

Computational Fact Checking

Table of Contents

- Introduction
- Knowledge Graphs
 - Properties
- Sources

Introduction

The ubiquity of the Internet has led to a proliferation of information, but not all of it good. Due to a variety of factors, including the uncontrolled nature of the Web, the low barrier to entry for social media platforms, and the ability of a single individual to hold multiple accounts, it is easier than ever to disperse unreliable information [1](#). Unfortunately, we are currently in the Information Age and just as information has revolutionized the world and been responsible for billions in profits, we are now seeing that misinformation can have an effect of the similar magnitude. In recent years, we have seen individuals and organizations spread misinformation to make profits, alter public opinion for financial or political gain, exert influence for a cause, and defame public figures [1](#). Journalists and fact-checking websites have been working to check the factual validity of statements for years, but as the volume of information and the number of agents which we consider "sources" of information both rise, there is no way their methods can keep up.

Determining the degree of truth of a claim is an extremely complex task. Even if this can be accomplished, one needs to consider the context of the claim to decide if despite being factually accurate, the claim misrepresents the bigger picture. This is a common practice in finance and politics, where statistics or numbers are *cherry-picked* in order to show the company or politician in the best possible light, even if the best possible light is the only bright spot. Fact-checking then is the process of putting a claim into context, gathering relevant information, conducting thorough analysis, and reporting a conclusion with explanations and evidence [1](#). The amount of time and work it takes to accurately and thoroughly fact-check a claim leads to misinformation having anywhere from minutes to days to spread before any corrective action can be taken. In order to increase the efficiency of fact-checking the field of computational journalism has sprung up, hoping to tackle problems in journalism with Computer Science tools such as natural language processing, information extraction, data integration, and information visualization [1](#).

Knowledge Graphs

Creating a system for automated fact-checking represents a huge computational challenge, but

recent advances in question-answering systems such as IBM Watson and PowerAqua not only give us hope, but also an avenue for exploration [1](#). Those systems are able to tackle question-answering because they use a data structure known as a Knowledge Graphs.

Properties

Homophily

Homophily is the tendency of individuals to associate and bond with similar and is one of the biggest contributors misinformation spreading across the Internet [1](#). In social networks, it leads to groups of users called *echo chambers* which do not appreciate contrary views or a diversity of opinion [1](#). However, it can also be used to our advantage when performing relational learning. This is helpful for us because we can cluster entities based on links and from there infer links based on the what is known about similar entities. Social media does this when they realize that you and 30 other people all like the same 10 bands, so the social media platform recommends a band to you that the other 30 people also all like.

Sources

[Computational Fact Checking by Mining Knowledge Graphs](#)