Detecting Fake News Spreaders in Social Networks via Linguistic and Personality Features

Notebook for PAN at CLEF 2020

Anu Shrestha, Francesca Spezzano, and Abishai Joy

Computer Science Department
Boise State University
{anushrestha, abishaijoy}@u.boisestate.edu
francescaspezzano@boisestate.edu

Abstract This paper addresses the problem of automatically detecting fake news spreaders in social networks such as Twitter. We model the problem as a binary classification task and consider several groups of features, including writing style, word and char n-grams, BERT semantic embedding, and sentiment analysis, which are computed from a set of tweets each user authored. Our proposed approach is evaluated on the dataset made available by the PAN at CLEF 2020 shared task on profiling fake news spreader, which provided labeled data in both English and Spanish. Experimental results show that we can detect fake news spreaders with an accuracy of 0.73 in English and 0.77 in Spanish when our approach is evaluated with 10-fold cross-validation on the provided training set, and with an accuracy of 0.71 in English and 0.76 in Spanish when the model is trained on the whole training set and tested on the provided test set. We also investigate the role of psycho-linguistic (LIWC) and personality features to detect fake news spreaders and find out that personality features do have a significant impact in user sharing behavior, achieving an accuracy of 0.72 in English and 0.80 in Spanish when evaluated with 10-fold cross-validation on the provided training set.

CEUR-WS.org/Vol-2696/paper_218.pdf

1 Introduction

Fake news sharing has become a concerning problem in online social networks in recent years. Research has found that fake news is more likely to go viral than real news, spreading both faster and wider [23] and is threatening public health [22], emergency management and response [21,6], election outcomes [9], and is responsible for a general decline in trust that citizens of democratic societies have for online platforms [1]. Surprisingly, bots are equally responsible for spreading real and fake news, and the considerable spread of fake news on Twitter is caused by human activity [23]. Fake news is successful mainly because people are not able to disguise it from truthful information [12,7] and often share news online without even reading its content [3]. Also, even if people recognize news as fake, they are more likely to share it if they have seen it repeatedly than the news that is novel [2].

Copyright © 2020 for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0). CLEF 2020, 22-25 September 2020, Thessaloniki, Greece.

Thus, being able to identify fake news spreaders in social networks is one of the key aspects to effectively mitigate misinformation spread. Examples of strategies that could be implemented include assisting fake news spreader with credibility indicators to lower their fake news sharing intent [24], and mitigation campaign, e.g., target the most influential real news spreader to maximize the spread of real news [19].

This paper describes the approach we implemented and submitted to profiling fake news spreaders on Twitter as part of the PAN at CLEF 2020 shared task [16]. Specifically, we propose a machine-learning based approach that considers several groups of features, including features capturing the Twitter writing style, words and chars n-grams, the BERT semantic embedding of the tweets, and the sentiment expressed in the tweets. Experimental results show that our approach can detect fake news spreaders with an accuracy of 0.73 on the English dataset and of 0.77 on the Spanish dataset when evaluated with ten-fold cross-validation on the training set while achieving an accuracy of 0.71 for English and 0.76 for Spanish when evaluated on the provided test set.

Furthermore, the paper also investigates the role of psycho-linguistic (LIWC) and personality features (including Big Five personality traits, needs, and values) in detecting fake news spreaders. Our additional experimental results on the provided training set show that we can achieve an accuracy of 0.72 with LIWC or personality features on the English dataset, while personality features achieve an accuracy of 0.80 on the Spanish dataset, outperforming our submitted approach on this specific language. These results suggest that psycho-linguistic and personality features are valuable characteristics to consider when profiling fake news spreaders in social media.

2 Related Work

Several studies have been conducted to understand the characteristics of regular (non-malicious) users that correlate with fake news spreading behavior in social networks. Vosoughi et al. [23] revealed that the users responsible for fake news spread had, on average, significantly fewer followers, followed significantly fewer people, and were significantly less active on Twitter. Shrestha and Spezzano showed that social network properties help in identifying active fake news spreaders [18]. Shu et al. [20] analyzed user profiles to understand the characteristics of users that are likely to trust/distrust fake news. They found that, in average, users who share fake news tend to be registered for a shorter time than the ones who share real news and that bots are more likely to post a piece of fake news than a real one, even though, users who spread fake news are still more likely to be humans than bots. They also show that real news spreaders are more likely to be more popular and that older people and females are more likely to spread fake news. Guess et al. [5] also analyzed user demographics as predictors of fake news sharing on Facebook and found out political orientation, age, and social media usage to be the most relevant. Specifically, people are more likely to share articles they agree with (e.g., right-leaning people tended to share more fake news because the majority of the fake news considered in the study were from 2016 and pro-Trump), seniors tend to share more fake news probably because they lack digital media literacy skills that are necessary to assess online news truthfulness, and the more people post on social media, the less they are likely to share fake news, most likely because they

are familiar with the platform and they know what they share. Yaqub et al. [24] analyzed open-ended responses where users who participated in the study explained the rationale behind their sharing intent of true, false, and satire headlines. Among the most frequent motivations for sharing/not sharing news there are (1) the interest/non-interest towards the news, (2) the potential of generating discussion among the friends, (3) the fact that the news is not relevant to the user's life, and (4) the perceived news credibility, especially as a motivation for not sharing news. Giachanou et al [4] addressed the problem of discriminating between fake news spreaders and fact-checkers (a problem slightly different from the one addressed in this paper) and proposed an approach based on a convolutional neural network to process the user Twitter feed in combination with features representing user personality traits and linguistic patterns used in their tweets. Beyond user and news characteristics, Ma et al. [11] also analyzed the characteristics of diffusion networks to explain users' news sharing behavior. They found opinion leadership, news preference, and tie strength to be the most important factors at predicting news sharing, while homophily hampered news sharing in users' local networks. Also, people driven by gratifications of information seeking, socializing, and status-seeking were more likely to share news in social media platforms [10].

3 Dataset

We carried out our experiments on the dataset provided by the PAN'20 shared task for Profiling Fake News Spreaders on Twitter [16]. This dataset contains a balanced set of users along with their Twitter feed and known ground truth about the users. Specifically, users that shared some fake news in the past are labeled as fake news spreader and real news spreader otherwise. The dataset has been collected in two languages, namely English and Spanish, and consists of a train and a test set. For each considered language, the training set includes 300 users with 100 tweets for each user resulting in 30,000 English tweets and Spanish 30,000 tweets. The test set contains data for 200 users in each language. As recommended by the shared task, the English and Spanish datasets have been treated separately in our experiments.

4 Features for Detecting Fake News Spreaders

This section presents the features we considered for detecting fake news spreaders. As the dataset files are in raw XML format, we parsed and formatted them by using the xml.etree.ElementTree[1] library. After extraction, we pre-processed the content (user tweets) as per requirement for each feature we considered as there are some features that require cleaned texts and some other that require the text as it is for incorporating underlying details of the author's writing. Basically, four different types of features were considered to address this task, as explained in detail below.

Style. This first set of features captures the writing style of the set of tweets authored by the same user. Specifically, we computed the average number of certain words, items,

[1] `https://docs.python.org/3/library/xml.etree.elementtree.html`

and characters per user tweet, which includes the average number of (1) words, (2) characters, (3) lowercase words, (4) uppercase words, (5) lowercase characters, (6) uppercase characters, (7) stop words, (8) punctuation symbols, (9) hashtags, (10) URLs, (11) mentions, and (12) emojis and smileys. Also, we considered the (13) percentage of user tweets that are a retweet and (14) the percentage of user tweets that are a sharing of breaking news.

N-grams. The second group of features includes TF-IDF based n-grams for both words and characters. We concatenated all the tweets by each user to form a single document per user. For each document, we removed words like 'RT,' 'Via,' and '&amp' and replaced emojis and smileys with the corresponding English words. Next, each document was converted to lowercase, stop words and punctuation symbols were removed, and the remaining words were lemmatized into root words the using the NLTK toolkit.[2] Finally, we computed the TF-IDF vector representation for both word and char n-grams.[3] We experimented with different parameters, including the number of terms and the length of grams, from uni-grams to tri-grams. We obtained the best results for both chars and words with uni-grams and by including all the terms (max_df = 1.0).

Tweet embedding. The third group of features includes the embedding of tweets as computed by using the BERT state-of-art NLP model. Specifically, we used the pre-trained multilingual model provided by SBERT [17] to address both languages. The extracted embedding was reduced to 10 features via principal component analysis (PCA)[4]. Then, we averaged the embedding of all the tweets of a single user to generate a single embedding representation per user that captures the semantic of all the user tweets. Before computing this embedding, each tweet was pre-processed to remove frequently Twitter used characters like 'RT,' 'Via,' and '&amp' and replace emojis and smileys with the corresponding English words.

Sentiment analysis. As people express their emotions, appraisals, and sentiments towards any news or article through the choice of words in their tweets, we leveraged sentiment analysis as another feature. Also in this case, we pre-processed each tweet to remove 'RT,' 'Via,' and '&amp.' However, we did not replace emojis and smileys with the corresponding English text for this feature since these emojis adds to the sentimental values of the text. Then for each user, we measured the average sentiment across all their tweets. We used the Valence Aware Dictionary and sEntiment Reasoner (VADER) [8], a library specifically built for capturing sentiments expressed in social media texts, for English tweets (compound value) and sentiment-analysis-spanish [5] for Spanish tweets.

---

[2] https://www.nltk.org/

[3] https://scikit-learn.org/stable/modules/generated/sklearn.feature_extraction.text.TfidfVectorizer.html

[4] https://scikit-learn.org/stable/modules/generated/sklearn.decomposition.PCA.html

[5] https://pypi.org/project/sentiment-analysis-spanish/

Table 1: Accuracy comparison of each group of features form Section 4 as input to four different classifiers, namely Random Forest (RF), Logistic Regression (LR), Support Vector Machine (SVM), and Extra Trees (ET). For each group of features, the best value is bolded.

(a) English dataset

Feature	Classifier	Accuracy
Style	RF	0.62
Style	LR	0.59
Style	SVM	0.63
Style	**ET**	**0.64**
N-grams	RF	0.71
N-grams	LR	0.71
N-grams	**SVM**	**0.72**
N-grams	ET	0.70
Tweet Embedding	RF	0.67
Tweet Embedding	LR	0.68
Tweet Embedding	**SVM**	**0.69**
Tweet Embedding	ET	0.66
Sentiment Analysis	RF	0.56
Sentiment Analysis	**LR**	**0.66**
Sentiment Analysis	SVM	0.64
Sentiment Analysis	ET	0.56

(b) Spanish dataset

Feature	Classifier	Accuracy
Style	RF	0.73
Style	LR	0.71
Style	**SVM**	**0.75**
Style	ET	0.70
N-grams	RF	0.73
N-grams	LR	0.74
N-grams	SVM	0.74
N-grams	**ET**	**0.75**
Tweet Embedding	**RF**	**0.74**
Tweet Embedding	LR	0.70
Tweet Embedding	SVM	0.73
Tweet Embedding	ET	0.73
Sentiment Analysis	RF	0.51
Sentiment Analysis	**LR**	**0.57**
Sentiment Analysis	SVM	0.55
Sentiment Analysis	ET	0.51

5 Experimental Results

We built our classification model as an ensemble of classifiers. Specifically, we considered each group of features separately and, by performing 10-fold cross-validation on the training set, we chose the best classifier for that group of features among Support Vector Machine (SVM) with linear kernel, Logistic Regression (with default parameter), Random Forest and Extra Trees (both with 500 estimators and minimum sample leaf parameter equal to 1).

According to the results shown in Tables 1a and 1b, for English, we chose Extra Trees as the best classifier for the style features, SVM for both n-grams and tweet embedding, and Logistic Regression for sentiment analysis. Likewise, for Spanish, we chose SVM as the best classifier for the style features, Extra Trees for n-grams, Random Forest for tweet embedding, and Logistic Regression for sentiment analysis.

We see that, for English Twitter users, the best performing set of features is n-grams (accuracy of 0.72), followed by BERT tweet embedding (accuracy of 0.69), sentiment (accuracy of 0.66), and style-based features (accuracy of 0.64). For Spanish Twitter users, style-based features and n-grams are equally the best sets of features (both with an accuracy of 0.75), followed by BERT tweet embedding (accuracy of 0.74), sentiment (accuracy of 0.57). Overall, the different performance of the features across English and Spanish may highlight a different cultural behavior in the use of Twitter in general, and in sharing news in particular.

Table 2: Accuracy values obtained by the final model.

Evaluation	Accuracy		
	English	Spanish	Average
Training Set (10-fold cross-validation)	0.73	0.77	0.75
Test Set	0.71	0.76	0.73

For the final model, we first trained each group of features with the corresponding selected best classifier from Tables 1a and 1b. Then, we used the prediction probabilities of all the four selected best classifiers as input features to a majority voting classifier that combined the predictions of four base estimators, namely SVM, Logistic Regression, Random Forest, and Extra Trees. Table 2 reports the resulting accuracy values of the final majority voting classifier in two cases. First, when we evaluate our model with 10-fold cross-validation on the training set (i.e., we trained on 90% of the training set and tested on the remaining 10%, repeated the experiment 10 times and averaged the results), we achieve an accuracy of 0.73 and 0.77 for the English and Spanish dataset, respectively. As we can see, the accuracy value of the features combined by our final model improves over the accuracy values achieved by each group of features individually, as reported in Tables 1a and 1b. Second, when we train on the whole training set and test on the test set, we achieve an accuracy of 0.71 for English and 0.76 for Spanish (run executed on TIRA [15]).

6 Investigating the Role of Psycho-linguistic and Personality Features in Detecting Fake News Spreaders

After the submission of our run, we performed some additional experiments to investigate how other groups of features, such as psycho-linguistic and personality features, would perform at detecting fake news spreaders. Specifically, we considered the set of psycho-linguistic features computed by the Linguistic Inquiry and Word Count (LIWC) tool and personality features extracted by using the IBM Watson Personality Insights service[6]. To compute these two sets of features, tweets were pre-processed in the same way as for the *tweet embedding* features from Section 4. These features are described in detail here below.

Linguistic Inquiry and Word Count (LIWC). LIWC is a transparent text analysis tool that counts words in psychologically meaningful categories. The tool works for different languages, including English and Spanish. We used LIWC2015 for English (93 features) and LIWC2007 for Spanish (90 features). LIWC computes different measures for analyzing the cognitive, affective, and grammatical processes in the text. The LIWC features can be divided into four main categories [14]:

Linguistics features refer to features that represent the functionality of text, such as the average number of words per sentence and the rate of misspelling. This cate-

[6] `https://cloud.ibm.com/apidocs/personality-insights`

gory of features also includes negations as well as part-of-speech (Adjective, Noun, Verb, Conjunction) frequencies.

Punctuation features include the occurrences of Periods, Commas, Question, Exclamation, and Quotation marks, etc. in the text.

Psychological features target emotional, social process, and cognitive processes. The affective processes (positive and negative emotions), social processes, cognitive processes, perceptual processes, biological processes, time orientations, relativity, personal concerns, and informal language (swear words, nonfluencies) can be used to scrutinize the emotional part of the text.

Summary features define the frequency of words that reflect the thoughts, perspective, and honesty of the writer. It consists of Analytical thinking, Clout, Authenticity, Emotional tone, Words per sentence, Words more than six letters, and Dictionary words under this category.

To compute the LIWC feature set for each user in the dataset, we first computed the LIWC features for each tweet and then averaged the features for the same user.

Personality Features. The IBM Watson Personality Insights service uses linguistic analytics to infer individuals' intrinsic personality characteristics, including Big Five personality traits, Needs, and Values, from digital communications such as social media posts. The tool is able to work for different languages, including English and Spanish. In our case, we concatenated all the tweets of a given user in a unique document to compute their personality characteristics. The features computed by this service are detailed in the following (we considered the raw scores provided by the service):

Big Five The Big Five personality traits, also known as the five-factor model (FFM) and the OCEAN model, are a widely used taxonomy to describe people's personality traits [13]. The five basic personality dimensions described by this taxonomy are openness to experience, conscientiousness, extraversion, agreeableness, and neuroticism. For each personality dimension, IBM Watson Personality Insights also provides a set of additional six facet features. For instance, agreeableness' facets include altruism, cooperation, modesty, morality, sympathy, and trust.

Needs These features describe the needs of a user as inferred by the text they wrote and include excitement, harmony, curiosity, ideal, closeness, self-expression, liberty, love, practicality, stability, challenge, and structure.

Values These features describe the motivating factors that influence a person's decision making. They include self-transcendence, conservation, hedonism, self-enhancement, and open to change.

Tables 3a and 3b report the accuracy achieved by the four considered classifiers with input LIWC and personality features for detecting fake news spreaders. We were able to evaluate these features only on the provided training set and reported results are the averaged values of 10-fold cross-validation. As we can see, for the English dataset, LIWC and personality features achieve both the best accuracy of 0.72 with Random Forest, which is the same as the accuracy achieved by n-grams features (cf. Table 1a) and slightly lower than the accuracy of 0.73 achieved by our final submitted approach on the training set (cf. Table 2). In the case of the Spanish dataset, LIWC achieves the

Table 3: Accuracy comparison of LIWC and Personality features form Section 6 as input to four different classifiers, namely Random Forest (RF), Logistic Regression (LR), Support Vector Machine (SVM), and Extra Trees (ET). For each group of features, the best value is bolded.

(a) English dataset

Feature	Classifier	Accuracy
LIWC	**RF**	**0.72**
LIWC	LR	0.63
LIWC	SVM	0.61
LIWC	ET	0.70
Personality	**RF**	**0.72**
Personality	LR	0.70
Personality	SVM	0.71
Personality	ET	0.70

(b) Spanish dataset

Feature	Classifier	Accuracy
LIWC	**RF**	**0.78**
LIWC	LR	0.73
LIWC	SVM	0.71
LIWC	ET	0.77
Personality	RF	0.77
Personality	LR	0.70
Personality	SVM	0.68
Personality	**ET**	**0.80**

best accuracy of 0.78 with Random Forest, while personality features perform with the best accuracy of 0.80 with Extra Trees. In this case, both LIWC and personality features outperform our submitted approach on the training set, which achieves an accuracy of 0.77 (cf. Table 2), with personality features having the best ever achieved accuracy among all the group of features we tried for this shared task.

7 Conclusions

We addressed the problem of automatically detecting Twitter users keen at spreading fake news as part of the PAN at CLEF 2020 shared task on profiling fake news spreaders in two languages, namely English and Spanish. We proposed a first approach leveraging several groups of features, including writing style, word and char n-grams, BERT semantic embedding, and sentiment analysis, which are computed from the Twitter feed provided for each user. This approach achieved an accuracy of 0.73 in English, resp. 0.77 in Spanish, when evaluated with 10-fold cross-validation on the provided training set, and an accuracy of 0.71 in English, resp. 0.76 in Spanish, when evaluated on the provided test set. We also investigated the role of psycho-linguistic (LIWC) and personality features on the same task. We showed that personality traits are important characteristics to consider when modeling user sharing behavior, as they achieved an accuracy of 0.72 in English, resp. 0.80 in Spanish, when evaluated with 10-fold cross-validation on the provided training set. Overall, the task of detecting fake news spreaders turned out to be very challenging. Even if we tried several sets of features, we could not achieve an accuracy value higher than 0.80. One possible motivation could be that some users keen to spread fake news do not do it intentionally; hence they are hard to differentiate from users who never shared fake news. Also, the accuracy gap between English and Spanish may indicate users have different news spreading behaviors across different cultures.

Future work will be devoted to analyzing the role of additional sets of features for detecting fake news spreaders in social networks, including behavioral features describing the user activity on Twitter and social network features.

Acknowledgements

This work has been partially supported by the National Science Foundation under Awards no. 1943370 and 1820685.

References

1. Barometer, E.T.: Edelman trust barometer global report. Edelman, available at: https://www. edelman. com/sites/g/files/aatuss191/files/2019-02/2019_Edelman_ Trust_Barometer_Global_Report_2. pdf (2019)
2. Effron, D.A., Raj, M.: Misinformation and Morality: Encountering Fake-News Headlines Makes Them Seem Less Unethical to Publish and Share:. Psychological Science (Nov 2019)
3. Gabielkov, M., Ramachandran, A., Chaintreau, A., Legout, A.: Social clicks: What and who gets read on twitter? ACM SIGMETRICS Performance Evaluation Review 44(1), 179–192 (2016)
4. Giachanou, A., Rissola, E.A., Ghanem, B., Crestani, F., Rosso, P.: The role of personality and linguistic patterns in discriminating between fake news spreaders and fact checkers. In: Natural Language Processing and Information Systems: 25th International Conference on Applications of Natural Language to Information Systems, NLDB 2020, Saarbrücken, Germany, June 24–26, 2020, Proceedings. p. 181. Springer Nature
5. Guess, A., Nagler, J., Tucker, J.: Less than you think: Prevalence and predictors of fake news dissemination on facebook. Science advances 5(1), eaau4586 (2019)
6. Gupta, A., Lamba, H., Kumaraguru, P., Joshi, A.: Faking sandy: characterizing and identifying fake images on twitter during hurricane sandy. In: Proceedings of the 22nd international conference on World Wide Web. pp. 729–736 (2013)
7. Horne, B.D., Nevo, D., O'Donovan, J., Cho, J., Adali, S.: Rating reliability and bias in news articles: Does AI assistance help everyone? In: Proceedings of the Thirteenth International Conference on Web and Social Media, ICWSM 2019. pp. 247–256 (2019)
8. Hutto, C.J., Gilbert, E.: Vader: A parsimonious rule-based model for sentiment analysis of social media text. In: Eighth international AAAI conference on weblogs and social media (2014)
9. Isaak, J., Hanna, M.J.: User data privacy: Facebook, cambridge analytica, and privacy protection. Computer 51(8), 56–59 (2018)
10. Lee, C.S., Ma, L.: News sharing in social media: The effect of gratifications and prior experience. Computers in human behavior 28(2), 331–339 (2012)
11. Ma, L., Lee, C.S., Goh, D.H.: Understanding news sharing in social media from the diffusion of innovations perspective. In: 2013 IEEE International Conference on Green Computing and Communications and IEEE Internet of Things and IEEE Cyber, Physical and Social Computing. pp. 1013–1020. IEEE (2013)
12. Mitchell, A., Gottfriedd, J., Barthel, M., Sumida, N.: Distinguishing between factual and opinion statements in the news. Pew Research Center (2018)
13. Neuman, Y.: Computational personality analysis: Introduction, practical applications and novel directions. Springer (2016)

14. Pennebaker, J.W., Boyd, R.L., Jordan, K., Blackburn, K.: The development and psychometric properties of liwc2015. Tech. rep. (2015)
15. Potthast, M., Gollub, T., Wiegmann, M., Stein, B.: TIRA Integrated Research Architecture. In: Ferro, N., Peters, C. (eds.) Information Retrieval Evaluation in a Changing World. The Information Retrieval Series, Springer, Berlin Heidelberg New York (Sep 2019)
16. Rangel, F., Giachanou, A., Ghanem, B., Rosso, P.: Overview of the 8th Author Profiling Task at PAN 2020: Profiling Fake News Spreaders on Twitter. In: Cappellato, L., Eickhoff, C., Ferro, N., Névéol, A. (eds.) CLEF 2020 Labs and Workshops, Notebook Papers. CEUR Workshop Proceedings (Sep 2020), CEUR-WS.org
17. Reimers, N., Gurevych, I.: Sentence-bert: Sentence embeddings using siamese bert-networks. arXiv preprint arXiv:1908.10084 (2019)
18. Shrestha, A., Spezzano, F.: Online misinformation: from the deceiver to the victim. In: ASONAM '19: International Conference on Advances in Social Networks Analysis and Mining. pp. 847–850. ACM (2019)
19. Shu, K., Mahudeswaran, D., Wang, S., Lee, D., Liu, H.: Fakenewsnet: A data repository with news content, social context and dynamic information for studying fake news on social media. arXiv preprint arXiv:1809.01286 8 (2018)
20. Shu, K., Wang, S., Liu, H.: Understanding user profiles on social media for fake news detection. In: 1st IEEE International Workshop on Fake MultiMedia (FakeMM 2018) (2018)
21. Spiro, E.S., Fitzhugh, S., Sutton, J., Pierski, N., Greczek, M., Butts, C.T.: Rumoring during extreme events: A case study of deepwater horizon 2010. In: Proceedings of the 4th Annual ACM Web Science Conference. pp. 275–283 (2012)
22. Vogel, L.: Viral misinformation threatens public health (2017)
23. Vosoughi, S., Roy, D., Aral, S.: The spread of true and false news online. Science 359(6380), 1146–1151 (2018)
24. Yaqub, W., Kakhidze, O., Brockman, M.L., Memon, N., Patil, S.: Effects of credibility indicators on social media news sharing intent. In: Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems. p. 1–14. CHI '20 (2020)