GDI

**Global
Disinformation
Index**

Rating
Disinformation
Risk: The GDI
Methodology

Author: Santhosh Srinivasan

The Global Disinformation Index is a UK-based not-for-profit that operates on the three principles of neutrality, independence and transparency. Our vision is a world in which we can trust what we see in the media. Our mission is to restore trust in the media by providing real-time automated risk ratings of the world's media sites through a Global Disinformation Index (GDI). For more information, visit www.disinformationindex.org

GDI | Global Disinformation Index

Table of contents

Introduction

The Global Disinformation Index's (GDI) vision is a world free of disinformation and its harms. One of the channels to achieve this vision is to identify a domain's disinformation risk and curtail its ability to spread disinformation.

To combat disinformation, we need to understand efforts to disinform – both upstream (where disinformation starts) and downstream (where and how it spreads). For the GDI, financial motivation is a connecting point that links together the upstream and downstream components of disinformation. One key financial motivation is the inadvertent funding provided to disinformation sites from programmatic advertising. To substantially reduce disinformation, we need to disrupt this funding source. For the GDI, this means providing programmatic ad buyers and other stakeholders with a trusted, neutral and independent assessment of a news site's disinformation risks. Such ratings are critical for brand safety and can be used by advertisers, ad tech companies and others to control where their advertising appears based on their individual appetites for disinformation risk.

The GDI risk assessment is based on structural, contextual, operational and content indicators. A site's risk rating is based on cutting-edge artificial intelligence (i.e. automated review) combined with a careful analysis of an expert-agreed set of disinformation signals (i.e. human review). Both the automated and human reviews have been piloted in 2019 and will be fully rolled out in the first half of 2020.

Conceptual framework

The GDI defines 'disinformation' as: to purposely and/or maliciously mislead by spreading inaccurate information (in terms of its content and context).

The index's framework uses a combination of signals that contribute to assessing a site's risk to disinform, based on structural, contextual, operational and content indicators (See Appendix A).

The index does not assess whether a specific news domain is actually carrying disinformation. This differentiation is critical.

The index does not determine which news stories are inaccurate, or whether a site should be labelled as a disinformation domain. The index should not be used in this way or to judge what is true news.

The focus of the index is on the site level, and it is designed to cover news domains. A news domain is defined as a range of sites: it can be "a single individual (e.g. a blogger), one team or department within a larger media organisation (e.g. a certain radio channel, a newsroom, a show, website or paper) or a whole media organisation with many outlets and brands".[1]
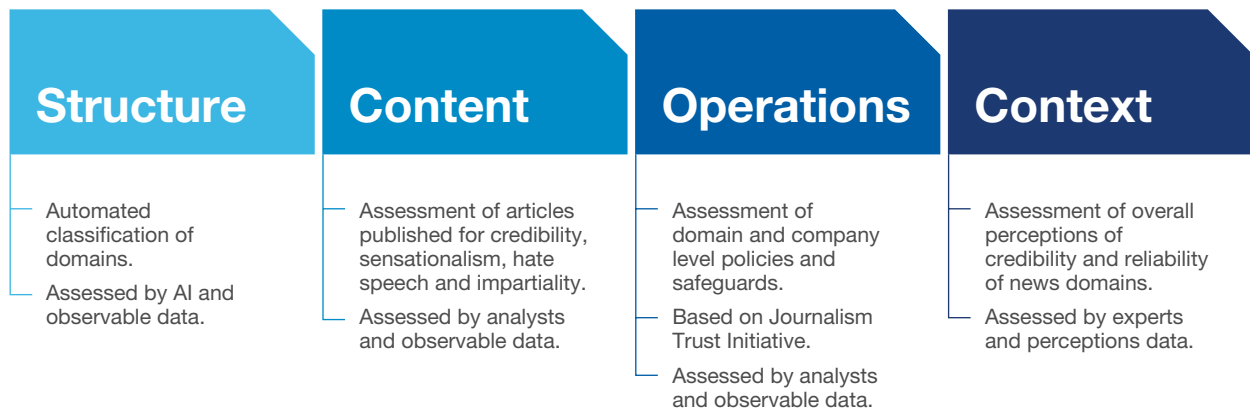
The GDI site-level ratings will give advertisers a trusted assessment of the potential disinformation risks associated with sites that carry their adverts. Advertisers and ad tech companies can use GDI's neutral and independent risk rating to decide where to best direct their ad spend based on their own brand safety and risk mitigation strategies.

Key pillars of the index

The GDI assessment is the first global disinformation risk rating of news sites. It is grounded on best practice standards for designing and developing an index. The GDI rating is based on a combination of automated and manual assessments of various indicators of disinformation risk spread over four pillars: **'Structure', 'Content', 'Operations' and 'Context'** (See Figure 1). These pillars and their indicators have been selected to focus on observable site features and practices that minimise the risk of bias in our assessment.

The development of this index began with a review of good index design[2] drawn from different sectors and using core principles needed to design a robust and reliable measurement metric. The index was also developed and finalised with the guidance of a technical advisory group (TAG) composed of leading experts from the fields of media, journalism, intelligence operations, governance and index design.

Figure 1. Pillars of the Global Disinformation Index

Structure	Content	Operations	Context
— Automated classification of domains. — Assessed by AI and observable data.	— Assessment of articles published for credibility, sensationalism, hate speech and impartiality. — Assessed by analysts and observable data.	— Assessment of domain and company level policies and safeguards. — Based on Journalism Trust Initiative. — Assessed by analysts and observable data.	— Assessment of overall perceptions of credibility and reliability of news domains. — Assessed by experts and perceptions data.

■ **The 'Structure' pillar** reviews news domains based on various metadata and computational signals. This pillar uses AI to assess these signals and rate a site's disinformation risk in real time. The AI-based classifier automatically tags news domains using machine learning that has been developed from a training dataset of 20,000 domains. This automated review is particularly effective for assessing the 'long tail' of domains with a high disinformation risk. This pillar of the index is constantly updated and revised to reflect the evolving nature of high-risk disinformation sites. Some of the current signals look at a site's advertisements, format, content, language, targeting (negative toward certain groups), and topics.[3]

■ **The 'Content' pillar** contains indicators that assess different elements of news articles published on a specific domain, including their *credibility, sensationalism, neutrality and impartiality*. As for all of the pillars, each of these indicators was chosen to identify and measure a specific disinformation signal or flag (see Appendix A). How a domain's content is presented and covered is an important indicator of the disinformation risk of the domain. Some of the more pernicious forms of disinformation occur when news domains present a variety of straight and accurate news with a few maliciously and purposefully inaccurate stories in order to gain and manipulate users' trust (often called "bait-and-switch" tactics). To assess this pillar, a country analyst reviews a set of ten anonymised articles randomly selected from the site's most shared articles over a period of two weeks. To anonymise the articles, they were presented as text files without any reference to the site. Any related photos, videos or other imagery were removed from the articles. The aim of this step was to eliminate any potential biases of the reviewer and to neutrally present the text.

■ **The 'Operations' pillar** assesses the underlying policies and rules that domains abide by to establish trust and reliability in the quality of the news being published. The integrity of a news organisation and its site is a good indication of whether checks and balances are in place to prevent or lower the risk of disinformation from appearing on a site. When certain editorial policies and guidelines are not in place, there is an elevated risk that journalistic integrity is compromised, and an elevated risk that disinformation could appear.

For this pillar, we apply a subset of five indicators selected from the Journalism Trust Initiative (JTI). These indicators serve as good proxies to assess the risk of disinformation arising from lapses in journalistic integrity. The JTI is a multi-partner initiative led and endorsed by journalists to set a global ISO standard and certification process for ethical norms and professional best practices for all media outlets, big and small.
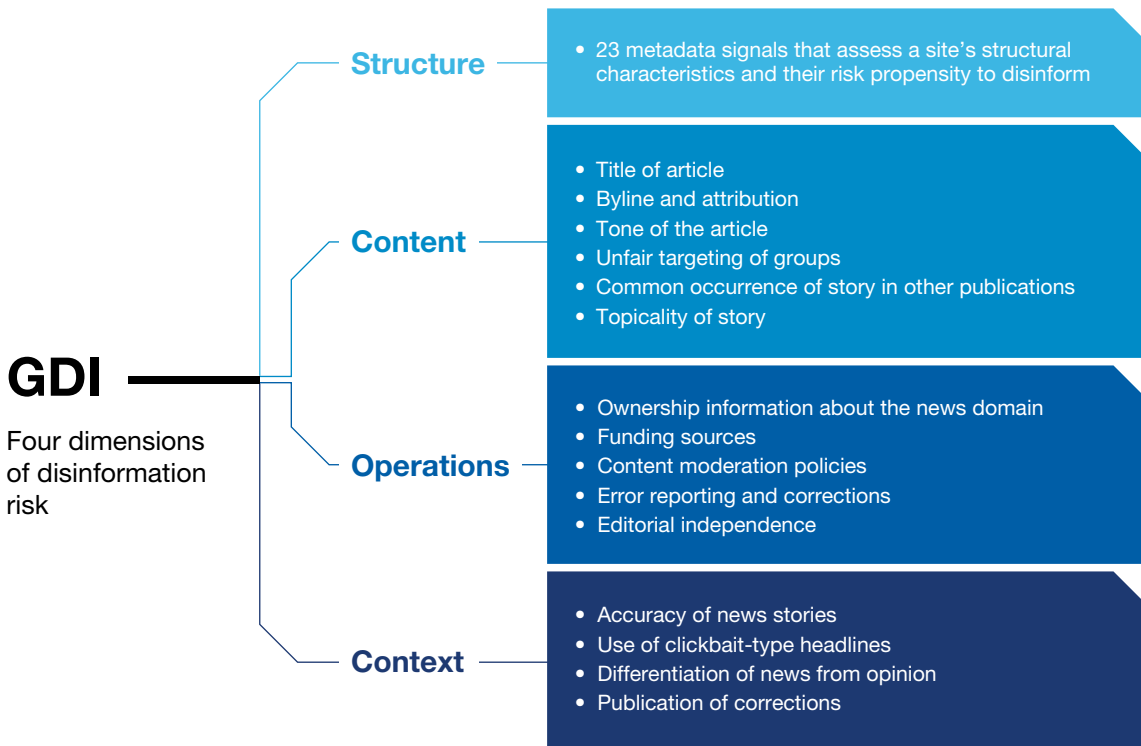
The operations pillar uses five indicators to assess four disinformation risks: *conflict of interest, lack of operational integrity, inaccurate reporting, and lack of accountability*. By adopting a subset of the JTI indicators, the index will be the first to collect data on these indicators and therefore contribute to the JTI database. The JTI standard and the associated indicators have been approved and will roll out in 2020.[4] Until news sites start reporting on the JTI standard, GDI will collect this data.

The 'content' and 'operations' pillar assessments are conducted by reviewers, who rely upon a Researcher's Notebook that provides clear and detailed instructions on how to review articles and/or domains.[5] This ensures standardisation of the review process. Additionally, analysts compile all related information on the domains assessed as evidence to support their assessment of any specific indicator.[6]

■ **The 'Context' pillar** assesses the overall credibility and reliability of news-related information provided by a specific domain. The overall conduct of a site can go beyond a sample of content and the operational policies in place. It relates to how the news domain is viewed: the overall perceived trustworthiness of the site. The disinformation flags assessed in this pillar are related to *credibility, trustworthiness, conflicts of interest and biasedness.* As these signals are not easily measurable by analysts, this pillar and the questions in it are assessed by country-level experts working on media and related issues. In each country, a panel survey of up to 100 experts is conducted to score a news domain's performance.[7] Expert responses provide perceptions-based data, while the other pillars provide observational-based data.

As an overall framework, the first three pillars ('structure', 'content' and 'operations') are snapshots of a site's disinformation risks at a moment *in time*. The fourth pillar ('context') is an attempt to draw on expert knowledge to capture long-term site trends *over time* that create disinformation risks. As a collective set of pillars and indicators, they are able to assess a site's disinformation risk. Indicators must be taken as an aggregate rather than focused on individually to have a multi-dimensional understanding of this risk (see Figure 2).

Figure 2. GDI dimensions and indicators

GDI

Four dimensions of disinformation risk

Structure
- 23 metadata signals that assess a site's structural characteristics and their risk propensity to disinform

Content
- Title of article
- Byline and attribution
- Tone of the article
- Unfair targeting of groups
- Common occurrence of story in other publications
- Topicality of story

Operations
- Ownership information about the news domain
- Funding sources
- Content moderation policies
- Error reporting and corrections
- Editorial independence

Context
- Accuracy of news stories
- Use of clickbait-type headlines
- Differentiation of news from opinion
- Publication of corrections

Note: For a detailed list of indicators under each pillar, please see Appendix A.

Domain selection

The AI classifier of the index uses a proprietary platform that crawls through thousands of news domains and tags them automatically.

The manual pillars of the index – content, operations and context – are applied to a subset of news domains per media market that are difficult for the automated process to classify. The automated classifiers are still being refined to include differentiated disinformation risk thresholds and scores (which are aligned with the methodology outlined here).

To assess a media market for disinformation risk, we develop a market sample of 30 media sites based on the sites' current reach and market relevance.

The market sample is chosen for manual review[8] using predefined metrics:

- the site's Alexa score/rank,[9]
- the site's number of Facebook followers, and
- the site's number of Twitter followers.

This approach allows for the selection of those news domains for analysis that have some of the highest levels of engagement and therefore relatively greater impact and reach on the overall media market. In addition to these metrics, we also look at sites that are considered to influence policy because they shape ideas among key decision-makers, groups or actors.[10]

These combined criteria limit the overall number of domains to be rated through human review to a manageable number, while still capturing a set of market-relevant domains for a country.

Assessment process

To calculate a site's disinformation risk rating, the following four steps are performed:

STEP 1

The first step in the assessment process is to review the assessment done by the automated classifier and score it. The automated assessment includes flags that were set based on site-level structural signals as well as metadata tags assigned to the sites by the classifier. All 30 domains per media market are scored for their level of disinformation risk based on the automated classifier findings. This is currently a binary rating system (high risk versus low risk). In 2020, this binary system will be developed into a risk rating on a 0–100 scale.

STEP 2

The 'Content' pillar is the first set of indicators to be assessed as part of the manual (human review) process. The selected domains per media market are assessed on the six disinformation indicators for the content pillar: *clickbait, fabrication, biasedness, hate speech, sensationalism, and credibility.*

The assessment is carried out by one analyst (reviewer) per media market. Each analyst reviews ten randomly selected and anonymised news stories or articles per domain. To compile the articles, the index samples articles published on a news domain over a two-week period. The top five stories each day are selected based on the number of Facebook shares for that story. A total of 70 articles are compiled per domain, from which ten articles are randomly selected for review for each domain.[11]

All of the assessed articles are scraped using a proprietary software programme. The software removes all domain-level attributes such as the name of the website or publication and any other identifying characteristics from the chosen articles. The articles are then saved as text files for the manual review. Additionally, a human crosscheck is conducted is to ensure that the articles are fully anonymised to prevent the reviewer from being able to discern the original source of the article. This process prevents any bias from entering into the analyst's assessment of the 'Content' pillar.[12]

STEP 3

The 'Operations' pillar contains four disinformation risks captured through five indicators: *conflict of interest, lack of operational integrity, inaccurate reporting, and lack of accountability*. This pillar assesses only whether a domain has publicly disclosed policies and other information related to the indicators; it does not assess whether these policies are also put into practice. The policies are considered necessary by the Journalism Trust Initiative to support high journalistic standards for any media site, regardless of its size.

Responses to the indicators must be documented to substantiate whether a policy or other information is publicly available. This includes screenshots, URL links, PDFs, and/or excerpts of relevant text.

The same country reviewer who completes the 'Content' pillar also assesses the 'Operations' pillar.[13]

STEP 4

The 'Context' pillar measures a domain's overall reliability and focuses on the following disinformation flags: *credibility, trustworthiness, conflicts of interest and biasedness.* This pillar is assessed by local media experts through a survey which is carried out by a third party.[14] Experts surveyed include experienced journalists, senior-level media researchers and academics, executives from non-governmental organisations (NGOs) and think tanks and business leaders.[15]

The decision to use an expert survey to gather the data was made for two reasons:

- The best way to assess the general trustworthiness of a news site is by asking a diverse group of senior-level media experts and practitioners who are more knowledgeable about journalistic standards and integrity than the average news consumer.

- Expert responses to the survey also validate some of the other indicators that are measured in the 'Content' (e.g. clickbait headlines) and 'Operations' (e.g. corrections policy) pillars, offering a crosscheck on these findings.

The unit of analysis for the 'Context' pillar is domains. Each expert is first asked to indicate all of the domains s/he is aware of from the total list of 30 domains, and then shown a subset of up to ten domains from this list. For each domain, the experts then rate seven indicators related to the disinformation flags measured.

The survey methodology ensures that each domain is rated by at least five experts.[16] We also controlled and tested the results for political bias, which showed one's political self-identification does not affect how they rate a site.

Ten steps to mitigate bias and ensure quality

The index assessment process has multiple mechanisms built in to mitigate bias and produce robust data. For an independent and trusted risk-rating tool, it is essential to ensure that such controls are in place.

The GDI risk assessment has created such a trusted framework by relying on validated analyst reviews and credible data from expert responses. Such a mixed approach of data sources helps to ensure that any biases are mitigated and that the index is a robust measure of domain-level disinformation risk.

As part of the GDI's methodology framework, the following strategies have been put in place to make sure that the index is reliable and provides an accurate estimate for the risk of disinformation.

1. The index methodology is informed by an in-depth review of the literature, both academic and non-academic, on misinformation, disinformation, ethics and integrity in journalism.

2. In designing the index, the GDI first reviewed the best practices for index construction to assist us in developing the index. A synthesis of this literature is available here.[17]

3. The media market list for each country is selected based on consistent metrics for site traffic and relevance (using both as proxies for potential reach and impact of a site).

4. All articles assessed for the 'Content' pillar are anonymised to prevent reviewers' personal biases from influencing their scoring.[18]

5. In the pilot, the 'Content' and 'Operations' pillar scores were reviewed by GDI staff to ensure that the scores were backed up by relevant information, and that domains and articles were scored consistently. Additionally, a subset of articles was also reviewed anonymously by a third reviewer to assess the inter-rater reliability of scoring by the analysts.

6. The expert survey used for the 'Context' pillar relies on senior-level journalists, researchers, academics and other professionals who have a strong knowledge of the principles and practices associated with free, credible and trusted media.

7. For the pilot, the survey results were tested statistically to ensure that experts' personal political views did not influence their responses.

8. Statistical tests have been conducted to validate the theoretical model underpinning the index and to check the relevance of all the indicators used to assess disinformation risk in our model.

9. A panel of experts from various fields have validated the tests used and their associated findings.

10. An advisory group consisting of experts from the fields of information security, journalism, index development, and survey research have provided expert guidance and support throughout the development of the index and its piloting.

Scoring methodology

The scoring scale for the index ranges from 0 (maximum risk of disinformation) to 100 (minimum risk of disinformation). A domain's score on the index is the mean of the scores earned across all four of the pillars.

■ The **'Structure'** pillar is composed of 23 metadata signals that are assessed to determine whether the site contains any disinformation flags. A domain is scored a 0 or 100 for the total set of signals, with 100 indicating low risk. Going forward, these metadata indicators will be expanded to include lexical and computational signals which will be used to compute a risk rating for each site. Given the nature of disinformation and neural networks, the specific signals used will continue to evolve and change.

■ The **'Content'** pillar contains indicators that assess news articles to provide an overall assessment of the content carried by the domain. Each article is assessed on six indicators. An article can receive a score between 0 and 100 for each indicator. A total of 10 anonymised articles is assessed per domain. The domain level score for each indicator in this pillar is the average score obtained across the ten articles. The pillar score for each domain is the average of all the scores for all of the pillar's indicators. The score for each domain in the 'Content' pillar ranges between 0 and 100.

■ The **'Operations'** pillar has five indicators, with each indicator assessed at the domain level. Similar to the 'Structure' pillar, the total domain score for the pillar is an average score across the five indicators. The total score a domain can obtain can range between 0 and 100 points, with a higher score indicating a lower risk of disinformation on this pillar.

■ The **'Context'** pillar for each domain is based on four indicators. The total score for each domain in this pillar can range from 0 to 100, based on an average score across the four indicators. Each domain contains multiple expert responses per question. Therefore, the domain level score per indicator will be the average score across all experts reviewing each domain.[19]

The final score per domain is a simple average of the scores for all four pillars.[20] To ensure that the scoring is transparent and understandable, there are no weights assigned to any of the pillars or indicators. Moreover, on the basis of statistical testing, it was determined that adding a weighting scheme does not significantly alter the final results.[21] The objective of this methodology is to crosscheck, limit and minimise any biases in the final scores to provide a neutral assessment of a site's disinformation risk.[22]

The domains are classified on a five-category risk scale based on their final index scores. In order to determine a domain's risk level, it is assumed that the factors influencing disinformation risk are similar across all domains, independent of the media market they belong to.

The final score per domain is standardised to fit a standard normal distribution with a mean of 0 and standard deviation of 1.[23] The resulting standardised scores and their distance from the mean are used to determine the overall risk rating for a domain. As such, it is important to use the aggregate score across all the pillars in order to understand the level of disinformation risk posed by a news site. The individual pillars and indicators help to understand how and where improvements can be made to mitigate risks, but should not be used alone to determine the risk level of a site (See Figure 3).

Figure 3. Calculation of risk levels

TOTAL DOMAIN SCORE	DISINFORMATION RISK LEVEL	DISINFORMATION RISK CATEGORY
< -1 SD from mean	5	Maximum risk
≥ -1 and < - 0.5 SD from mean	4	High risk
≥ -0.5 and ≤ 0.5 SD from mean	3	Medium risk
> 0.5 and ≤ 1SD from mean	2	Low risk
> 1 SD from mean	1	Minimum risk

The calculated scores and risk levels have been developed using a mathematical and statistical approach to assess a site's risk. The aim of this methodology is to ensure a neutral, independent, trusted and transparent rating of news domains and their disinformation risks. These ratings are seen as a critical input to aid advertisers and the ad tech community to best direct their ad monies to new sites that align with their brand safety and risk mitigation strategies.

The GDI methodology should been seen as a living document. It will continue to be refined and updated to reflect changes in how disinformation is funded, produced and spread.

Appendices

Appendix A: List of indicators by pillar

Content

	Assessment Area	Disinformation Risk Flag
1	Title Representativeness	Clickbait
2	Attribution	Fabrication
3	Tone	Bias
4	Targeting	Hate Speech
5	Recent Coverage	Sensationalism
6	Common Coverage	Credibility

Operations

	Assessment Area	Disinformation Risk Flag
1	Ownership of Site	Conflict of Interest
2	Funding	Conflict of Interest
3	Policy Guidelines	Operational Integrity
4	Errors and Corrections	Inaccurate Reporting
5	Editorial Independence	Lack of Accountability

Context

	Assessment Area	Disinformation Risk Flag
1	Accuracy	Trustworthiness
2	Labeling of Opinion and News	Sensationalism/Credibility
3	Clickbait	Biasedness/Sensationalism
4	Correction Practices	Credibility

Appendix B: Expert survey questionnaire

Media expert - Profile questions

*Question type: **Single***

[scr1_media] Thinking about the main paper that you contribute to, is your media outlet global, national, regional or local?

<1>	Global
<2>	National
<3>	Regional
<4>	Local – CLOSE

*Question type: **Multiple***

[scr2_media] And which of these subject areas do you personally cover in your role?

<1>	Current affairs
<2>	Politics
<3>	Business & finance
<4>	Culture
<5>	Lifestyle
<6>	Sport
<7>	Technology
<95>	Other [open] please specify

*Question type: **Single***

[scr3_media] Which of these media outlets do you work for, or if you are freelance which do you mostly tend to work for?

<1>	Broadsheet newspaper
<2>	Mid-market newspaper
<3>	Tabloid newspaper
<4>	Television
<5>	Radio
<6>	Current affairs or business magazine
<7>	Other magazine (trade, specialist, consumer etc)
<8>	Online publication
<95>	Other

*Question type: **Single***

[scr4_media] Please indicate your job function:

<1>	Editor
<2>	Correspondent/Reporter
<3>	Presenter
<4>	Columnist
<5>	Freelance
<95>	Other [open] please specify
<99 xor>	None of these – CLOSE

*Question type: **Single***

[scr5_media] How many years have you worked in media?

<1>	Less than 3 years – CLOSE
<2>	4-5 years
<3>	6-7 years
<4>	8-9 years
<5>	10+ years

*Question type: **Multiple** | #Row order: **Randomise***

[scr6_media] Do you have a professional interest in any of the following issues? Please select all that apply.

CLOSE IF NOT CODED AT LEAST ONE CODE IN RED

<1>	Poverty
<2>	Education
<3>	Healthcare/Disease
<4>	Human Rights
<5>	Climate/Environment
<6>	Technology
<7>	Democracy and Governance (including anti-corruption)
<8>	Media rights
<9>	Digital rights
<10>	Access to information
<11>	Transparency
<12>	Freedom of expression
<13>	Cybersecurity
<14>	Terrorism
<15>	Extremism
<95 fixed xor>	None of the above

NGO Expert - Profile Questions

*Question type: **Single***

[q1_NGO] Which of these types of organisation do you work for?

<1>	NGO – International Level
<2>	NGO – National Level
<3>	Charity
<4>	Think Tank
<5>	Pressure/Advocacy Group
<6>	Academic Research Institute
<7>	Association/Federation
<95>	Other

*Question type: **Single***

[q2_NGO] Do you sit on the Executive Board of your organisation or do you report directly to an Executive Board member?

<1>	I sit on the Executive Board
<2>	I report directly to an Executive Board Member
<3>	Neither of these - I do not report directly to an Executive Board Member – CLOSE

*Question type: **Single***

[q3_NGO] Please indicate your job function:

<1>	Analyst or researcher
<2>	Fundraising Manager or Director
<3>	General Management/Administration
<4>	Campaign Manager or Director
<5>	Policy Manager or Director
<6>	Communications Manager or Director
<7>	Research Manager or Director
<8>	Scholar
<9>	Other
<10>	CEO
<11>	Chairman
<12>	Founder
<97>	None of these

*Question type: **Multiple** | #Row order: **Randomise***

[q4_NGO] Do you have a professional interest in any of the following issues? Please select all that apply.

CLOSE IF NOT CODED AT LEAST ONE CODE IN RED

<1>	Poverty
<2>	Education
<3>	Healthcare/Disease
<4>	Human Rights
<5>	Climate/Environment
<6>	Technology
<7>	Democracy and Governance (including anti-corruption)
<8>	Media rights
<9>	Digital rights
<10>	Access to information
<11>	Transparency
<12>	Freedom of expression
<13>	Cybersecurity
<14>	Terrorism
<15>	Extremism
<95 fixed xor>	None of the above

Academic Expert - Profile Questions

*Question type: **Single***

[academic_title_screener] Which of the following best describes your professional title?

<1>	PhD Student – CLOSE
<2>	Postdoctoral Research Assistantship/Fellowship – CLOSE
<3>	Assistant or Associate Lecturer
<4>	Lecturer
<5>	Senior Lecturer or Reader
<6>	Associate Professor
<7>	Professor
<8>	Dean
<96>	Other – CLOSE
<97>	None of these – CLOSE

*Question type: **Multiple** | #Row order: **Randomise***

[scr1_academic] Do you have a professional interest in any of the
following issues? Please select all that apply.

CLOSE IF NOT CODED AT LEAST ONE CODE IN RED

<1>	Poverty
<2>	Education
<3>	Healthcare/Disease
<4>	Human Rights
<5>	Climate/Environment
<6>	Technology
<7>	Democracy and Governance (including anti-corruption)
<8>	Media rights
<9>	Digital rights
<10>	Access to information
<11>	Transparency
<12>	Freedom of expression
<13>	Cybersecurity
<14>	Terrorism
<15>	Extremism
<95 fixed xor>	None of the above

*Question type: **Single***

[politics_scale_profile_update] Some people talk about 'left', 'right' and 'centre' to describe
parties and politicians. With this in mind, where would you place yourself on this scale?

<1>	Very left-wing
<2>	Fairly left-wing
<3>	Slightly left-of-centre
<4>	Centre
<5>	Slightly right-of-centre
<6>	Fairly right-wing
<7>	Very right-wing
<8>	Don't know

Appendix C: Statistical analysis for model validity

The index framework, with its four pillars extending over 13 indicators, was developed based on lessons learned from a synthesis of best practices around index construction, and with the help of and input from a ten-member technical advisory group (TAG). Upon completion of data collection across the three manual review pillars – 'Content', 'Operations', and 'Context' – several data validation and statistical tests were conducted using the data from the pilots in the UK and South Africa, both to test the assumptions behind the index framework and to ensure high standards of data quality.

The correlation matrices for both the UK and South Africa show a fairly high degree of correlation between indicators, both within specific pillars as well as for indicators across pillars.[24] For example, the findings for the UK show a positive correlation among the 'Content' pillar indicators related to the title, tone and targeting of the article, and a significant correlation between the tone and the title and target indicators. Similarly, within the 'Operations' pillar, the ownership, funding, error, and independence indicators show a positive correlation. And finally, all of the indicators within the 'Context' pillar measured using the expert survey are positively and significantly correlated with each other. Similarly, the indicators from the 'Context' pillar such as accuracy and corrections are also well correlated with associated indicators such as title and tone from the 'Content' pillar, and with the error and independence indicators from the 'Operations' pillar. This indicates that brand reputation assessed by the expert survey is influenced by the core indicators captured in the 'Content' and 'Operations' pillars, and that the overall assessment framework is robust.

Figure C.1 Correlation matrix for the UK

	Title	Byline	Tone	Target	Recent	Common	Owner	Funding	Policy	Error	Independence	Accuracy	Opinion_v_news	Clickbait	Corrections
Title	1.00	0.01	0.49**	0.26	-0.06	0.25	0.06	-0.05	0.25	0.13	0.24	0.42*	0.38*	0.46*	0.49*
Byline	0.01	1.00	-0.14	-0.05	0.02	-0.07	0.03	-0.09	0.07	0.30	0.06	-0.26	-0.14	-0.24	-0.13
Tone	0.49**	-0.14	1.00	0.64**	0.18	0.47**	0.34	0.37*	0.52**	0.46*	0.36	0.40*	0.49**	0.47**	0.65**
Target	0.26	-0.05	0.64**	1.00	0.15	0.23	0.34	0.33	0.29	0.48**	0.39*	0.32	0.33	0.42*	0.512**
Recent	-0.06	0.02	0.18	0.15	1.00	0.17	0.05	-0.21	0.32	0.10	0.15	-0.24	-0.43*	-0.23	-0.17
Common	0.25	-0.07	0.47**	0.23	0.17	1.00	0.25	0.03	0.40*	0.30	0.00	-0.04	0.00	-0.08	0.17
Owner	0.06	0.03	0.34	0.34	0.05	0.25	1.00	0.58**	0.33	0.41*	-0.02	0.20	0.36*	0.22	0.43*
Funding	-0.05	-0.09	0.37*	0.33	-0.21	0.03	0.58**	1.00	0.18	0.375*	0.26	0.37*	0.44*	0.40*	0.50**
Policy	0.25	0.07	0.52**	0.29	0.32	0.40*	0.33	0.18	1.00	0.47**	0.30	0.09	0.20	0.10	0.44*
Error	0.13	0.30	0.46*	0.48**	0.10	0.30	0.41*	0.38*	0.47**	1.00	0.52**	0.17	0.31	0.23	0.42*
Independence	0.24	0.06	0.36	0.39*	0.15	0.00	-0.02	0.26	0.30	0.52**	1.00	0.48**	0.38*	0.55**	0.43*
Accuracy	0.42*	-0.26	0.40*	0.32	-0.24	-0.04	0.20	0.37*	0.09	0.17	0.48**	1.00	0.87**	0.96**	0.82**
Opinion_v_news	0.38*	-0.14	0.49**	0.33	-0.43*	0.00	0.36*	0.44*	0.20	0.31	0.38*	0.87**	1.00	0.86**	0.86**
Clickbait	0.46*	-0.24	0.47**	0.42*	-0.23	-0.08	0.22	0.40*	0.10	0.23	0.55**	0.96**	0.86**	1.00	0.81**
Corrections	0.45*	-0.13	0.65**	0.51**	-0.17	0.17	0.43*	0.50**	0.44*	0.42*	0.43*	0.82**	0.86**	0.81**	1.00

* indicates $p < .05$ | ** indicates $p < .01$

The correlations between indicators for South Africa are similar to those for the UK.

Figure C.2 Correlation matrix for South Africa

	Title	Byline	Tone	Target	Recent	Common	Owner	Funding	Policy	Error	Independence	Accuracy	Opinion_v _news	Clickbait	Corrections
Title	1.00	0.27	0.76**	0.77**	0.20	0.31	0.25	0.12	0.15	0.18	0.22	-0.07	-0.35	0.01	0.14
Byline	0.27	1.00	0.30	0.39*	0.16	0.01	0.36	0.43*	0.30	0.29	0.20	0.20	0.17	0.19	0.25
Tone	0.76**	0.30	1.00	0.87**	0.36	0.37*	0.37*	0.32	0.28	0.09	-0.05	-0.35	-0.49*	-0.31	-0.02
Target	0.77**	0.39*	0.87**	1.00	0.32	0.36*	0.36*	0.26	0.30	0.19	0.10	-0.16	-0.09	0.10	-0.21
Recent	0.20	0.16	0.36	0.32	1.00	0.36	-0.20	-0.02	0.21	0.28	0.00	0.09	-0.06	0.07	0.05
Common	0.31	0.01	0.37*	0.36*	0.36	1.00	-0.19	-0.27	0.13	0.53**	0.23	-0.08	-0.08	-0.06	-0.04
Owner	0.25	0.36	0.37*	0.36*	-0.20	-0.19	1.00	0.70**	0.22	-0.04	0.06	-0.20	-0.06	-0.34	-0.14
Funding	0.12	0.43*	0.32	0.26	-0.02	-0.27	0.70**	1.00	0.37*	0.23	-0.20	0.00	0.02	-0.25	0.30
Policy	0.15	0.30	0.28	0.30	0.21	0.13	0.22	0.37*	1.00	0.34	-0.24	-0.18	0.13	0.02	-0.01
Error	0.18	0.29	0.09	0.19	0.28	0.53**	-0.04	0.23	0.34	1.00	0.27	0.36	0.24	0.07	0.38
Independence	0.22	0.20	-0.05	0.10	0.00	0.23	0.06	-0.20	-0.24	0.27	1.00	0.06	-0.18	0.19	0.00
Accuracy	-0.07	0.20	-0.35	-0.16	0.09	-0.08	-0.20	0.00	-0.18	0.36	0.06	1.00	0.77**	0.76**	0.64**
Opinion_v_news	-0.35	0.17	-0.49*	-0.09	-0.06	-0.08	-0.06	0.02	0.13	0.24	-0.18	0.77**	1.00	0.50*	0.29
Clickbait	0.01	0.19	-0.31	0.10	0.07	-0.06	-0.34	-0.25	0.02	0.07	0.19	0.76**	0.50*	1.00	0.65**
Corrections	0.14	0.25	-0.02	-0.21	0.05	-0.04	-0.14	0.30	-0.01	0.38	0.00	0.64**	0.29	0.65**	1.00

* indicates p < .05 | ** indicates p < .01

As a further step to test the validity of indicators used to assess disinformation risk, a principal component analysis (PCA) was conducted for both the UK and South Africa to verify whether and how well the indicators capture underlying constructs of disinformation risk. The PCA results support the three pillars of the framework. Based on the PCA results, multiple weighting scenarios were also tested. In light of the results of this scenario analysis, a decision was made to not weight the data based on the pilot exercise results.

Figure C.3 PCA results for the UK (left) and South Africa (right)

	C1	C2	C3
Title	0.50	0.45	-0.53
Tone	0.33	0.71	0.03
Target	0.26	0.50	0.21
Common	-0.29	0.82	-0.06
Owner	0.02	0.26	0.74
Funding	0.28	-0.01	0.78
Policy	-0.06	0.76	0.10
Error	0.05	0.57	0.39
Independence	0.52	0.23	-0.04
Accuracy	0.99	-0.11	-0.03
Opinion_v_news	0.88	-0.01	0.14
Clickbait	0.99	-0.07	-0.01
Corrections	0.77	0.27	0.17

	C2	C1	C3	C4
Title	0.14	0.94	-0.02	-0.01
Tone	-0.21	0.82	0.20	0.10
Target	0.01	0.82	0.15	0.17
Common	-0.14	0.29	-0.46	0.71
Owner	-0.12	0.33	0.70	-0.18
Funding	0.08	0.15	0.90	0.02
Policy	-0.14	-0.05	0.49	0.63
Error	0.28	0.08	0.02	0.78
Independence	0.21	0.43	-0.47	-0.09
Accuracy	0.93	-0.09	0.01	0.05
Opinion_v_news	0.59	-0.44	0.21	0.29
Clickbait	0.89	0.10	-0.19	-0.08
Corrections	0.82	0.13	0.16	0.00

Based on the results above, it is clear that: the indicators for title, tone and target capture a key disinformation risk related to a domain's narrative; the indicators for ownership and funding capture another aspect of disinformation risk related to transparency and accountability; and the indicators for accuracy, opinion versus news differentiation, clickbait and corrections point to a third category around brand reputation and trust. Additionally, the factor analysis results also support the assumptions behind the current index framework and its underlying model.

Since the 'Context' pillar is measured on the basis of expert survey responses and incorporates multiple perception metrics within the pillar, a statistical test was conducted to test the independence of experts' personal political "opinions" and their responses to the survey questions. A chi-square test for independence as well as a one-way analysis of variance test established that respondents' personal political opinions did not have a significant impact on the way they responded to the questions in the survey. In addition, for domains that had fewer than ten experts rating them on this pillar, a decision was made to exclude this pillar score from their total domain score.[25]

Endnotes

1   This definition comes from the Journalism Trust Initiative
    (JTI): https://jti-rsf.org/en/#!the-solution.

2   *Designing the Index: A review of good practices*, GDI, May 2019.

3   An earlier prototype established a strong link between the site structure and
    the quality of an online outlet. Site structure was a good proxy in determining
    whether a website had been set up to just generate clicks and generally presents
    junk news: a news domain that "deliberately publish(es) misleading, deceptive
    or incorrect information purporting to be real news about politics, economics
    or culture." For definition of 'junk', see: https://newsaggregator.oii.ox.ac.uk

4   Between July and October, 2019 JTI had invited both the general public
    and professional communities to participate in a public consultation and
    to comment on the proposed JTI indicators. All feedback received will be
    incorporated and a standard document will be released by the end of 2019.

5   The Researcher's Notebook was developed by the GDI team to provide clear standardised
    guidelines to the analysts assessing the 'content' and 'operation' pillar indicators.

6   Such evidence includes screenshots or any other textual
    information available from the site and/or article reviewed.

7   In the pilot phase, 50 experts were surveyed for South Africa
    and 100 experts for the United Kingdom.

8   In both the UK and South Africa, we sampled 31 news domains to
    provide a buffer in case of data collection issues or gaps.

9   Alexa rank is a measure of website popularity. It ranks millions of websites
    in order of popularity, with an Alexa Rank of 1 being the most popular. Alexa
    Rank reveals how a website performs relative to all other sites, which
    makes it a great KPI for benchmarking and competitive analysis.

10 In the pilot, we also included a few sites that were flagged as carrying low-quality, untrusted
    news so we could better understand how well the framework assessed these types of sites.

11 For the pilot, one analyst per country has been used. This may change as
    the index is rolled out to other countries to be two analysts per country. For
    the index pilot in the UK and South Africa, the articles were selected over a
    period of six weeks between August 1, 2019 and September 15, 2019.

12 To make this as objective as possible and limit potential bias, a subset of
    articles was also reviewed by another independent analyst to investigate any
    variation in the different analysts' ratings of the same articles. The second
    analyst also reviewed a sample of non-anonymised articles in order to
    determine the effectiveness and usefulness of the anonymisation process.

13 As noted, once sites begin reporting to the JTI, the dataset for the operations pillar
    will come directly from the JTI rather than being collected by the country reviewers.

14 For the pilot, the surveys were conducted by YouGov, which is a polling company that conducts polls about politics, public affairs, and other topics of general interest. See: www.yougov.com.

15 The survey is administered either online or by phone. In the case of the pilot countries, it was done online in the UK (100 experts), but administered through phone interviews in South Africa (50 experts). See Appendix B for more information on the types of expert profiles that responded to the survey.

16 In the pilot, most of the sites were rated by a statistically relevant number of experts. For example, 93% of the domains in the UK and 20% of the domains in South Africa were rated by at least 25 experts. Due to challenges with the survey in South Africa, eleven sites in the sample did not meet the threshold of ten expert assessments per question per site. In future iterations of the survey, this outcome will be prevented by including additional survey respondents (up to 100 per country) and/or modifications to the market sample that is to be assessed.

17 See note 2.

18 There were discussions with the Technical Advisory Group (TAG) regarding whether or not anonymising articles and using only text files removed too much of the framing in which stories are presented on sites. For example, the text of an article might be straight news, but photos, videos or other visuals accompanying the story may be biased. However, it was decided to pilot the approach of anonymised articles to assess the results and benefits of text-only articles (to limit potential biases on the part of the reviewer).

19 The expert survey used a five-point Likert scale, which was converted to a 0 to 100 scale using the formula (maxnew - minnew)/(maxold - minold) x (v - maxold) + maxnew.

20 For the pilot in the UK and South Africa, the final score was a simple average of scores across three pillars: 'Content', 'Operations', and 'Context'. This was based on the fact that all sites received a 100 score on the 'Structure' pillar, given the current indicators and scoring in place for them.

21 This finding also substantiates the fact that the GDI framework holds up well and provides a robust assessment of a news site's risk of carrying disinformation.

22 For information on the types of validity checks performed, please refer to Appendix C.

23 The domains from all media markets are combined to compute the mean and standard deviation of the distribution of index scores, and those estimates are used for standardisation.

24 Indicators that were not well correlated, such as byline and topicality in the 'Content' pillar, were excluded from the score calculations for the pilot. However, the relevance and validity of these indicators will be tested again when the index is scaled up to cover more countries in 2020.

25 This affected 11 news sites, all of them in South Africa.

GDI | Global Disinformation Index

www.disinformationindex.org