

GDI

Global
Disinformation
Index



Disinformation Risk Assessment: The Online News Market in Italy



Authors: Marinella Petrocchi and Angelo Spognardi.

Researchers: Giulia Consorti, Michaela Fazzolari and Manuel Pratelli.

Design: www.designbythink.co.za

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 free from disinformation and its harms. Our mission is to catalyse industry and government to defund disinformation. We provide disinformation risk ratings of the world's news media sites.

For more information, visit www.disinformationindex.org



SAPIENZA
UNIVERSITÀ DI ROMA



ISTITUTO
DI INFORMATICA
E TELEMATICA



Consiglio Nazionale
delle Ricerche



January 2022. Published under a Creative Commons License (CC BY-NC-SA 4.0)

Table of contents

Executive summary	4
The Italian media market: Key features and scope	6
Disinformation risk ratings	8
Conclusion	15
Appendix: Methodology	16
Endnotes	20

Executive summary

Since the news business has expanded to the online world, transformations in news production and distribution have exposed the industry to new disinformation risks.

Websites have financial incentives to spread disinformation, in order to increase their online traffic and, ultimately, their advertising revenue. Meanwhile, the dissemination of disinformation has disruptive and impactful consequences. The COVID-19 pandemic offers a recent – and deadly – example. By disrupting society's shared sense of accepted facts, these narratives undermine public health, safety and government responses.

To combat ad-funded disinformation, the Global Disinformation Index (GDI) deploys its assessment framework to rate news domains' risk of disinforming their readers. These independent, trusted and neutral ratings are used by advertisers, ad tech companies, and platforms to redirect their online ad spending, in line with their brand safety and disinformation risk mitigation strategies.

GDI defines disinformation as 'adversarial narratives that create real-world harm'. The GDI risk rating provides information about a range of indicators related to the risk that a given news website will disinform its readers by spreading these adversarial narratives. These indicators are grouped under the index's **Content** and **Operations pillars**, which respectively measure the quality and reliability of a site's content and its operational and editorial integrity.¹ A site's overall risk rating is based on that site's aggregated score across all the indicators, and ranges from zero (maximum-risk level) to 100 (minimum-risk level).

The GDI risk rating methodology is not an attempt to identify and label disinformation sites or trustworthy news sites. Rather, GDI's approach is based on the idea that a combined set of indicators can reflect a site's overall risk of carrying disinformation. The ratings should be seen as offering initial insights into Italy's media market and its overall levels of disinformation risk, along with the strengths and challenges the sites might face in mitigating disinformation risks.

The following report presents the findings pertaining to disinformation risks for the media market in Italy, based on a study of 34 news domains. These findings are the result of the research led by the GDI in collaboration with the Institute of Informatics and Telematics, National Research Council, Pisa, and the Computer Science Department, Sapienza University of Rome, from April through October 2021. All sites included in the report were privately informed of their individual scores and risk ratings, to allow for engagement and feedback.

The need for a trustworthy, independent rating of disinformation risk is pressing. This risk-rating framework for Italy will provide crucial information to policy-makers, news websites, and the ad tech industry, enabling key decision-makers to stem the tide of money that incentivises and sustains disinformation.

Key findings: Italy

In reviewing the media landscape for Italy, GDI's assessment found that:

More than a third of the sites in our sample fell within the high- or maximum-risk categories.

- Ten of the sites in our sample have a high risk of disinforming their users online, while three present a maximum-risk rating.
- Many of these sites tend to publish biased and sensational content, thus creating a risk of manipulating their audience. Some of them, especially those belonging to the maximum-risk rating, also produce content that negatively targets individuals, groups or organisations.
- These sites perform poorly on most indicators from the **Operations pillar**. The results highlight that significant effort is required in order to improve or introduce transparency about their policies concerning **Attribution, Ensuring accuracy, comment regulation, Funding and Ownership**, as well as adherence to **Editorial principles and practices**.
- Many of these sites tend to publish news not covered by any other media outlets, which is also frequently not based on recent, newsworthy events.

A quarter of the sites present low- or minimum-risk levels of disinformation.

- The two minimum-risk sites achieve strong results on both the **Content** and **Operations pillars**.
- Seven sites were rated with a low level of disinformation risk. These sites also score well on **Content pillar** indicators; they tend to carry unbiased, non-sensational content, with no adversarial narratives.
- These low-risk sites perform slightly worse on the **Operations pillar** indicators, due mainly to the lack of transparency over policies covering **Attribution, Ensuring accuracy, and Editorial principles and practices**.

Twelve of the media sites in our sample scored a medium-risk rating.

- The findings suggest that for these sites, the most significant improvements to their risk rating can be achieved through improving their performance on the **Operations pillar**. In particular, improvements can be achieved by publishing policies and statements that provide readers with information about **Funding** sources, as well as information about the guidelines used for content **Attribution, Ensuring accuracy** and regulating user-generated content (i.e. **Comment policies**).
- On the **Content pillar**, most of these sites achieve a low score on the **Byline information** indicator. This indicates a lack of information about the author (or news agency) that wrote the content. It is worth noting that publishing the identity of the author is an easy way to ensure transparency and accountability. Including a byline is key to improving readers' awareness of who produces the content they are reading.

Considering the overall picture, even sites with low-risk scores have significant room for improvement on the **Operations pillar**. In fact, while the content they publish adheres to sound journalistic and editorial principles, even the best-performing sites often fail to publish an adequate code of ethics, disclose the author of an article, and clearly detail the ownership and financial structure of the media outlet.

The Italian media market: Key features and scope

The year 2020 was an extraordinary one from several points of view, and had a significant impact on the digital consumption of news.

The COVID-19 pandemic severely affected almost all sectors, causing huge losses in terms of human lives. And with these unconscionable losses there were also huge waves of economic and social instability. The social containment measures (mainly lockdowns) rolled out globally by governments also forced many people to stay home, with great effects on their digital lives. As revealed by the 'Digital 2021' report by WeAreSocial, "many people connected to the internet for the first time as they adapted to the challenges of COVID-19, while existing users embraced new digital tools and rediscovered old favourites".² This is also true in Italy, where 2020 saw more than 1 million people become 'new internet users' (an increase of 2.2 percent), reaching an estimated number of 50.54 million internet users, corresponding to 83.7 percent of the Italian population.³

However, Italy is a country characterised by great diversity in digital savviness and access to computing technologies among the population. According to a 2019 report, there are regions where one in three families does not even have an internet connection, a phenomenon called the 'digital divide'.⁴ There are, indeed, several rural areas where the average wealth is below the national average, as is the level of education, which is tightly bound to the 'early school leaving' phenomenon, as reported in the Education and Training Monitor 2020 report of the European Commission.^{5,6} Regarding Italian users' internet news consumption, the Italian National Institute of Statistics (ISTAT) revealed that in 2020, around 63 percent of respondents between the ages of 20 and 74 use the internet to read online newspapers, news websites or online magazines. IPSOS research about the frequency of media usage for news shows that in Italy, social media are the second source of information (after

television), followed by news sites online.^{7,8} Thus, even though the Italian statistics have strong geographical fluctuations, the internet is the second-most used source of information and, consequently, a critical means of communication for citizens. The IPSOS survey also showed that only two out of ten Italians usually trust news shared with them by people they know only through the internet (e.g., influencers, bloggers, leaders, celebrities on social media).

The survey also highlighted that, in the past five years, the overall trust placed in online news websites and platforms by Italians has decreased by 19 percent, with 69 percent of respondents claiming that 'fake news' is prevalent in the news and information provided by online news websites and platforms. Disinformation campaigns in Italy have largely focused on issues such as immigration and anti-EU rhetoric, and more recently, the COVID-19 pandemic. Originating from both local and foreign actors alike, these campaigns are generally part of broader narratives in support of right-wing and anti-establishment political movements within the country.⁹ For instance, ahead of the 2019 European Parliament election, investigative research by Avaaz followed the activity of a network of Facebook pages and accounts which spread divisive content (frequently in support of the Five Star Movement or Northern League) about topics such as immigration and national security, many of which have since been shut down by Facebook.¹⁰ More recently, during the COVID-19 pandemic, both Russia and China launched social media campaigns in Italy to improve their reputation among the Italian population, while also downplaying and distorting the roles and efforts of other actors in the crisis, like the EU.^{11,12} Partially as a result of such campaigns, eighty-eight percent of respondents to a Monitor Italia survey, carried out from March 12 to 13, 2020, believed that the EU was not helping Italy during the COVID-19 pandemic.¹³ Twitter also played an important role in spreading disinformation about the pandemic among right- and center/right-wing

communities in Italy. Posts containing URLs linking to news published on dubious media outlets shared on Twitter during the peak of the pandemic in Italy (February 2020 – April 2020) reached 22.1 percent of Twitter users on the right and center-right wing, with 96 percent of these URLs shared by political groups emerging from this community.¹⁴

The proliferation of disinformation is a phenomenon best addressed by free and independent journalists. The Council of Europe monitors the state of press freedom in its 47 member countries.¹⁵ According to the 2020 report, ‘free and independent journalism suffered enormous damage.’ Indeed, the number of reported physical attacks on journalists doubled between 2016 and 2020, with Italy ranking second only to Russia. Example cases include the attack on a Sky TG24 crew in Naples during an anti-lockdown demonstration, the attack on journalist Mimmo Rubio, who was also a victim of anti-lockdown protesters, and the attacks on several journalists following far-right protests in Rome in June 2020. In recent years, Italy has also witnessed an increase in strategic lawsuits against public participation (SLAPPs) – e.g. defamation suits – leveraged by influential actors.¹⁶ These lawsuits – abusive of the legal system – are brought against individuals and organisations with ‘an aim to shut down critical speech by intimidating critics and draining their resources’, tying them up in court and undermining their ability to hold the powerful to account.¹⁷

A characteristic feature of the Italian news media market is that almost all private sources of news (TV/radio/newspapers, local or national, printed or online) are effectively owned by only a few big media groups. These, in turn, are controlled by major companies active primarily outside the media sector: GEDI, RCS, the Caltagirone Group, Monrif and Fininvest. One of the most active is the GEDI group (Exor corporation, controlled by the Agnelli family) which publishes *La Repubblica*, *La Stampa* and dozens of local newspapers, as well as magazines, and also owns several radio stations nationwide. Estimates show that the GEDI group currently commands 25.4 percent of the Italian print market in terms of revenue. RCS MediaGroup is owned by Cairo Communication (controlled by Urbano Cairo), which owns *Il Corriere della Sera*, sport newspapers, several local newspapers and the La7 television channel. Cairo Communication is now the second largest player in the print market and the fifth by audience in the audiovisual segment. Similarly the Caltagirone Group (Caltagirone family) publishes, among several others, *Il Mattino* and *Il Messaggero*,

the latter of which is the third-most visited online news website in Italy. Monrif (Monti and Riffeser) publishes *QN Quotidiano Nazionale*, *Il Resto del Carlino*, *La Nazione*, and *Il Giorno*, the last of which was the fifth-most popular newspaper in Italy by circulation in 2019. Fininvest (Berlusconi family) owns *Il Giornale* and *Il Foglio* and controls the Mediaset company, which is the second-most influential TV broadcasting group, with a market share of 31.6 percent in 2019, only behind the public service broadcaster RAI.

While the Italian government has mechanisms in place to support print newspapers and promote the plurality of information sources, funding from large companies helps newspapers to face the ups and downs of this media sector. In Italy, this has become the de facto standard, creating grounds for suspicion that editors and journalists constantly deal with the influence of their main shareholders. This highly concentrated corporate media ownership can have a substantial influence on both the quality of the published content and the broader editorial policies which guide content creation and selection. Yet even independent newspapers are not immune to the influence of large corporations, for example, when threatened with cuts in advertising revenue.

While the exceptional rise in online consumption of media in 2020 did not lead to a similar level of growth in online advertising revenue, the online ad market is nevertheless expected to maintain the steady growth of recent years, as also noted by AGCOM in its research focused on the 2020 online ad market.¹⁸ The COVID-19 pandemic has impacted the entire ad media market, with an expected downturn of around 2 million euros for the Italian market and a reduction of 11 percent in revenue for the online market. Thus, ‘for the first time there was a reversal of a positive trend that seemed unstoppable if we consider absolute values, and it is assumed that the gap between online advertising and sales in other sectors will continue to widen’. In fact, since 2015, online advertising revenue in Italy has increased by 18 percent, from 24 percent to 42 percent of the market, overtaking the previous principal medium, namely TV in 2019, with an impressive growth from 1.66 to 3.33 million euros.¹⁹

Disinformation risk ratings

This study focuses on a sample of 34 Italian news websites.

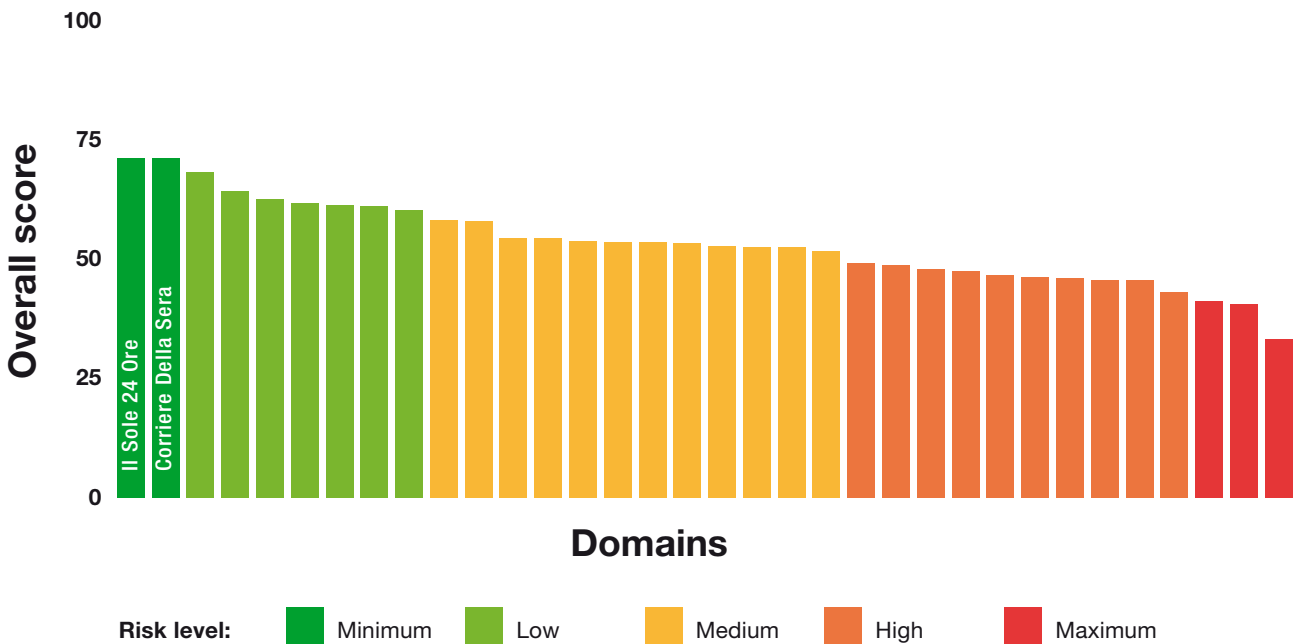
Market overview

The final sample of websites (see Table 1) was defined after performing two successive steps. A first version of the list was produced based on the sites' reach (using each site's Alexa rankings, Facebook followers, and Twitter followers). Then other features such as 'diffusion' (national or local), 'location' (north, centre, south) and 'political orientation' (left, right, neutral) were used to produce the final sample of domains reviewed. The basic idea was to produce a list that was balanced in terms of diffusion, location and political orientation to ensure unbiased and complete coverage over all these aspects.

Table 1. Media sites assessed in Italy (in alphabetical order)

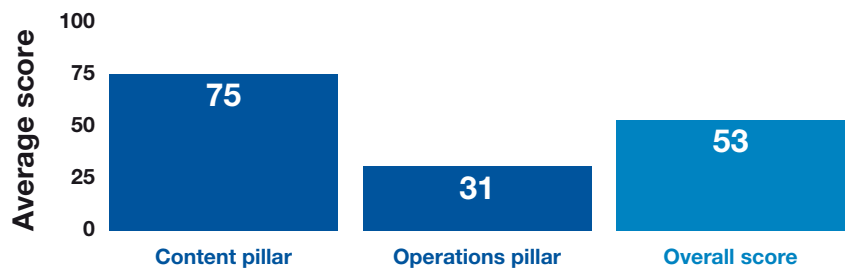
News outlet	Domain	News outlet	Domain
Avvenire	www.avvenire.it	Il Resto Del Carlino	www.ilrestodelcarlino.it
Corriere Della Sera	www.corriere.it	Il Secolo D'italia	www.secoloditalia.it
Domani	www.editorialedomani.it	Il Sole 24 Ore	www.ilsole24ore.com
Il Corriere Del Giorno	www.ilcorrieredelgiorno.it	Il Tirreno	www.iltirreno.gelocal.it
Il Fatto Quotidiano	www.ilfattoquotidiano.it	L'unione Sarda	www.unionesarda.it
Il Foglio	www.ilfoglio.it	La Gazzetta Del Mezzogiorno	www.lagazzettadelmezzogiorno.it
Il Gazzettino	www.ilgazzettino.it	La Nazione	www.lanazione.it
Il Giornale	www.ilgiornale.it	La Nuova Ferrara	www.lanuovaferrara.gelocal.it
Il Giornale Di Sicilia	www.gds.it	La Nuova Padania	www.lanuovapadania.it
Il Giorno	www.ilgiorno.it	La Nuova Sardegna	www.lanuovasardegna.it
Il Manifesto	www.ilmanifesto.it	La Repubblica	www.repubblica.it
Il Mattino	www.ilmattino.it	La Stampa	www.lastampa.it
Il Messaggero	www.ilmessaggero.it	La Verità	www.laverita.info
Il Piccolo	www.ilpiccolo.gelocal.it	Libero	www.liberoquotidiano.it
Il Post	www.ilpost.it	Libertá	www.liberta.it
Il Primato Nazionale	www.ilprimatonazionale.it	Open	www.open.online
Il Quotidiano Del Molise	www.quotidianomolise.com	Stopcensura	www.stopcensura.online

Figure 1. Disinformation risk ratings by site



The findings for Italy show media domains nearly equally divided into three groups (see Figure 1). The first group consists of minimum- or low-risk domains and contains nine sites, around 26 percent of the sample. At the other end of the risk spectrum, thirteen sites, around 38 percent, fall in the high- or maximum-risk categories. Twelve sites, around 35 percent, are in the medium-risk category. This group contains sites with the greatest likelihood of improving their scores in the short term. Overall, some of the main risk indicators in Italy that influence the classification come from the **Operations pillar**, including indicators such as **Ensuring accuracy**, **Attribution** and **Comment policies**.

Figure 2. Overall market scores, by pillar



In Italy, only two sites received a minimum-risk rating: Il Sole 24 Ore (www.ilssole24ore.it) and Corriere Della Sera (www.corriere.it). The sites performed very well on all of the content indicators: all of the articles assessed were neutral and unbiased, carried bylines, and had headlines which matched the story’s contents and did not negatively target groups or individuals. The sites also had many of the key operational policies in place, including information about **Funding** and **Ownership**, and statements of editorial independence.

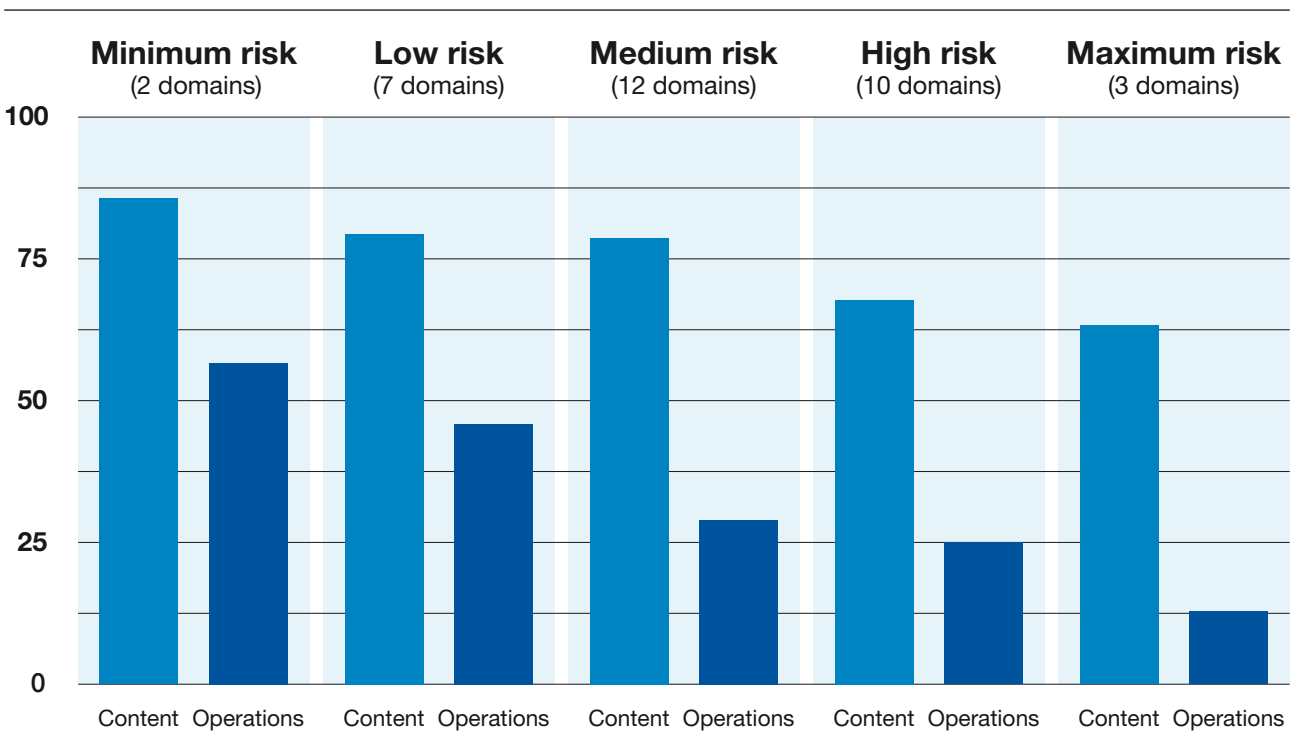
Seven sites in Italy were rated as low-risk sites. These sites achieve good performance on the content indicators, especially for carrying unbiased and non-sensational content, without any adversarial narrative against any specific individuals or groups. However, they performed slightly worse on the operations indicators, due mainly to the lack of statements defining content **Attribution**, **Ensuring accuracy**, and **Editorial principles and practices**.

Twelve sites were assessed with a medium-risk rating. These sites generally performed well on **Content pillar** scores, providing unbiased and non-sensational language, although they achieved, on average, a low

score concerning the use of **Byline information**. It is worth noting that the main difference between these and the low-risk sites was due to the **Operations pillar** results (see Figure 3). In order to improve the scores for that pillar, the key aspects to work on are the lack of transparency about policies and statements detailing the processes used for content **Attribution**, **Ensuring accuracy**, a declaration of **Funding** sources, and **Comment policies** for user-generated content.

The 13 remaining sites received a high- or maximum-risk rating. Ten sites received a high-risk rating, while three sites were in the maximum-risk category. They all achieved very poor results for all of the **Operations pillar** indicators, namely **Attribution**, **Ensuring accuracy**, **Comment policies**, **Editorial principles and practices**, as well as **Funding** and **Ownership**. These sites also tend to produce more biased and sensational content. Furthermore, two of the sites classified as maximum-risk – and several of the high-risk domains – occasionally negatively target individuals, groups or organisations in their narratives.

Figure 3. Average pillar scores by risk rating level



Pillar overview

Content pillar

This pillar focuses on the reliability of the content provided on the site. Our analysis for the **Content pillar** is based on an assessment of ten anonymised articles for each domain. These articles are drawn from (1) the most frequently shared pieces of content during the data collection period; and (2) articles on topics that are likely to polarise and carry disinformation. All article scores are based on a scale of zero (worst) to 100 (best), as assessed by the country reviewers.

Figure 4 shows the average values obtained from the various content indicators. In general, we can say that domains performed well on a majority of the indicators. The indicators that achieved the highest scores are **Visual presentation** (i.e., the reviewers' judgement on how sensationalist the page where the article appears is in terms of, for example, the presence of bold text, and images or videos that are sensationalised), as well as the use of **Sensational language** and the presence of bias in the article text.

On the other end of the spectrum, the indicator that obtained the lowest score was the **Byline information** indicator; this is an indication that, on average, the name of the author (or news agency) that wrote the news is largely missing in articles from Italian domains. For the 10 articles reviewed for each of the 34 news sites, the average value was 57. Other low-performing indicators that ended up below the overall **Content pillar** average were **Common coverage**, **Lede present** and **Recent coverage**. The **Common coverage** indicator, with an average score of 65, shows that in some cases, there were no other reliable sources of information covering the same event as the stories published by the source. The **Lede present** indicator, with an average score of 68, suggests that not all of the sites contained an opening paragraph – or subheadline – summarising the main points of the news story (typically based on the 'five Ws': Who, What, When, Where, and Why) without emotion, editorialising or analysis. The **Recent coverage** indicator, with a score of 72, highlights that, in some cases, the produced stories covered news events or developments that did not occur within 30 days prior to the article publication dates, suggesting that some domains presented outdated news content or re-used it out of its original context. The average overall score for the **Content pillar** was 75.

Figure 5 shows the average values for each site resulting from the **Content pillar** review. More than half of the sites analysed (18 out of 34) obtained values above the overall pillar average (75). The highest value was 89.67, achieved by Il Sole 24 Ore. As for the sites with below-average scores, 7 sites received scores with values above 70. The worst-performing sites achieved scores of 48.33 and 48.39. These sites performed particularly poorly on the indicators **Article bias**, **Negative targeting** and **Sensational language**, suggesting non-neutral reporting, and extensive favouring of some actors and organisations, as well as denigration of others in their news coverage. These two sites are significantly behind the third-worst performing domain, which received a score of 61.56.

Figure 4. Average Content pillar scores by indicator

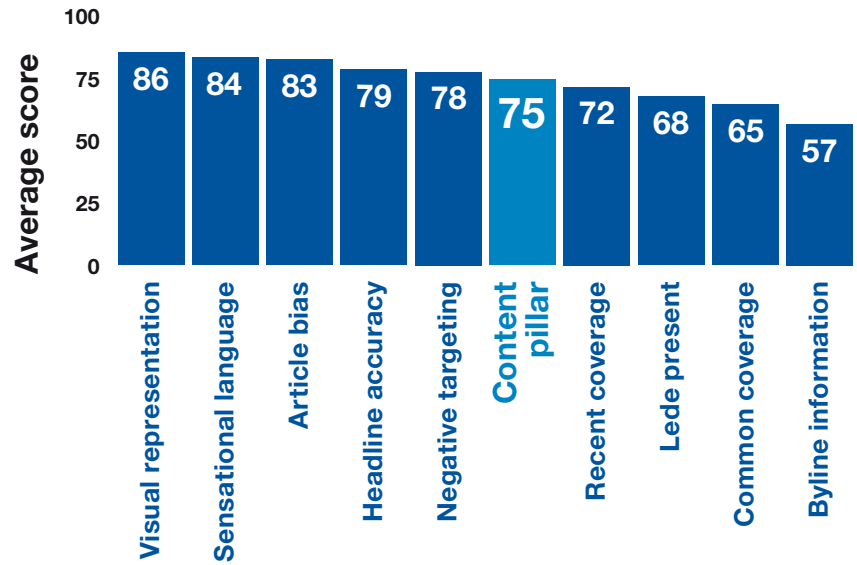
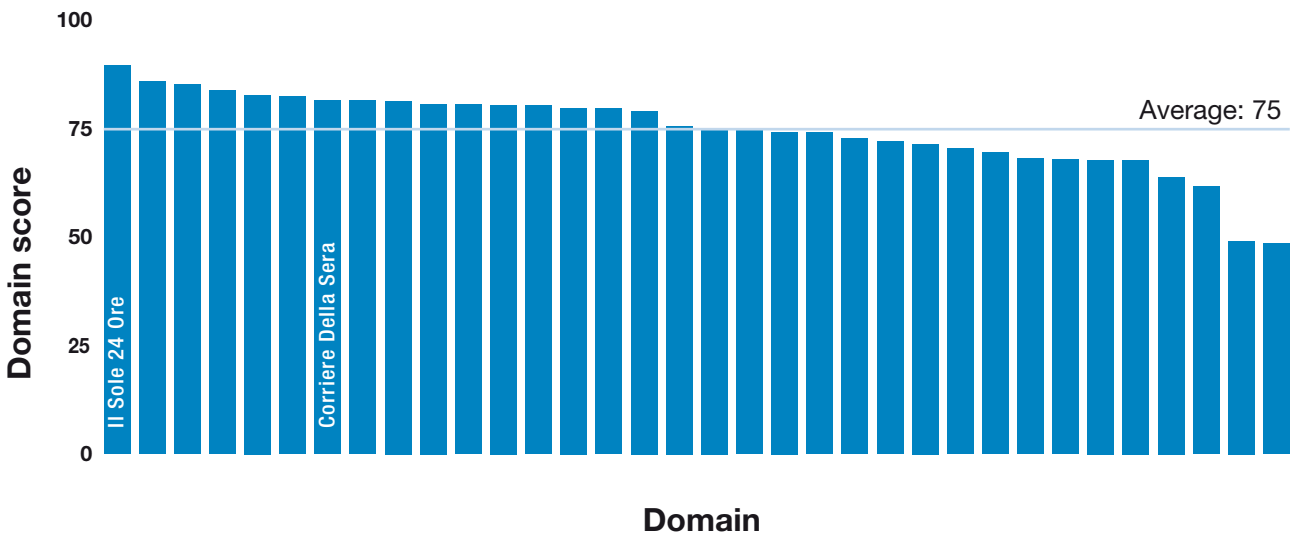


Figure 5. Content pillar scores by site



Operations pillar

This pillar assesses the operational and editorial integrity of a news site. All scores are based on a scale of zero (worst) to 100 (best), as scored by the country reviewers according to the information available on the site. The **Operations pillar** looks at whether relevant policies are made transparent to the public. It is not able to measure whether or how well the policies are implemented.

All 34 sites in our sample have the potential to score perfectly on all the indicators of the **Operations pillar** if they transparently disclose such operational policies and information. The indicators for the **Operations pillar** are taken from the standards which have been set by journalists as part of the Journalism Trust Initiative (JTI).²⁰ As the JTI points out, adopting these standards raises credibility in the eyes of the public, compels traditional media to reassess their practices in the digital age, and encourages news media outlets to be more transparent about their business models.

From Figure 6, it can be seen that the only indicator for which the operational review gives an encouraging result relates to the **Ownership** of the media outlet. This indicator covers transparency in declaring who owns the site and how it is organised (i.e., board of directors, editor-in-chief and editorial staff). For the **Ownership** indicator, 20 of the 34 sites scored 50 or higher, while 11 sites scored 75 or higher. This means that the companies and organisations that own the sites could typically be identified on the websites (mainly from the contact section). It is worth noting that there are few major publishing groups in Italy. If a news website in our sample is owned by one of these groups, this is usually stated explicitly. However, the average score for this indicator shows that a significant number of Italian domains still do not publish sufficiently clear information regarding their owners.

Domains performed significantly worse on the remaining **Operations pillar** indicators, including **Funding**, meaning that the site includes a declaration of the media outlet's main sources of income, with an average score of 41; the declaration of adherence to fundamental principles of journalism (**Editorial principles and practices**), with an average score of 40; the presence of an explanation of the rules governing the management of comment sections (**Comment policies**), with an average score of 30; pre- and post-publication fact checking (**Ensuring accuracy**), with an average score of 12; and **Attribution**, which assesses the presence of policies for attributing the author, source and media contained in an article, with an average score of 11. As a result of low scores on individual indicators, the average value for the **Operations pillar** is 31.

Low scores for **Operations pillar** indicators can significantly undermine readers' confidence in a website, due to 1) a lack of transparency regarding the news site's revenue, which does not allow readers to determine that there are no conflicts of interest in a domain's coverage, 2) a lack of information regarding the editorial principles and practices which are meant to guide the domain's coverage and instill trust in the readers, and 3) insufficient information regarding attribution policies and pre- and post- publication fact-checking policies, which hinders accountability for erroneous or poor reporting.

Figure 7 shows the results of the overall score obtained by each site for the **Operations pillar**. The average score for the sites included in the sample, as mentioned above, is 31. Only 4 sites receive an average score above 50. The highest value is 60.64, achieved by Corriere Della Sera.

Figure 6. Average Operations pillar scores by indicator

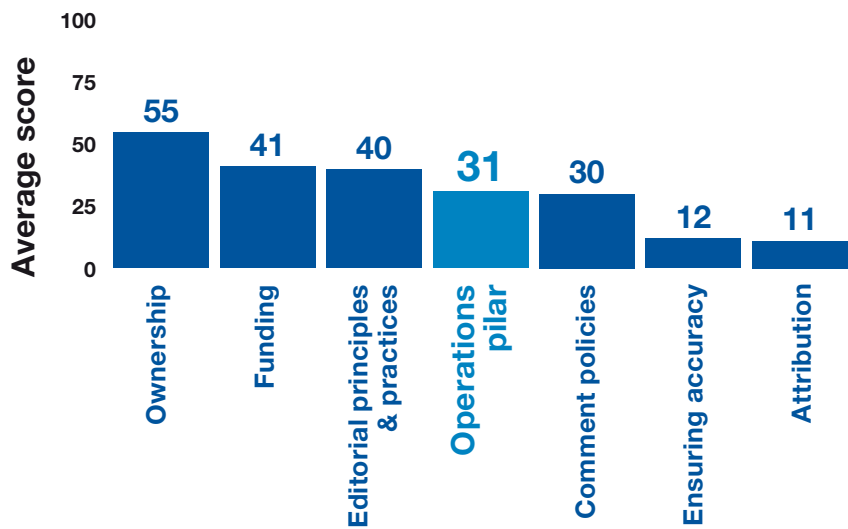
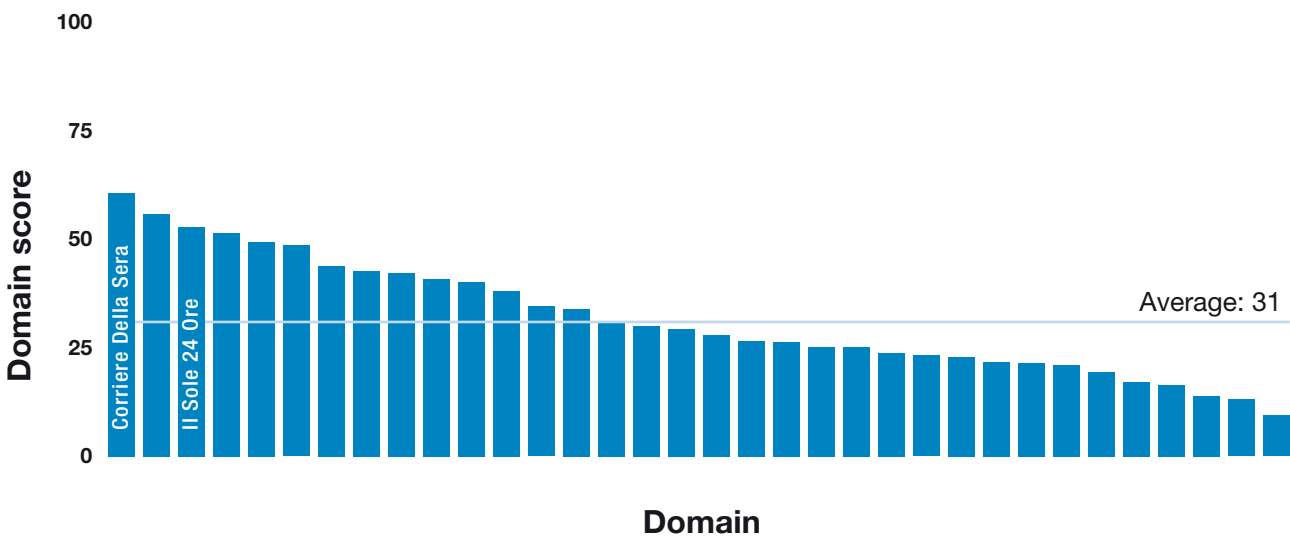


Figure 7. Operations pillar scores by site



Conclusion

Our assessment of the disinformation risk of news sites in Italy finds an array of risk classifications.

Taking the content and operational review results together, 22 of the 34 sites analysed in this study have a medium- to high-risk level. Only two sites achieved a minimum-risk rating, while three sites performed very poorly, receiving a maximum-risk rating.

Italy's media sites typically demonstrate low-risk in GDI's framework when it comes to indicators that assess the reliability, neutrality and relevance of content published. Based on our findings, however, domains should attempt to consistently provide timely coverage of recent events which are of interest to the broader public, while also implementing a consistent use of ledes. Still, these domains' overall ratings are brought down by operational shortcomings, especially regarding transparent information about **Funding**, **Editorial principles and practices**, **Comment policies**, pre- and post-publication fact-checking, and **Attribution** policies.

News sites could address these shortcomings by taking concrete actions, such as:

- Adopt and publish journalistic and operational standards like those set by the Journalism Trust Initiative that make transparent information about overall policies of the site
- Clearly publish their sources of funding on their page rather than a parent company site; this information helps to build trust in the site and dispel doubts about how it is funded
- Publish a statement of editorial independence
- Ensure that there are clear policies around moderation for sites that have commenting available to readers
- Improve and make more visible a site's practices for correcting errors; it is important that such site corrections are clearly seen and understood, rather than being hidden on a web page 'below the fold'
- Publish bylines and policies regarding content attribution; publishing the identity of the author or a policy as to why this information was withheld is an easy way to ensure transparency and accountability
- Ensure that every story includes a fact-based lede that immediately gives the reader a good sense of the overall story.

Appendix: Methodology

The Global Disinformation Index evaluates the level of disinformation risk of a country's online media market. The country's online media market is represented by a sample of 30 to 35 news domains that are selected on the basis of their Alexa rankings, their number of social media followers, and the expertise of local researchers. The resulting sample features major national news sites with high levels of online engagement, news sites that reflect the regional, linguistic and cultural composition of the country, and news sites that influence ideas among local decision-makers, groups or actors.

The index is composed of the **Content** and **Operations pillars**. The pillars are, in turn, composed of several indicators. The **Content pillar** includes indicators that assess elements and characteristics of each domain's content to capture its level of credibility, sensationalism, and impartiality. The **Operations pillar's** indicators evaluate the policies and rules that a specific domain establishes to ensure the reliability and quality of the news being published. These policies concern, for instance, conflicts of interest, accurate reporting and accountability.

Each of GDI's media market risk assessments are conducted in collaboration with a local team of media and disinformation experts who develop the media list for the market sample, contribute to the sampling frame for the content included in the **Content pillar** review, conduct the data collection for the **Content** and **Operations pillars**, vet and interpret the index results, and draft the market report.

Site selection

The market sample for the study is developed based on a mix of quantitative and qualitative criteria. GDI begins by creating a list of the 50 news websites with the greatest traffic in the media market. This list is provided to the country research team, along with data on the number of Facebook and Twitter followers for each site, to gauge relevance and reach. The local research team then reduces the list to 35 sites, ensuring that the sample provides adequate geographic, linguistic and political coverage to capture the major media discourses in the market. International news outlets are generally excluded, because their risk ratings are assessed in the market from which they originate.²¹ News aggregators are also excluded, so that all included sites are assessed on their original content. The final media market sample reflects the complete set of between 30 to 35 sites for which complete data could be collected throughout the review process.

Data collection

The **Content** indicators are based on the review of a sample of ten articles published by each domain. Five of these articles are randomly selected among a domain's most frequently shared articles on Facebook within a two-week period. The remaining five articles are randomly selected from a group of a domain's articles covering topics that are likely to carry disinformation narratives. The topics, and the associated set of keywords used to identify them, are jointly developed by GDI and the in-country research team. Each country team contributes narrative topics and the keywords used to identify them in the local media discourse to GDI's global topic classifier list, developed by GDI's data science and intelligence teams. Country teams also manually verify the machine translation of the entire topic list in the relevant study languages.

The sampled articles are anonymised by stripping them of any information that allows the analysts to identify the publisher or the author of the articles. The anonymised content is reviewed by two country analysts who are trained on the GDI codebook. For each anonymised article, the country analysts answer a set of 13 questions designed to evaluate the elements and characteristics of the article and its headline, in terms of bias, sensationalism and negative targeting. The analysts subsequently review how the article is presented on the domain and the extent to which the domain provides information on the author's byline and timeline. While performing the **Content pillar** reviews, the analysts are required to provide a thorough explanation and gather evidence to support their decisions.

The **Operations pillar** is based on the information gathered during the manual assessment of each domain performed by the country analysts. The country analysts answer a set of 98 questions aimed at evaluating each domain's ownership, management and funding structure, editorial independence, principles and guidelines, attribution policies, error correction and fact-checking policies, and rules and policies for the comments section. The analysts gather evidence to support their assessments as they perform each **Operations pillar** review.

Data analysis and indicator construction

The data gathered by the country analysts for the **Content pillar** are used to compute nine indicators. The **Content pillar** indicators included in the final risk rating are: **Article bias**, **Byline information**, **Common coverage**, **Headline accuracy**, **Lede present**, **Negative targeting**, **Recent coverage**, **Sensational language**, and **Visual presentation**. For each indicator, values are normalised to a scale of zero to 100. 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, and ranges from zero to 100.

For the **Operations pillar**, the answers of the country analysts are translated into a set of sub-indicators. The six indicators are calculated as the averages of these sub-indicator scores. The resulting **Operations pillar** indicators are: **Attribution**, **Comment policies**, **Editorial principles and practices**, **Ensuring accuracy**, **Funding**, and **Ownership**. For each indicator, values are normalised to a scale of zero to 100. The domain score for the **Operations pillar** is the average score across indicators.

Table 2. Global Disinformation Index pillars and indicators

Pillar	Indicator	Sub-indicators	Unit of analysis	Definition	Rationale
Content	Headline accuracy	None	Article	Rating for how accurately the story's headline describes the content of the story	Indicative of clickbait
	Byline information			Rating for how much information is provided in the article's byline	Attribution of stories creates accountability for their veracity
	Lede present			Rating for whether the article begins with a fact-based lede	Indicative of fact-based reporting and high journalistic standards
	Common coverage			Rating for whether the same event has been covered by at least one other reliable local media outlet	Indicative of a true and significant event
	Recent coverage			Rating for whether the story covers a news event or development that occurred within 30 days prior to the article's publication date	Indicative of a newsworthy event, rather than one which has been taken out of context
	Negative targeting			Rating for whether the story negatively targets a specific individual or group	Indicative of hate speech, bias or an adversarial narrative
	Article bias			Rating for the degree of bias in the article	Indicative of neutral, fact-based reporting or well-rounded analysis
	Sensational language			Rating for the degree of sensationalism in the article	Indicative of neutral, fact-based reporting or well-rounded analysis
	Visual presentation			Rating for the degree of sensationalism in the visual presentation of the article	Indicative of neutral, fact-based reporting or well-rounded analysis
Operations	Attribution	None	Domain	Rating for the number of policies and practices identified on the site	Assesses policies regarding the attribution of stories, facts and media (either publicly or anonymously); indicative of policies that ensure accurate facts, authentic media and accountability for stories
	Comment policies	Policies		Rating for the number of policies identified on the site	Assesses policies to reduce disinformation in user-generated content
		Moderation		Rating for the mechanisms to enforce comment policies identified on the site	Assesses the mechanism to enforce policies to reduce disinformation in user-generated content
	Editorial principles and practices	Editorial independence		Rating for the number of policies identified on the site	Assesses the degree of editorial independence and the policies in place to mitigate conflicts of interest
		Adherence to narrative		Rating for the degree to which the site is likely to adhere to an ideological affiliation, based on its published editorial positions	Indicative of politicised or ideological editorial decision-making
		Content guidelines		Rating for the number of policies identified on the site	Assesses the policies in place to ensure that factual information is reported without bias
	Ensuring accuracy	News vs. analysis		Rating for the number of policies and practices identified on the site	Assesses the policies in place to ensure that readers can distinguish between news and opinion content
		Pre-publication fact-checking		Rating for the number of policies and practices identified on the site	Assesses policies to ensure that only accurate information is reported
	Funding	Post-publication corrections		Rating for the number of policies and practices identified on the site	Assesses policies to ensure that needed corrections are adequately and transparently disseminated
		Diversified incentive structure		Rating for the number of revenue sources identified on the site	Indicative of possible conflicts of interest stemming from over-reliance on one or few sources of revenue
		Accountability to readership		Rating based on whether reader subscriptions or donations are identified as a revenue source	Indicative of accountability for high-quality information over content that drives ad revenue
	Ownership	Transparent funding		Rating based on the degree of transparency the site provide regarding its sources of funding	Indicative of the transparency that is required to monitor the incentives and conflicts of interest that can arise from opaque revenue sources
		Owner-operator division		Rating based on the number of distinct executive or board level financial and editorial decision-makers listed on the site	Indicative of a separation between financial and editorial decision making, to avoid conflicts of interest
	Transparent ownership	Rating based on the degree of transparency the site provides regarding its ownership structure	Indicative of the transparency that is required to monitor the incentives and conflicts of interest that can arise from opaque ownership structures		

Risk ratings

The overall index score for each domain is the average of the pillar scores. The domains are then classified on the basis of a five-category risk scale based on the overall index score. The risk categories were defined based on the distribution of risk ratings from 180 sites across six media markets in September 2020.

This cross-country dataset was standardised to fit a normal distribution with a mean of 0 and a standard deviation of 1. The standardised scores and their distance from the mean were used to determine the bands for each risk level, given in Table 3. These bands are then used to categorise the risk levels for sites in each subsequent media market analysis.

Table 3. Disinformation risk levels

Risk level	Lower limit	Upper limit	Standard deviation
Minimum risk	69.12	100	> 1.5
Low risk	59.81	69.11	> 0.5 and ≤ 1.5
Medium risk	50.5	59.8	> -0.5 and ≤ 0.5
High risk	41.2	50.49	≥ -1.5 and ≤ -0.5
Maximum risk	0	41.19	< -1.5

Endnotes

- 1 The GDI assessment framework is outlined in the annex of this report.
- 2 See <https://wearesocial.com/digital-2021>.
- 3 See <https://datareportal.com/reports/digital-2021-italy>.
- 4 See <https://www.istat.it/it/files/2019/12/Cittadini-e-ICT-2019.pdf>.
- 5 The 'early leavers from education and training' (ELET) are 18-24 year-olds with, at most, lower secondary educational attainment (i.e. ISCED 0-2 levels), who are no longer in formal or non-formal education or training.
- 6 See <https://op.europa.eu/webpub/eac/education-and-training-monitor-2020/en/chapters/chapter2.html#ch2-1>.
- 7 See http://dati.istat.it/Index.aspx?DataSetCode=DCCV_ICT.
- 8 See <https://www.ipsos.com/sites/default/files/ct/publication/documents/2020-10/the-future-of-trust-in-media-graphic-reference-report.pdf>.
- 9 See [https://www.europarl.europa.eu/RegData/etudes/IDAN/2021/653641/EXPO_IDA\(2021\)653641_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/IDAN/2021/653641/EXPO_IDA(2021)653641_EN.pdf).
- 10 See https://s3.amazonaws.com/avaazimages.avaaz.org/Networks_Report_Update_Page_July_2019.pdf.
- 11 Sciubba Caniglia, C. (2020). Signs of a new world order: Italy as the COVID-19 disinformation. Harvard Kennedy School (HKS) Misinformation Review, 1(3).
- 12 See <https://euvsdisinfo.eu/report/the-eu-left-italy-face-to-face-with-an-invisible-and-terrible-enemy/>.
- 13 See <https://dgap.org/en/research/publications/eu-response-disinformation-russia-covid-19>.
- 14 Guido Caldarelli, Rocco De Nicola, Marinella Petrocchi, Manuel Pratelli, Fabio Saracco: *Flow of online misinformation during the peak of the COVID-19 pandemic in Italy*. EPJ Data Sci. 10(1): 34 (2021). See <https://dblp.org/pid/183/6149.html>.
- 15 See <https://www.coe.int/en/web/media-freedom/home>.
- 16 See <https://www.rcmediafreedom.eu/Tools/Legal-Resources/SLAPPs-the-Italian-Case>.
- 17 See <https://www.ecpmf.eu/slapp-the-background-of-strategic-lawsuits-against-public-participation/>.
- 18 See <https://www.agcom.it/documents/10179/19267334/Allegato+6-7-2020+1594044962316/36cae229-dcac-4468-9623-46aab47964f?version=1.0>. Figure 13.
- 19 See <https://www.agcom.it/documents/10179/19600536/Allegato+31-7-2020/56faf730-d62e-4cbf-a0b7-7fb60fa0fa70?version=1.0>.
- 20 For more information on the JTI, which has adopted an ISO standard for the industry, please see: <https://jti-rsf.org/en/>.
- 21 In select cases, international news outlets may be included in a study if the domestic market is small, the sites are considered highly relevant, the content on the site is specific to the market assessed, and GDI has not developed a risk rating for that site elsewhere.



www.disinformationindex.org