

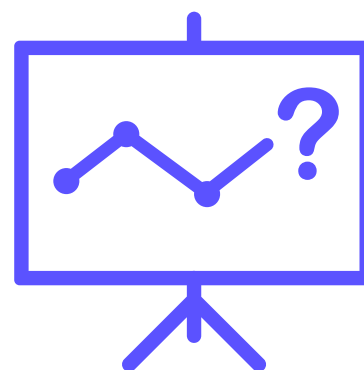
How Well Can AI Software Capture Brands' Media Sentiment?

Interpreting your media coverage without sentiment analysis is like trying to paint in color using a black and white reference image.

Sure, you could make an educated guess based on the photo's grayscale values and your preexisting knowledge of the colors of the sky and the earth, and so on. But you'll likely get some of it wrong and miss out on the invaluable nuances that can illuminate essential information about the scene. Like the time of day (the subtle hues of the sky could indicate the difference between dusk and dawn) or the season (think autumnal oranges versus vibrant spring greens).



Now imagine you're evaluating a messaging campaign by analyzing your earned media coverage. Measuring coverage volumes by topic can only tell you so much. Without knowing the tone of that coverage, how can you determine whether the campaign was a success or a failure?



You may think using an AI tool to perform sentiment analysis will improve your odds of getting it right. But experts explain that relying solely on a tech tool to measure sentiment “can be like flipping a coin – or **only 50% accurate** – since these platforms often struggle to measure more nuanced posts or are unable to filter and interpret the information through the lens of a company.”

So, how do you know if AI sentiment analysis is really “good enough” to evaluate your work?

The Test

We put AI sentiment analysis to the test by comparing a human-analyzed dataset – the gold standard for accuracy in sentiment analysis – with a leading provider’s AI analysis of the same coverage.

We had our experts analyze hundreds of articles mentioning one of our clients for sentiment towards the company. We then performed a comparative analysis against the leading provider’s data to highlight any discrepancies in tone assignment.



The leading provider has two methods for applying sentiment: to the article level and the entity (company) level. We’ve based our comparison on the leading provider’s entity-level analysis because it’s a more directly comparable methodology and was more accurate.

Here’s what we found:

The leading provider’s entity-level toning was accurate 46% of the time, and its standard article-level toning was 44% accurate.

Essentially, if you’re relying exclusively on the leading provider’s automated sentiment, you’re working with incorrect data **more than half** the time.

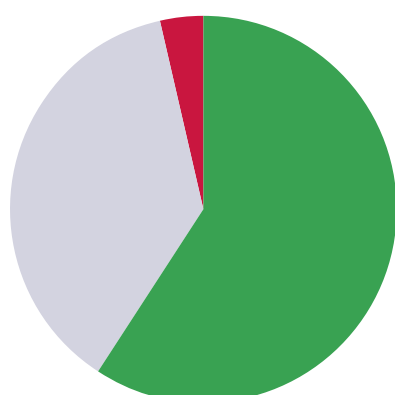
Let’s zoom in on the details to understand how this could skew an assessment of your PR work:

In overall results, we found **59%** of our client’s coverage positive, **compared with the leading provider’s finding of 33%.**

Overall Company Media Coverage by Tone

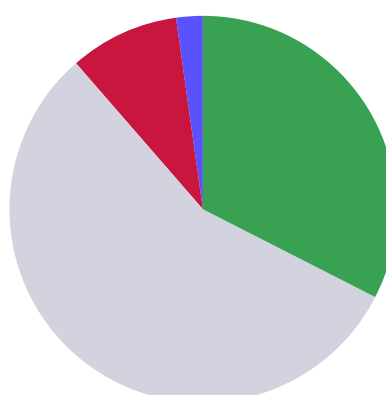
Analyzed by Trained Humans

PublicRelay



Analyzed by AI

Leading Provider



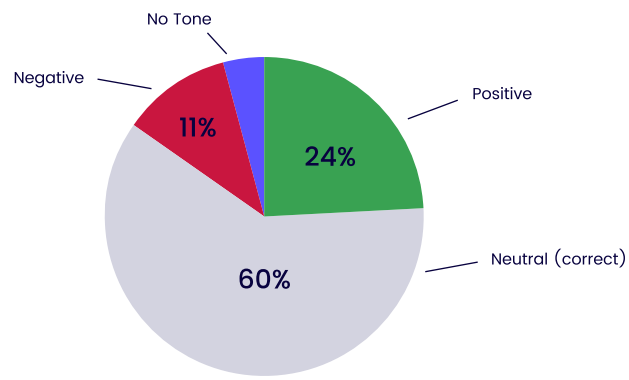
Client's Neutral Coverage

As toned by trained human analysts

For neutral coverage (as assigned by a trained human analyst), the AI toning was the most likely to be correctly toned (60%). However, the leading provider's AI mislabeled 24% of neutral coverage as positive and 11% as negative.

Tone Assigned to Neutral Stories by AI

If toned correctly, the entire pie should be neutral.



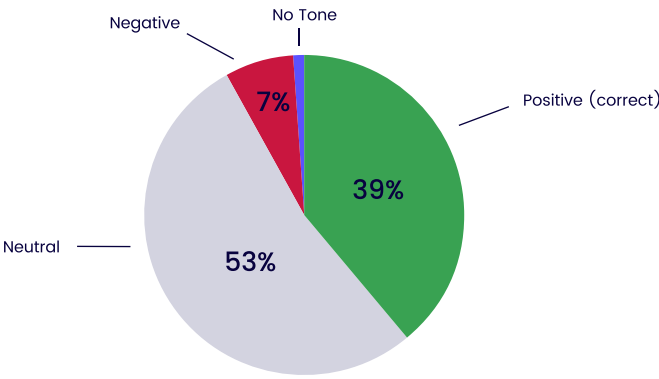
Client's Positive Coverage

As toned by trained human analysts

For positive coverage (as assigned by a trained human analyst), the AI toning was correct 39% of the time. However, the leading provider's AI mislabeled 53% of positive coverage as neutral and 7% as negative. In other words, your team wouldn't get credit for nearly **two-thirds** of the positive coverage you generate.

Tone Assigned to Positive Stories by AI

If toned correctly, the entire pie should be positive.



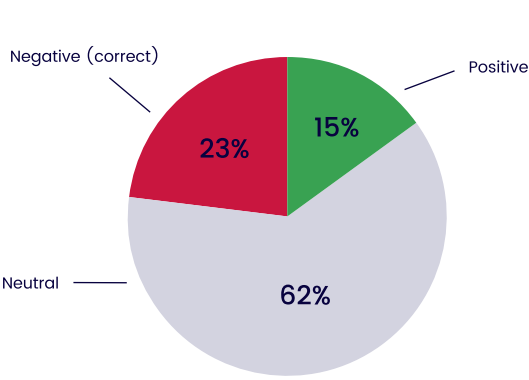
Client's Negative Coverage

As toned by trained human analysts

For negative coverage (as assigned by a trained human analyst), the AI toning was correct 23% of the time. However, the leading provider's AI mislabeled 62% of negative coverage as neutral and 15% as positive. Such an error could lead your team to ignore potential PR crises or mistakenly claim credit for what turns out to be a negative story.

Tone Assigned to Negative Stories by AI

If toned correctly, the entire pie should be negative.



The leading provider's system showed a strong bias towards neutral sentiment and **mislabeled more than half** of both negative and positive coverage as neutral.

Finally, AI found **2.5x more negative mentions** and about **half as many positive mentions** as trained human analysts. Overall, the leading provider's technology could make a successful quarter look bad.



Other Considerations

We also found that **the leading provider's software cannot perform key message identification** or recognize conceptual topics like ESG, innovation, thought leadership, and DE&I. The reason is that artificial intelligence relies on Boolean strings and their ability to capture concepts using binary logic (i.e., categorizing all information as either "true" or "false"). This approach can also lead to an over-inclusion of irrelevant coverage.

Lastly, the leading provider can only apply sentiment at the article and entity level. But high-level sentiment analysis limited to article and entity toning fails to capture the nuances of the multiple messages appearing in a single article. It means that if an article discusses more than one reputational driver, it can't distinguish variations in tone across them.

The PublicRelay Difference

Our capabilities allow us to apply sentiment not only to entities but also to topics and key messages associated with those entities through our **relational tagging system**. In a single article, we could identify that the overall discussion about the company was positive due to favorable mentions of the CEO, thought leadership, and ESG practices, despite a minor negative mention of a workplace issue.

With this **depth of data**, our dashboards and reports could show you that the company's coverage was positive overall while allowing you to isolate each topic and sub-topic associated with the brand. For instance, you could determine what percentage of its total coverage positively mentioned its ESG practices versus innovation.

Why Human-Augmented AI?

While AI programs save time and improve scalability, **fully automating your sentiment analysis can compromise its accuracy.**

If you're working with 50/50 odds of accuracy when compiling the full picture of your media coverage, you risk misinterpreting positive coverage as negative. Your successes as failures. Dusk as dawn. Spring as Autumn.

And are you really comfortable telling your C-suite that you are only 50% confident in the results of your work?

Despite its advantages, even the most advanced artificial intelligence can't keep pace with social contexts and meanings that can change overnight. Yet **human analysts can** identify abstract ideas and concepts and **ensure highly accurate sentiment analysis** by applying a human understanding of context, linguistic devices, and nuance.

That's why a combined approach of **human-augmented technology** can offer the best of both worlds. ▼

