden posed by refugees. Whether such apprehensions are grounded in reality or are retrofitted to rationalise pre-existing prejudices is beyond the current analysis. Noteworthy is the prevalence of the term "fortune seeker," featured in 157 distinct comments across the dataset, all of which question the motives ascribed to incoming individuals. Additionally, the term "gratis" emerges organically in numerous comments, despite not being a selected keyword. Found in 62 comments, the majority of which convey negative perspectives, "gratis" serves to accentuate the perceived association between individuals of foreign origin and economic motivations.

Original	Translated
<ul style="list-style-type: none">• “En buitenlanders krijgen gratis hotels en woningen deze wereld is ziek maat”	<ul style="list-style-type: none">• “And foreigners get free hotels and houses this world is sick mate”
<ul style="list-style-type: none">• “En dat volk krijgt hier gratis huis/appartement”	<ul style="list-style-type: none">• “And those people get free houses/appartements here”
<ul style="list-style-type: none">• “Ze krijgen alles gratis en nog ondankbaar en wij nederlanders zitten in de armoe boodschappen niet meer te betalen benzine en diesel zo duur dat we straks niet meer naar ons werk kunnen jesus wat n kutland 😞”	<ul style="list-style-type: none">• “They get everything for free and are still ungrateful while us Dutch people are in poverty and can’t afford the groceries, gasoline and diesel are so expensive that soon we won’t be able to get to work jesus what a shithole 😞”

This small selection gives a decent overview of the general sentiment of the comments that use the word “gratis”.

3.5 Dog Whistles

At 861 comments, Dog Whistles are the least present within the data. Dog Whistles are more rare because they are more difficult to identify. Such expressions are often subtle and identification relies heavily on contextual cues.

Many of the comments in this category are associated with the concept of "wokeness," a term identified through qualitative interviews as a potential indicator of racist undertones. However, most comments that explicitly include the word “woke” express opposition to left-leaning political ideologies, including progressive climate and gender policies. Consequently, these remarks cannot be unequivocally categorised as racist dog whistles. When we exclude comments referencing "woke," the dataset is reduced to 571 comments. It is imperative to acknowledge that the ensuing examples are not necessarily representative of the entire category, as false positives persist within the dataset.

Original

- *“je bent je Nederlands snel afgeleerd zie ik al wel. Komt de aap uit de mouw, of mag ik dat niet zeggen link je jezelf aan apen? Max word nog 5 keer wereldkampioen, lewis is dan al lang vergeten! Geen record blijft er over en er zal een Nederlander de geschiedenis boeken in gaan als beste coureur. Succes met kijken man, echt veel plezier er mee hahahaha”*
- *“Dat is vast en zeker geen aap.”*
- *“Altijd die finnen 🤔”*
- *“De religie van vrede 😄😄😄🙏”*
- *“Zwarte aap jij hebt niet eens recht op te spreken slaafje”*
- *“Boeien het was maar een aap oe oe a a 🙄”*

Translated

- *“You’ve unlearned your Dutch quickly I see. Is the monkey coming out of the sleeve [this is a Dutch saying, roughly equal to “the cat’s out of the bag”], or can I not say that, do you link yourself to monkeys? Max is going to be worldchampions another 5 time, lewis will be long forgotten by then! No record will be left standing and a Dutch man will go down in history as the best driver. Good luck with watching, really have fun hahahaha”*
- *“That surely wont be a monkey”*
- *“Always those Finns 🤔”*
- *“Religion of peace 😄😄😄🙏”*
- *“Black monkey you have no right to speak little slave”*
- *“Who cares it was only a monkey oe oe a a 🙄”*

What we have found is still quite overtly racist, and could justifiably be categorised as Racist Slurs. Many of the comments link black people to monkeys or apes. There are also quite some mentions of the “Fins”. This is used as a coded message indicating that crime was committed by someone of colour.

4. Conclusion

The findings from this social media analysis reconfirm the relevance of The Digital Us. Our examination of 309,083 comments on Instagram posts where racist remarks were expected uncovered a considerable number of such instances. It's important to note that this finding holds true even after considering that some of the comments may use certain terms in a non-racist context, because they are likely balanced out by instances of racism our analysis did not catch. Having 12,000 likely racist comments within this relatively small sample of Instagram posts demonstrates the scale and urgency of the problem. This is especially true because the public comments are likely the tip of the racist iceberg.

The explicitly and severely racist remarks found in individual comments confirms what Build Up has already been arguing for other contexts: content moderation alone is insufficient to address the problem of racism and divisive language online.¹³ The results of this analysis therefore serve as an urgent call to action. Simply ignoring and scrolling past online hatred is the path of least resistance, but it is not a viable path. If we want to live in a society that does not tolerate racist attacks, we need to stand up to racism in the digital space as well. The call to action is to all of us. We cannot leave this responsibility to people who are directly impacted by racism.

Standing up online can feel like a drop in the bucket given the overwhelming volume of racist content. This is why this analysis set out to understand which accounts publish Instagram posts where racist comments are most likely, and what types of racist comments are most prevalent for which themes. The conclusions of this analysis point to recommendations that can guide people on where and how it is most effective to intervene:

- **National news posts attract the biggest volume of racist comments, and themes vary across actors.** We categorised the potentially relevant Instagram accounts into thirteen types of actors, such as politicians, national news outlets, or sports accounts. This categorization helped us analyse the frequency of posts related to specific themes of interest and the level of interaction these posts received. Among all the actors, national news outlets were the most active, receiving the highest number of comments, likes, and overall engagement on their posts. A detailed look at themes for all actors revealed that the themes vary for each actor type.
 - For The Digital Us, this implies that interventions will reach the largest audience if they focus on addressing comments under posts from national news outlets. However, participants might prefer to engage with topics related to their interests – for example, someone keen on sports might choose to interact with comments on sports accounts. This analysis can then guide them on the likely themes associated with their chosen area (for instance, crime-related topics for regional news outlets). Participants could also choose to concentrate their efforts on comments for a specific Instagram account. CESTMOCRO, which posts the most about our themes of interest, could be particularly important for engaging with a large number of commenters.
- **Most posts are about crime, discrimination, protest, Islam and immigration, with varied levels of comments.** The analysis of themes likely to provoke racist comments revealed that crime-related posts are both the most common and the most commented upon. It is noteworthy that posts about Islam rank fourth in terms of frequency, yet they receive the second-highest number of comments. This indicates that the topic of Islam tends to elicit greater reader engagement in the form of comments, and consequently, a higher potential for racist comments.

¹³ See, for example: <https://shorturl.at/gpzC2> or <https://shorturl.at/hjST2>

→ For The Digital Us, this indicates that engaging in discussions about crime or Islam is relevant, as these topics generate considerable debate. However, for genuine and meaningful engagement, it's just as important for participants to get involved in topics that are personally significant to them, even if these subjects are less prevalent.

- **There are racist comments in all defined racism classes.** To understand the frequency of various forms of racism in the comments, we sorted commonly used racist keywords into four groups: Worthiness for Work or Asylum; Dutchness and Whiteness; Dog Whistles; and Racist Slurs. Our detailed analysis revealed that Worthiness for Work or Asylum and Dutchness and Whiteness were the most common, with the fewest “false positives”. This means the keywords in these categories were consistently employed in a racist manner. Fewer comments were categorised as Dog Whistles, which are inherently challenging to analyse due to their nature of concealing racist content behind seemingly harmless words. Nevertheless, within the relevant ones, there were expressions so racist they could be classified as Racist Slurs. Considering that Racist Slurs should ideally be filtered out initially as hate speech, the fact that this analysis still uncovered a significant number of them indicates that this remains an important category to monitor.

→ For The Digital Us, it means that all racism classes need attention. Participants can opt for a specific type of racist expression where they feel they are able to hold conversations best.

- **Each racist comment class has specific discursive patterns.** Racist Slurs are used to belittle individuals and predominantly target those of African descent, with the term "n*gger" being the most frequently used. This finding is somewhat unexpected, considering the Dutch debate often focuses on people of Middle Eastern descent. Comments about Dutchness and Whiteness usually question someone's Dutch identity based on their origin. They revolve around criteria defining a “real Dutch” person, or discuss foreign influences perceived to undermine key elements of an assumed “Dutch identity”. Comments categorised under Worthiness for Work or Asylum display two main trends: they challenge the motives of newcomers to the Netherlands as “inappropriate”, and focus largely on economic aspects, accusing them of freeloading. The most pertinent Dog Whistles related to racism involve likening black people to monkeys, or using the coded term “Fin” to imply that a crime was committed by non-White individuals. Many comments in this category employed the term 'woke', although not in a racist context.

→ The deeply offensive nature of these racist comments necessitates that The Digital Us adopt a cautious strategy. We now understand the patterns within comment threads and can craft tailored responses. This includes addressing claims that asylum seekers exploit social systems, or participating in debates about a supposedly deteriorating - yet vague - “Dutch identity”. Current research indicates that empathy is the most effective strategy for responding to hateful comments.¹⁴ Consequently, a promising approach would involve engaging with the underlying fears expressed in these comments. This will undoubtedly be challenging. The process of analysing these comments has been emotionally taxing, even for research team members who are not personally affected by racism. It's crucial that participants in The Digital Us have access to a supportive environment for debriefing after participating in these discussions.

This analysis provided guidance where and how to intervene in racist comments so that this labour is actually effective. The Digital Us offers a platform to prepare for and do this very necessary but hard work in solidarity with other people who feel connected to the Dutch context and care about racism in the digital space. We look forward to getting started.

¹⁴ ETH Zurich. 2021. Empathy-based counter speech can reduce hate speech. <https://shorturl.at/qvFV3>