# Project Checkpoint 1: Disinformation Challenge

- **Have your data set on hand and start any manual annotation that you may need to do.**

  We have identified the following annotated datasets for the analysis. Each of the datasets come from different sources such as twitter,Politifact and BSDetector. We describe each of them below.

  **LIAR dataset:** This is publicly available and used for the task of fake news detection. This dataset was collected over a decade and contains 12,836 human-labeled short statements, which are sampled from various contexts from POLITIFACT.COM, which provides detailed analysis report and links to source documents for each case. The labels for news truthfulness are fine-grained classes: pants-fire, false, barely-true, half-true, mostly true, and true.

  **BSDetector dataset on Kaggle:** This has text and metadata scraped from 244 websites tagged as "bullshit" by the BS Detector, a Chrome Extension by Daniel Sieradski. Data presents 12,999 posts in total pulled using the webhose.io API. The label for each website is the output of the BS detector that includes labels such as ‘bias’, ‘conspiracy’, ‘fake’, ‘hate’, ‘satire’, rather than human annotations. Data sources that were missing a label were assigned a label of "bs".

  **FakeNewsNet:** This contains two datasets collected using ground truths from Politifact and Gossipcop. The output labels include ‘real’ and ‘fake’. PolitiFact is a website that provides fact-checking evaluation for news articles that were annotated as fake or real by journalists and domain experts. GossipCop is a website for fact-checking entertainment stories collected from different media platforms. It rates a news story on the scale of 0 to 10 to classify the degree of fakeness.

- **Write the introduction section of your project report: Introduction should describe the problem that you are trying to solve, and explain why it should be interesting (to the audience), why it is challenging (technically).**

  Identifying disinformation has been a challenge although many methods have been proposed recently by identifying emerging trends in the spread of disinformation. The disinformation has been growing rapidly due to the heavy use of social media by people and high level of social engagement. Social media owners are finding ways to tackle the spread of disinformation online due to its spread and global impact. Existing research has been carried out on two broad categories of disinformation that are opinion-based which include fake reviews and fact-based which include fake news. The scope of this project is to develop methods and to experiment with existing methods to identify fake news. These days there are few open source browser extensions for real time detection of misinformation. But the performance of such automatic detection engines has not proven to be reliable due to the real challenge in defining what fake news is and also the limited availability of annotated data with gold standard labels. Human labellers who annotate the dataset also have an impact/bias of the misinformation because they filter the information they read (For example, people read/watch what they like and ignore what they don't) .

- **Write the related work section of your project report: Related work section should discuss some of the previous attempts to the problem you are trying to solve. It should highlight the best solutions, why they work, where they fall short, and how these approaches relate to your proposed solution to the problem.**

  Lot of research in this area has been done to identify real and fake news using fact checking tools and to understand the reliability of the news by using surface level linguistic based methods that identify patterns in the writing style. Other research has based analysis by building real time analyzers to crawl data from social media sources such as twitter to collect the tweets regarding a claim to identify if it is true. Besides the news text, the meta data can contribute to identification of fake news which include social context and socio temporal information.

  In addition to research on developing methods and tools for automatic detection of fake reviews there were surveys published that focussed on survey on fake news in social media that assess how an user would perceive if a social media post is real or fake and survey of false information on the web and social media which included three topics - fake reviews in e-commerce platforms, hoaxes on collaborative platforms, and fake news in social media.

  While twitter has been on the top list for research in this area where many publications discussed methods on identifying features based on topic that help in early detection of misinformation, there are also research methods focussed on personal messaging platforms such as whatsapp. Research on Fact-checking WhatsApp rumors suggests that raising a signal of doubt on a claim can suppress the effect of its spread. An interesting system used in social media to reduce circulation of misinformation is to rate the news sources based on their authenticity. For instance, in reddit's Manchester united fan community, they list 'tiers' for various news sources that report transfer news during the transfer season.

- **Describe in a paragraph some of the potential approaches you will explore for your solution to the problem. How are your solutions different from what has already been done?**

  Given the annotated data from several sources, the challenge is to use the right metric that could evaluate the performance of deep learning models. This is possible by using a combination of metrics. We also propose to use socio-political analyses to better understand the nature of information and its distribution. We believe a combination of human input to state of the art deep learning methods could reap significant improvements in detecting misinformation.