

1. Overview

- Over 11% of all tweets posted in Twitter are deleted
- **Is tweet deletion equally prevalent among all Twitter users?**
- **Do tweets, that are deleted later, have distinctive properties?**
- Study of 200 thousand Twitter users, over a period of four weeks
- Personality based characterization of user differences
- Function word usage based characterization of tweet differences

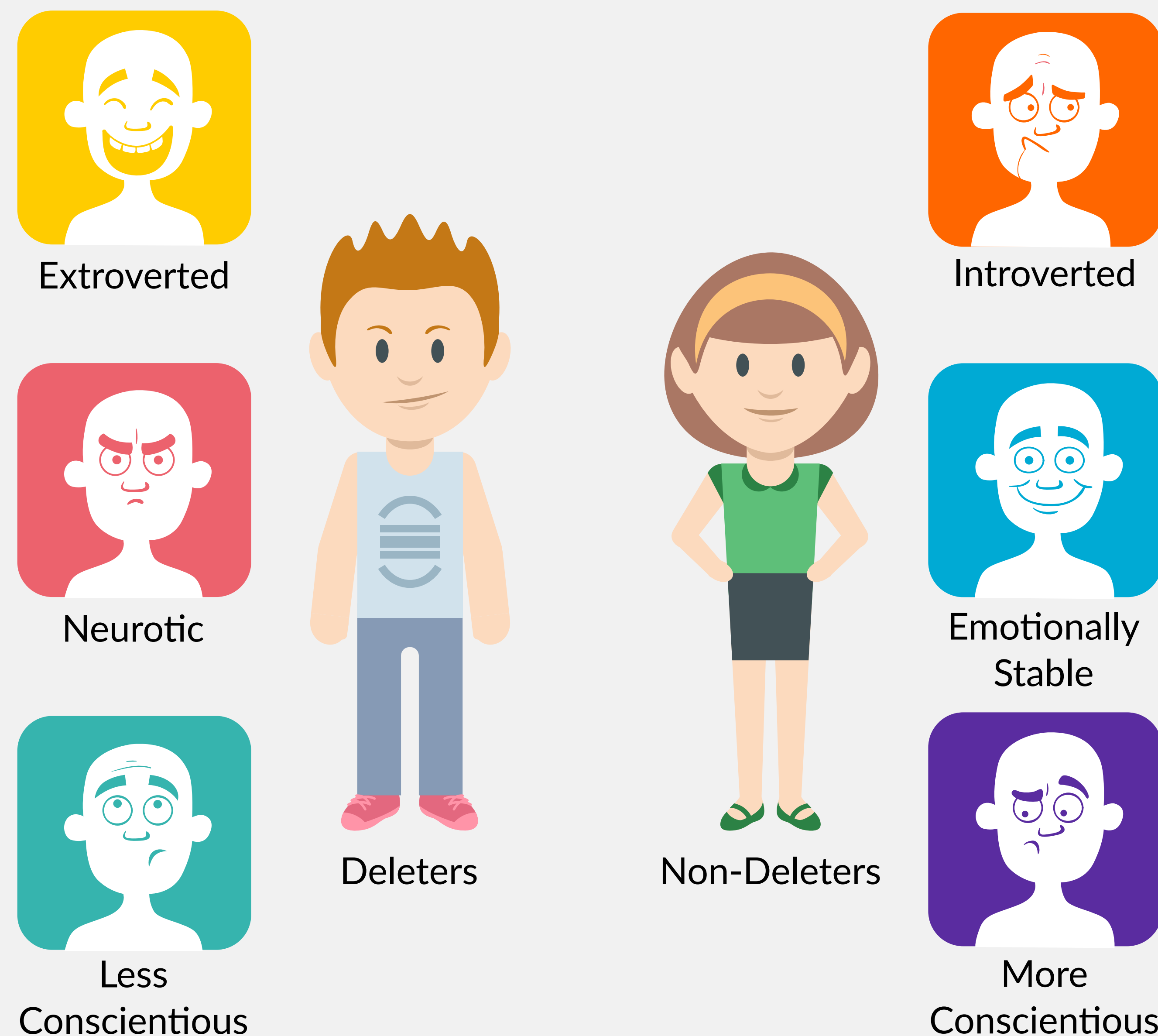
2. Motivation & Challenges

- Understanding deletion is fundamental for privacy aware systems
- Deletion is a personal choice and depends heavily on user
- Difficult for third party to understand why a tweet was deleted
- Lack of tweet editing feature major cause of superficial deletions
- Careful user selection to avoid spammers and bot accounts
- Extensive data cleanup to remove automated tweets and retweets

3. Characterizing User Differences

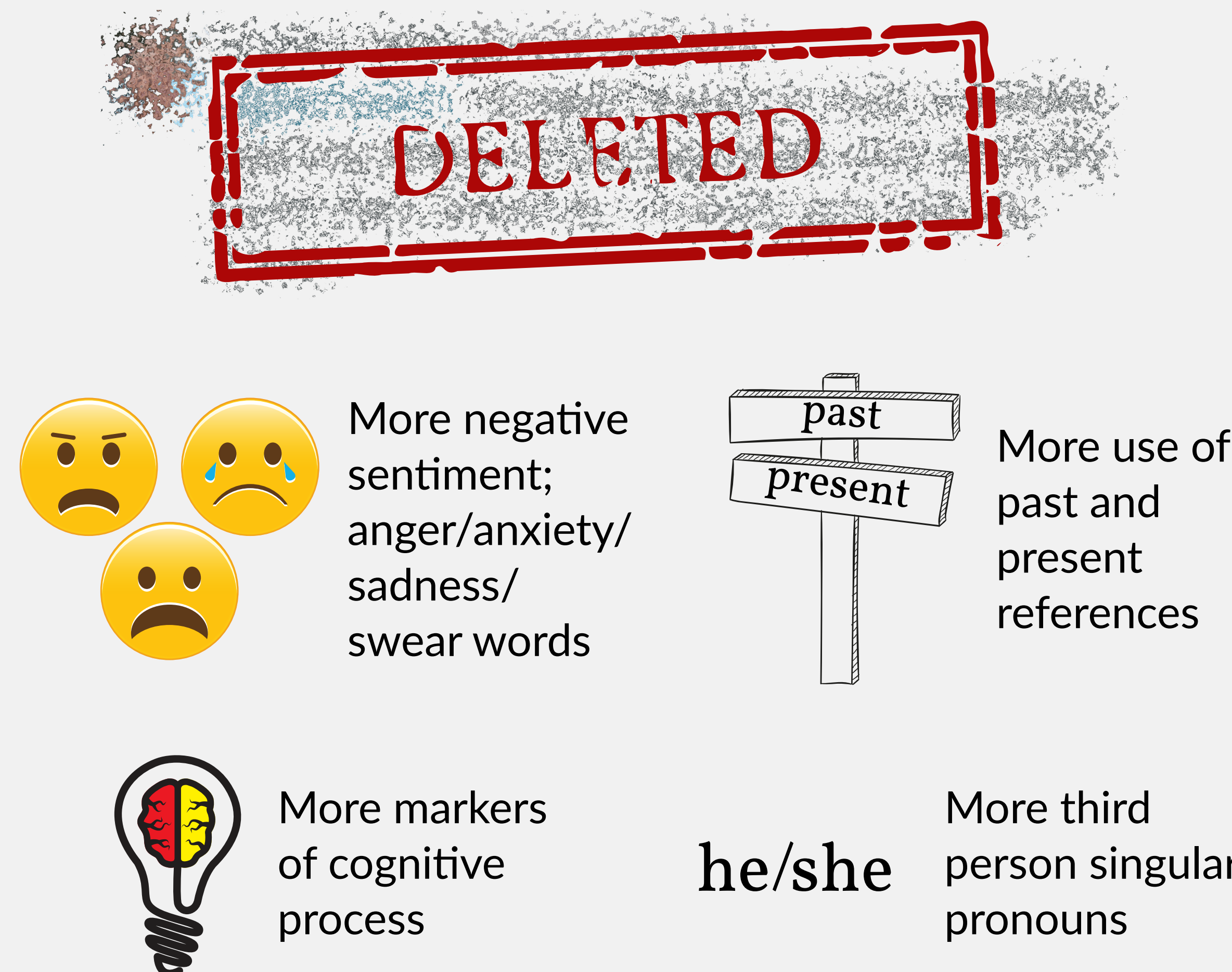
Big Five Personality Traits

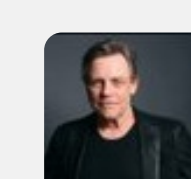
- **Openness:** open attitude towards new experiences and diverse ideas
- **Conscientiousness:** high degree of self-discipline, high achievers, meticulous planners
- **Extraversion:** highly social, generally friendly, energetic in social situations
- **Agreeableness:** optimistic outlook, preference towards maintaining social harmony
- **Neuroticism:** significant mood fluctuations, high propensity to feel negative emotions

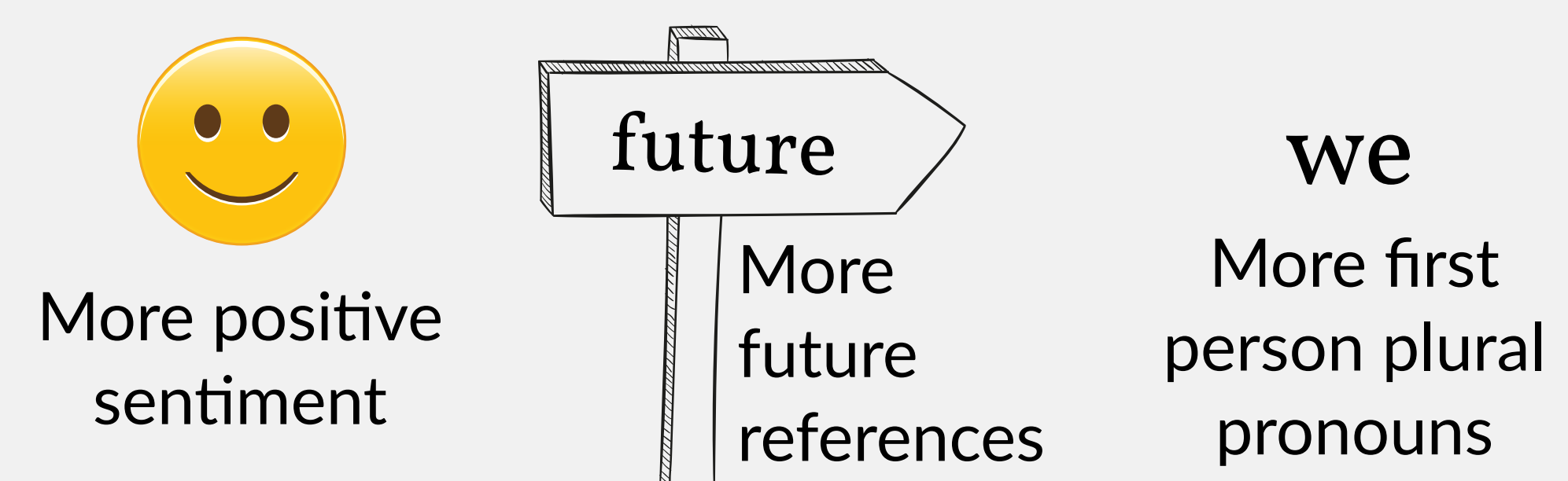


- Earlier works have shown strong and significant correlations between social characteristics [1] and linguistic style [2] with a user's personality
- We utilized the above reported strong and significant correlations to predict user personality
- **Deleters:** 92 thousand users, who had made at-least one non-superficial tweet deletion
- **Non-Deleters:** 102 thousand users, for whom all deletions (if any) were classified as superficial deletions

4. Characterizing Tweet Differences



 **Mark Hamill** @HamillHimself
Hey #Hamsters-When I make the effort to answer your questions please don't delete UR tweet so my answer makes no sense-Thank you #TimeWaster



- We used LIWC toolkit to analyze different function word usage in deleted and non-deleted tweets from deleters
- Dataset had 1.2 million tweets that were deleted, and 15.9 million non-deleted tweets

5. Key Insights

- To delete one's posted tweet is a very personal decision, and depends on the author's personality and writing style.
- Signals about an author's personality can be obtained from her deletion practices, even if number of deletions are low.
- Deleted tweets have distinctive features, although it is difficult for a third party to understand the reasons behind their deletion.
- Understanding post deletion is important for any privacy aware system that attempts to help its users not leak information.

6. References

- Quercia et al. "Our Twitter Profiles, Our Selves: Predicting Personality with Twitter", Proc. IEEE PASSAT/SocialCom, 2011
- Golbeck et al. "Predicting Personality from Twitter" Proc. IEEE PASSAT/SocialCom, 2011

Contact

CNeRG.org - Complex Networks Research Group
Email: parantapa@cse.iitkgp.ernet.in