

# Rating the disinformation risks of news domains: the Global Disinformation Index

**Santhosh Srinivasan**

GDI / San Francisco, CA

santhosh.srinivasan@disinformationindex.org

**Craig Fagan**

GDI / Berlin, Germany

## Abstract

Disinformation is the shadowy side to the open internet. It undermines faith in institutions, economies, governments, and even democracy itself. The Global Disinformation Index (GDI) is an assessment of the risk of news domains in media markets around the globe of disinforming their public. The index is structured as a neutral, independent and transparent assessment of a news sites disinformation risk at the domain level. The following sections set out the methodology and approach of the index, including the use of different disinformation flags to score and determine the disinformation risk rating of a news site. .

## 1 Approach

As part of the assessment, each news site that has been reviewed will be assigned a risk-level for disinformation: *low, medium-low, medium, medium-high or high*. Domains are an important step in transferring disinformation from its origins, and are key beneficiaries of ad monies. As such, domains are a good measurement unit for the GDI and its first iteration. The index is focused on determining and assessing the risk factors that could expose or allow a news domain to carry disinformation in a given media market. For the pilot, the media market is defined as the top 30 news sites in a country based on their Alexa rankings and other indicators such as their social media following (on Facebook and Twitter). The index itself is constructed using:

- an automated, machine-learning classifier which categorises large volumes of low-quality high-risk news sites in real time, and
- a manual assessment suitable for lower-risk disinformation news outlets that may not be easily discernible by automated or technical

means. We have termed these the "gray area" sites as they fall between the different extreme risk thresholds: high and low. We associate "high-risk" with low quality, "junk" news sites that are hastily assembled. We consider "low risk" to be "quality" news sites that are more established, trusted and credible.

## 2 Index Framework and Data Collection

The GDI index framework has selected 13 indicators or "disinformation flags": **1.) Accountability, 2.) Biasedness, 3.) Conflict of Interest, 4.) Credibility, 5.) Fabrication, 6.) Hate Speech, 7.) Impartiality, 8.) Inaccurate Reporting, 9.) Operational Integrity, 10.) Reputation of Brand, 11.) Sensationalism, 12.) Trustworthiness, and 13.) Verifiability.**

As a composite, they provide a strong proxy of the disinformation risk level of a news site. These indicators are spread over four dimensions or pillars: Structure, Content, Operations and Context.

- **Structure:** This pillar contains the domain level metadata signals, which are evaluated using an automated algorithm based assessment process. Based on a sample of 20,000 known disinformation domains, the GDI is building a tagging and classifying system (i.e. filter). The tags look at similar metadata and other computational signals. These signals are effective for detecting in real time high-risk sites. In addition to the disinformation flags mentioned above, this pillar will have metadata flags (still being defined).
- **Content:** This dimension contains indicators that assess the quality, tone, and verifiability of articles published on a specific domain. We will take a random sample of 10 of the

top shared articles on a news domain over a period of two weeks. These stories (and their sites) will be anonymised and assessed by a country analyst. A combined score for the 10 stories will be assigned to each domain for this pillar.

- **Operations:** This dimension assesses the underlying policies and rules that domains abide by to establish trust and reliability in the quality of news being published. We will apply a sub-set of five indicators selected from the Journalism Trust Initiative (JTI) that serve as proxies to assess the disinformation risk arising from lapses in journalistic integrity. Until the JTI data is available for a media market, a country analyst will collect this information for the GDI.
- **Context:** This dimension focuses on assessing overall credibility and reliability of a specific news site. It looks at signals for understanding the more nuanced risks of a domain disinforming its users - currently or in the future. It also tests against some of the findings in the other pillars to assess correlation and coherence. All data is collected through a survey of media experts (from academia, media companies, the press, think tanks and civil society organisations). A series of questions are posed through a 15-minute online survey conducted by an external polling company. The panel is 50 to 100 experts per country (depending on the size of the media market).

Each site receives an overall score per pillar. Each pillar has a score from 0 (high-risk) to 100 (low-risk). For each site, all pillar scores are aggregated to calculate a site score (i.e. risk).

The index has been developed with a technical advisory group drawn from the disinformation and fact-checking communities. Members include Camille Francois (Graphika), Dr. Scott Hale (Meedan/Oxford Internet Institute), Olaf Steenfadt (Reporters without Borders/Journalism Trust Initiative), and Cristina Tardaguila (International Fact Checking Network/Poynter).

### 3 Current Assessment Models

We reviewed current initiatives to assess the approaches used to measure disinformation. Overall, current initiatives can be broadly categorised as doing the following:

- understanding the reach and impact of disinformation;
- fact-checking various domains;
- defining and measuring the quality of journalism; and
- developing indicators to measure disinformation measurement.

Most efforts are focused on assessing the quality of journalism or the measures in place that contribute to high-quality journalism. None conduct metadata assessments of news domains or use a hybrid approach such as being proposed for the GDI. Those initiatives that use automated methods are focused on analysing the news content and do not seem to take a comprehensive view of various disinformation risks.

#### 3.1 Findings

The GDI is piloting its methodology in two countries: South Africa and the United Kingdom. We expect all data to be finalised and ready for publication by November 2019. For now, we can offer some initial findings regarding the process of developing and running the index:

- there is a strong need to use questions and indicators that can be cross-checked to test the correlation between them as valid "flags" of disinformation risk;
- it is important to create a diverse media market sample in terms of perceived levels of disinformation risk;
- the use of a risk-rating tool is a good way to move away from a binary assessment of whether a site is disinforming or not; and
- defining what is a news site can be challenging and subjective (i.e. should satire sites, personal blogs and sports sites be included in a market sample?).

#### 3.2 Way Forward

The GDI plans to use the lessons learned from the pilot to go to scale to up to 10 countries in 2020. The pilot findings may require a reworking of the disinformation flags. However, the overall approach suggests that using "risk ratings" rather than binary scores is the better approach for assessing disinformation at the site level.