

MisinfoMe: Who is Interacting with Misinformation? *

Martino Mensio^[0000-0002-9875-6396] and Harith Alani^[0000-0003-2784-349X]

Knowledge Media institute, The Open University, UK
{martino.mensio,h.alani}@open.ac.uk

Abstract. Misinformation is a persistent problem that threatens societies at multiple levels. In spite of the intensified attention given to this problem by scientists, governments, and media, the lack of awareness of how someone has interacted with, or is being exposed to, misinformation remains to be a challenge. This paper describes MisinfoMe; an application that collects ClaimReview annotations and source-level validations from numerous sources, and provides an assessment of a given Twitter account, and the followed accounts, with regards to how much they interact with reliable or unreliable information.

Keywords: Misinformation · Twitter · Fact Checking · ClaimReview.

1 Introduction

Misinformation is an old problem, amplified in recent years by social media. Much research tackled this problem from social as well as technical perspectives, to better understand the phenomenon, how it spreads online, how it can be automatically detected, its impact on opinions, etc. Nevertheless, research shows that making people aware of how they interact with misinformation remains to be a challenge [2]. In spite of this, there is hardly any research and technologies that aim at raising awareness of how a given user and their network have interacted with false or unreliable information [1].

This paper introduces MisinfoMe;¹ an application that uses ClaimReview semantic annotations and other information validators to highlight the tweets that point to reliable or unreliable sources and content. MisinfoMe also expands the assessment to the Twitter users followed by the given one, to reflect the amount of misinformation they pass on to the given user. More specifically, the objectives of the MisinfoMe tool are:

1. Create a (mis)information collection of 68K URLs and 1.3K domains, generated by 86 fact-checking organisations,
2. Collect the Twitter timeline of a given user and the accounts they follow,

* Copyright © 2019 for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

¹ <http://socsem.kmi.open.ac.uk/misinfo/>

3. Highlight tweets with URLs pointing to false/true/mixed information,
4. Compare the given user with the “average user” with respect to the level of interaction with reliable and unreliable information,
5. Highlight the levels of misinformation interaction by followed twitter accounts and rank them accordingly.

In the following sections we briefly describe related work, the dataset used by MisinfoMe, and the main components of the application.

Demo: MisinfoMe will be demoed at the conference. Users can enter a Twitter handle to see the assessment of their levels of misinformation interaction, and the tweets that belong to each category (misinforming, valid, mixed). A video recording of the demo is available at: <https://tinyurl.com/yxeun2nt> and the full application is accessible at: <http://socsem.kmi.open.ac.uk/misinfo/>

2 Related Work

Several applications have been developed to help validating different types of content (images, bot accounts, news sources, reviews, etc.). NewsGuard (<https://newsguardtech.com>) is a browser plugin that uses journalists to assess the reliability of websites (e.g., BBC, Wikipedia, Breitbart) following generic credibility and transparency indicators. B.S.Detector (tinyurl.com/y368cbhw) is another browser plugin that offers validations of news sources. ClaimBuster (idir.uta.edu/claimbuster/) detects and ranks claims, and matches them to fact-checks. Botometer (botometer.iuni.iu.edu) calculates the likelihood of a Twitter account being a bot based on various activity patterns. TinEye (tineye.com) provides a reverse image search to help track out of context image usage. Rbutr (rbutr.com) connects web pages that present disputes or contradictions, provided by the user community, and provide them to users when browsing such pages. Fakespot (fakespot.com) assesses the validity of online reviews on sites such as Amazon and Tripadvisor.

Unlike the tools above, MisinfoMe is focused on revealing how a given Twitter account, and the accounts it follows, have been interacting with misinformation. MisinfoMe does not offer new content validations per se, but rather it uses third-party fact-checking and source validations to assess all the URL-bearing Tweets in the timeline of the accounts in question.

3 MisinfoMe

Below we briefly describe the main components of MisinfoMe and its user interface. An in-depth description can be found on the MisinfoMe site itself.

3.1 Validation Dataset

MisinfoMe collects ClaimReview² annotations (see example below) from different fact-check publishers and some aggregators.³ ClaimReview is a standard schema used for fact-checking reviews of claims, and from it we collect

² <https://schema.org/ClaimReview>

³ <https://www.datacommons.org/factcheck/download>

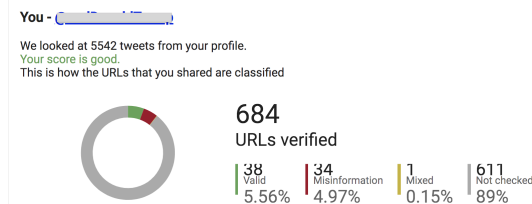
the URLs (if present) of the fact-checked article, and the validation outcome (rating). We then add other publicly available source-assessments, e.g., from OpenSources,⁴ Wikipedia,⁵ and BuzzFeedNews,⁶ combined with the approach described in [3]. A complete list of the sources used is on the MisinfoMe site at <http://socsem.kmi.open.ac.uk/misinfo/about>. In total, MisinfoMe collected 34764 ClaimReviews so far, constructing a dataset of 68K URLs and 1.3K domains, generated by 86 fact-checking organisations.

```
{
  "@context": "http://schema.org",
  "@type": "ClaimReview",
  "author": { "@type": "Organization", "url": "https://www.politifact.com/" },
  "claimReviewed": "President Donald Trump hasnt condemned David Duke and Richard Spencer.",
  "datePublished": "2019-08-26",
  "itemReviewed": { "@type": "Claim", "firstAppearance": {
    "url": "https://www.facebook.com/TND/videos/358402408404240/" } },
  "reviewRating": { "alternateName": "Mostly False", "ratingValue": "3", },
  "url": "https://www.politifact.com/truth-o-meter/statements/2019/aug/27/joe-biden/biden-wrong-when-he-says-trump-hasnt-condemned-dav/" }
}
```

3.2 Assessing Twitter Timelines

Given a Twitter handle, MisinfoMe uses the Twitter API to collect its timeline. URLs appearing in the tweets are then matched with the dataset described above. MisinfoMe displays how many tweets were collected, how many contain a URL (Figure 1), and how many of these URLs point to reliable, unreliable, or mixed pages and sources. To unshorten the URLs, we perform a HTTP HEAD request and follow the HTTP redirections. We also extract the domain name of each URL to compare with domain-level validations.

Fig. 1: URLs are extracted from tweets and categorised according to reliability.



As shown in Figure 2, users can scroll the list of tweets that were assessed, see evidence to support their placement into the reliable, unreliable, or mix category, and the source of this evidence (e.g., Snopes.com). A bar-chart comparison of the given Twitter account with the average of over 770 randomly selected account, is also shown on MisinfoMe (not shown in figures).

3.3 Assessing Social Network

In addition to assessing the given Twitter account with regards to sharing misinformation, MisinfoMe also assesses the accounts followed by the given one with

⁴ <http://www.opensources.co/>

⁵ https://en.wikipedia.org/wiki/List_of_fake_news_websites

⁶ <https://github.com/BuzzFeedNews/2018-12-fake-news-top-50>

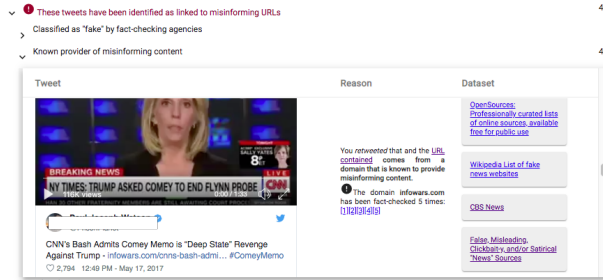
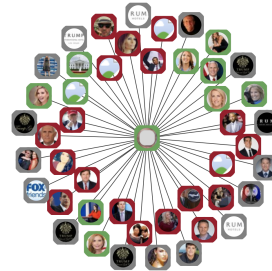


Fig. 2: Tweet pointing to an article from an unreliable publisher. Reason for the assessment and the dataset that provides this judgment are provided.

regards to their spread of misinformation, and ranks them accordingly. Figure 3 shows a selection of the accounts followed by the given account, colour-coded to reflect their tendency to sharing reliable (green), unreliable (red), or mixed (orange) information. Grey indicate that none of their shared URLs were found in our dataset.

Fig. 3: Followed users, coloured to reflect their activities in sharing unreliable information.



4 Conclusions and Future Work

This paper introduced MisinfoMe; an application for assessing the sharing of unreliable information posted by a Twitter account and the accounts it follows. We described the data collection and the analysis performed, and the main user interface components. We are currently planning several user-based experiments, to evaluate the interface, the datasets, and to assess the impact of MisinfoMe on people's awareness of their interactions with misinformation, or of those they are interested in (e.g., celebrities, politicians).

Acknowledgment Co-Infom, H2020, grant agreement 770302.

References

1. Fernandez, M., Alani, H.: Online Misinformation: Challenges and Future Directions. In Companion Proc. The Web Conference, Lyon, France (2018).
2. Southwell, B. G., Thorson, E. A., Sheble, L.: Misinformation and mass audiences. University of Texas Press (2018)
3. Mensio M., Alani H.: News Source Credibility in the Eyes of Different Assessors. In Conference for Truth and Trust Online (2019)