

ANNOTATION RULES AND CODEBOOK

Paper: #WhoAmI in 160 Characters? Classifying Social Identities Based on Twitter Profile Descriptions

Accepted at NLPCSS Workshop (EMNLP 2016)

HOW TO CODE

1 = the profile fits the identity category and the “Not applicable” category

0 = the profile does not fit the identity category and the “Not applicable” category

Use the “Notes” column to annotate any further clarification.

GENERAL RULES

1. What to annotate: **Twitter profiles descriptions, ScreenName and DisplayName** only. The last two features (ScreenName and DisplayName) are useful to identify non-individual users (see point 5). Annotating the profile description makes the focus on social identity expressed on Twitter, that is, how people describe themselves in terms of roles/memberships on Twitter only.
 - a. **Include profiles with emoticons** and make an evaluation. If no decision can be made, code as “Not applicable”.
 - b. **Blank profiles are excluded** from the annotation dataset.
2. **Three datasets:** random tweets from 2015, Movember 2013, Movember 2014.
3. **Language: English** (feature)
4. **Country: US, UK and AUS** (country classifier)
5. **Include only profiles of individuals.** If the profile is not an individual, code as “Not applicable” and specify the type of user in brackets.

SPECIFIC RULES

Follow the social identity classification (Table 1) to annotate the profiles. Some general remarks:

- Keep your concentration while annotating: it is easy to make typos or to lose attention after annotating for a while. Errors due to a lack of attention have to be avoided as they may affect reliability.
- Read the syntax of the Tweet and take the text literally; do not think for too long about the meaning of words.

Table 1. Social identity classification: categories, definitions and examples.

Category of social identity	Definition	Dimension	Examples
1. Relational	<p><i>Relationship identities</i> are self-definitions based on relationships and social roles played by the individual. These identities are both</p> <ul style="list-style-type: none"> Means of categorizing oneself with <u>others who are similar</u>. They target social groups (from <i>Social identity theory</i>/Self – categorization theory) A way of defining a <u>relationship</u> to one specific other person. They target a relation or role (from <i>Identity theory</i>) 	Categorization with others who are similar/have the same social role	I am a Woman, Man, Teenager, Child
		Categorization according to relational social roles (family, friends, love)	I am a Mother Father Son Daughter Sister Brother, Grandmother Grandfather Uncle Aunt, Friend, Girlfriend, Boyfriend, Wife, Husband , Widow, Divorced Person ! abbreviation (bro, sis)
		Identification/not identification/rejection with another person/other people	@MattBellamy is perfection @MattBellamy is magnificent not really sure if i am barbra streisand or celine dion
		Real/unreal relationships and uni/bidirectional relationships with someone/group of people	My relationship with my (supervisor/coworker/co-founder-spouse) is... I love my family, mother, husband, boss, ... @Born2KillH2O is my Waterboy.
		Being a fan of a person/group of people	I am a fan of Justin Bieber I love Harry Styles OneDirection are my life I support/have an interest in LFC (football club)
		Being part of a collective (association/team/group/club etc.) where the stress is on the relation that the individual has with the collective	I love Football Club Twente! It is my family! TEAM BREEZY DreamTeam Team Douwe Bob
		Tagging someone/group of people	@Onedirection 4life?? @Real_Liam_Payne @Louis_Tomlinson @harry_styles @zaynmalik & @Niallofficial
		Mentioning someone/group of people	Khleo Thomas Justin Bieber RIP ...

2. Occupational, vocational	<i>Occupational identities</i> relate to the occupation and profession of the individual.	Occupations, career Identification with an organization that can be related to work (e.g., university)	I am Intellectual Military veteran Student Collector Musician Gardener Teacher Supervisor Secretary Scientist Psychologist Salesperson Business person Athlete I am/Graduate from UTwente
	<i>Occupational identities</i> also reflect individual vocations/avocations. <i>Note from annotators' meeting:</i> It is not easy to understand whether writing "musician, gamer, YouTuber, etc." is actually consider an occupation/profession or not. Therefore, occupational identities also deal with vocation/avocation, hobbies and interests can fit in this category because people spend time in doing what they like, even if they are not paid for it (as for occupations or professions)	Interests and hobbies Being <u>fan</u> of something Loving/liking doing something	I am a Gamer, YouTuber, musician amateur yogi avid pop culture consumer a big fan of pizza, olives, animals I am TV/food/animals lover/fan/enthusiast I love gaming, playing videogames, cooking, baking, travelling, researching
3. Political	<i>Political identities</i> reflect political affiliation, being member of/play role in of a social movement, being politically active, being member of/play role in a collective (association/team/group/club etc.) that is based on political ideology/party or movement	Self-statement on political affiliations, parties, social movements, collectives	I am ... Feminist Political Independent Democrat Republican, conservative, democrat, environmentalist, feminist, liberal, pacifist, voter mo bro, mo sis LIVESTRONG FOUNDATION I'm also a council candidate in local elections for UKIP. I'm a union rep and H&S rep for my job. occasional UN ambassador

		Movement's hashtags	<p>Welsh rep @CoopPartyYouth, @WelshYoungLab & @CoopParty Community Cllr #Tonyrefail, @UniteWales rep = representative</p> <p>No Tolerance 4 Socialist Libtards.</p> <p>YES voter & SNP member</p> <p>#movember #BlackLivesMatter #UniteRight DIY Hack</p> <p>#ORPUW, #TUGC, #NoAmnesty</p>	
4. Ethnic, religious	<i>Ethnic/religious identities</i> refer to self-definitions based on memberships of ethnic/religious groups. it is also valid for being member of/play role in a collective (association/team/group/club etc.) that is based on religion/ethnicity	Self-statement on ethnicity and religious, and related affiliations, movements, collectives	<p>I am ... Jewish Christian Catholic Southerner New Yorker American Hispanic Asian-American African-American, White, Black</p> <p>Native washingtonian US expat Nordic Down Under = Australian</p> <p>fluent in Grade A chilling hardwork and Espanol</p> <p>GodFirst</p> <p>Godless</p>	
	Stating the name of a country or a town is a symbol of sense of membership/belonging. Therefore, it is <i>ethnic identity</i>	Names of towns, countries	<p>London. Italy. in west midlands Pacific island Singapore notes Espanol India Florida Seattle</p>	<p>native washingtonian North Melbourne Loughton London Welsh rugby Doncaster England Rogers, MN Western australia</p>
5. Stigmatized	<i>Stigmatized identities</i> refer to people who belong to stigmatized groups. In sociological theory, a stigma is a deeply discrediting feature in an individual, although this is largely dependent on groups. In other words, a stigma is a mainly negative label attached to people who are considered different from what the society consider normal at the level of social and cultural norms.		<p>Unemployed Homeless Fat, Looking to lose it again with clean eating & exercise Person with AIDS Lesbian, Gay, Queer</p>	<p>addicted/dependent to something/addiction geek cynic grumpy old git hipster</p>

	Typical stigmatized groups are fat people, alcoholics, homosexuals. However, in contemporary society the concept of stigma has changed. Some stigmas have disappeared or been accepted a normal in many societies. New stigmas have emerged (e.g., people with ADHD). The practice of self-stigma is often common, that is, people attach a stigma to themselves (sometimes with an ironic tone).	Heavy Drinker Smoker Alcoholic mentally ill Idiot savant Junkie (E.g. food junkie) Workaholic Snob Nigga	Breast cancer survivor social awkwardness Legal alien nerd Gay McYaoison damaged people Social anxiety Fitfam Arseglow
Not applicable	Quote/citations/values is <i>value identity</i> from religious book (Bible, Coran, Psalm, ...) related to ethnicity	Love for ALL animals- consciousness resides in all beings- respect it “Blessed are you when people insult you, persecute you and falsely say all kinds of evil against you because of me. (Matthew 5:11 NIV) Love your country, but never trust its Government. - R. A. Heinlein	
	Self-promotion/Links to other social media accounts	instagram ashp8 follow me...	
	Individual attributes	procrastinator, cynic, lame, boring	
	Interest/love for SOMETHING	I love food, music, tv, animals, I love food..... I love coffee..... I love chocolate..... and I have a license to love DRUGS!	
	Ambiguous cases for which we cannot define which category they belong to	Brisbane Lions Football Club #LakerNation #Lakers4Life #Lakers Kik im a goddamn robot Does it come in black? barb. 14 y/o. bangtan (esp seokjin) af. i follow back, just ask! nl YES	
	Incomprehensible sentences for which no decision can be made	who gives a fuck. imma siner 315 / 19 / IG-katrynae /❤️❏	