

WHITE PAPER

# AI ASSISTANCE FOR REPUTATION MANAGEMENT

---

This white paper is written by Sune Debel, data scientist at Hypefactors.

It's about media reputation management: What it is, and how Hypefactors is using AI to improve the ability of businesses and brands to react and adapt to what is being said about them in social and news media.

---

# AI Assistance for Reputation Management

**A look at how the limitations of sentiment analysis for media reputation management can be overcome, and how Hypefactors is pushing the state-of-the-art in this field.**

Monitoring what people and media are saying about your organization is the first step in avoiding the proverbial shitstorm and taking advantage of opportunities for improving your reputation.

Most organizations recognize that reputation management is an integral part of running a successful business. There is no shortage of examples of companies that have suffered significant losses in both revenue and market value as a direct consequence of changes in public opinion about their brand in the negative direction. At the same time, the ability to cultivate a positive public opinion about your brand is a strong competitive advantage. Something similar can be said for the public opinion of political and governmental organizations. Monitoring what people and media are saying about your organization is the first step in avoiding the proverbial shitstorm and taking advantage of opportunities for improving your reputation.

In the age of social media and digital news however, monitoring what is being said about your brand on a global scale is becoming increasingly difficult. This is true both because of the sheer number of media mentions an organization must read and react to, but also because of the multilingual and multicultural nature of the global media landscape.

As a consequence, many organizations are adopting software tools to assist them with tackling PR and reputation management at scale. Hypefactors provides a state-of-the-art PR automation software that replaces time-consuming manual work with intelligent automation. It's a comprehensive one-stop platform, driven by powerful Artificial Intelligence (AI) that automates Public Relations and Earned Media tasks.

In this white paper we will review how AI can help you identify threats and opportunities for reputation management of your brand, and how Hypefactors is improving the state-of-the-art in this field. Our focus will be on one of the most widely used AI techniques for reputation management: sentiment analysis.

# Sentiment Analysis

Sentiment analysis is a collection of techniques for automatically detecting subjective opinions and feelings

Sentiment analysis is a collection of techniques for automatically detecting subjective opinions and feelings expressed by some source towards some target in natural language. Online product and service reviews is a good example of a domain where sentiment analysis can be made useful, since the source and the target are easily identified as the author and the subject of the review respectively.

Take for example the following review of the Hypefactors app taken from the review site capterra.com:

*“Really easy interface and configuration. The customer service team is very helpful and the set-up was done in one day.”*

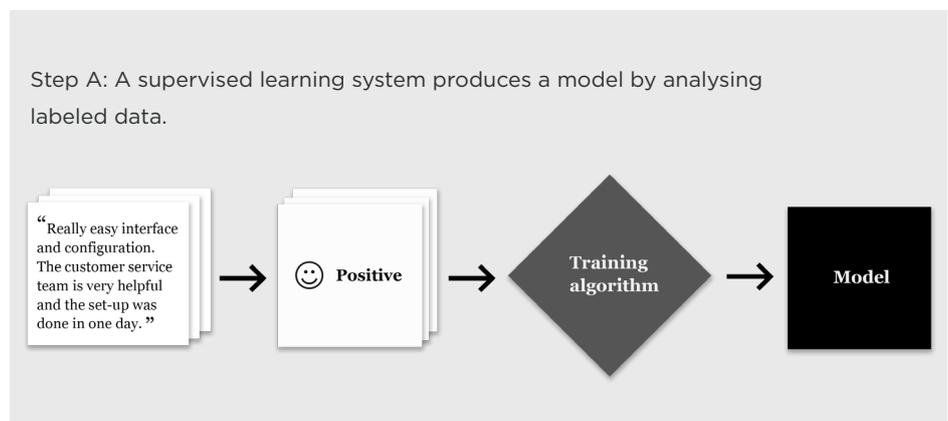
The subjective opinion expressed by the author towards Hypefactors is positive in this case, as indicated by phrases such as easy interface and configuration and helpful customer service team.

## Machine Learning techniques for Sentiment Analysis

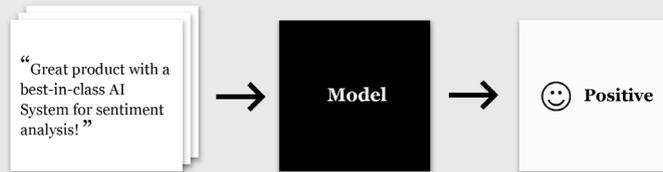
Most modern sentiment analysis systems are based on machine learning

Using a large collection of reviews and associated sentiment labels, it's possible to train a machine learning system to recognize such linguistic indicators of sentiment, and use them to accurately predict whether the sentiment of any review is negative, neutral or positive. Most modern sentiment analysis systems are based on a machine learning technique called supervised learning. In this learning paradigm, the system is produced by a process referred to as training in which a training algorithm is applied to a large collection of inputs (text in this case) and desired associated outputs. The output of training is a model. The model is a mapping from text

**Figure 1: Machine learning for sentiment analysis**



Step B: Once the model is trained it can be used to classify new data automatically.



to its most likely label. Once the model is trained it can be used to classify new, unlabeled data without human input. Once the model is trained it can be used to classify new data automatically.

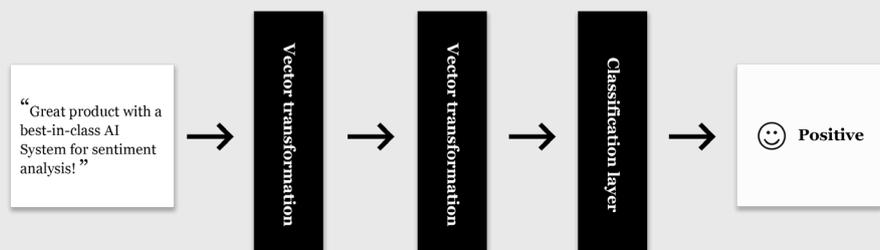
Deep learning models are by far the most common models used in supervised learning for natural language processing problems such as sentiment analysis.

Today, deep learning models are by far the most common models used in supervised learning for natural language processing problems such as sentiment analysis. Deep learning is an umbrella term that covers a number of different techniques with the common trait that the model consists of a series of nonlinear vector-to-vector transformations of the input, until the output of the final vector transformation is used as input to a classification algorithm

The training algorithm for deep learning models is called backpropagation. It works by adjusting relevant parameters of the model's vector transformation layers and classification layer in a direction that improves the accuracy of the model on the labeled input data. In other words, the model learns to recognize patterns in the training data that are evidence of negative, neutral or positive sentiment.

**Figure 2: Deep learning model**

Deep learning models consists of layers of vector-to-vector transformations.



# Limitations of Sentiment Analysis for Reputation management

Sentiment analysis systems have a number of huge shortcomings as a tool for reputation management.

## **Sentiment Neutral Statements with Negative or Positive Reputational Impact**

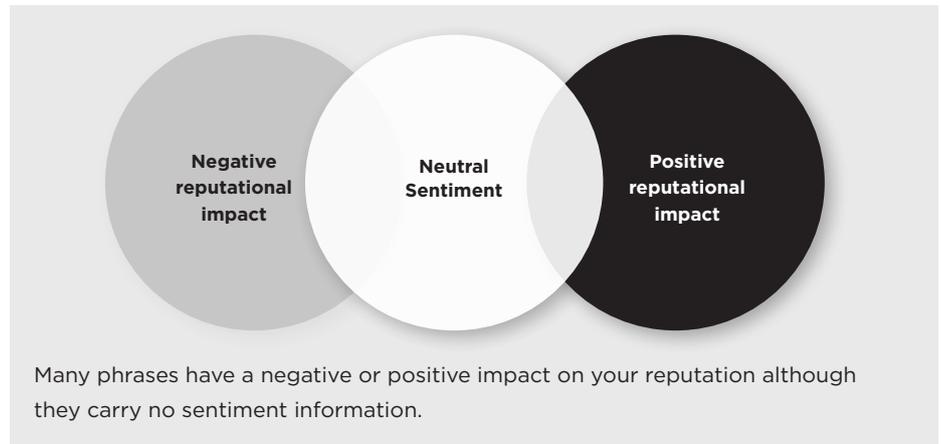
Because of the abundance of review data available online, most machine learning systems for sentiment analysis are trained on language data from this domain. Consequently, most software tools for media monitoring and reputation management provide automatic sentiment analysis of brand mentions using machine models trained on review data.

However these sentiment analysis systems have a number of huge shortcomings as a tool for reputation management. Firstly, subjective opinions and feelings expressed towards your brand represent only a fraction of the information your care about when managing the reputation of your organization. Very often, online mentions of your brand will have an impact on your reputation regardless of whether any sentiment is expressed in the text.

Consider for example the following sentence taken from a news article about the Deepwater Horizon oil spill in 2010:

*“Lingering oil residues have altered the basic building blocks of life in the ocean by reducing biodiversity in sites closest to the spill, which occurred when a BP drilling rig exploded in April 2010, killing 11 workers and spewing about 4m barrels of oil into the Gulf.”*

The sentiment of the above phrase is close to neutral: no positive or negative feelings or opinions are directly expressed by the author or any other sources, it is simply a statement of facts. However, the reported facts clearly have a negative impact on the reputation of British Petroleum (BP), and obviously it would be extremely useful if their reputation management software platform could alert them to such mentions. Traditional sentiment analysis is fundamentally useless in this situation.

**Figure 3: Reputation impact**

Most machine learning systems trained on data from this domain make no special effort to distinguish between different targets of sentiment.

### Incorrectly Targeted Sentiment

Since the target of reviewer sentiment is so easily identified in a review, most machine learning systems trained on data from this domain make no special effort to distinguish between different targets of sentiment. For this reason, the analysis provided by these models when applied to data outside of the review domain, for example news or social media data, is not granular enough to be useful for reputation management.

As an illustration, consider the following phrase taken from a news article about the incident in which a United Airlines passenger was forcibly removed from an overbooked flight in 2017:

*“After a further uproar, (even President Trump described United’s treatment of [the passenger] as “horrible”), United issued what was perceived as a more sincere apology, with [United’s CEO] promising such an incident would never happen again on United.”*

Both United Airlines and Donald Trump are mentioned in this sentence. A sentiment analysis system that does not distinguish between sentiment expressed towards different targets would most likely classify the sentence as expressing negative sentiment because of the presence of words such as horrible and uproar. This is correct if the target is considered to be United Airlines, but not if the target is considered to be Donald Trump. As a consequence, a reputation management system that assigns sentiment irrespective of which target is considered would give misleading information about Donald Trump in this case.

## **Review Language is Dissimilar to Social Media and News Language**

Sentiment analysis systems trained on review data are often much less accurate when applied to data from other domains such as news or social media because of the differences in how people express themselves in these domains: journalists generally do not express sentiment in the same way a reviewer does, which are both different from how a poster on social media expresses sentiment. As a consequence, the patterns a machine learning system trained on review data has learnt to recognize as evidence for predicting sentiment in this domain will generally not be useful for predicting sentiment in other domains.

At Hypefactors, we are pushing the-state-of-the-art of AI assisted reputation management by going beyond the limitations of sentiment analysis.

Even more problematically, most online review data is in English. For global organizations, successful reputation management requires monitoring media sources in many languages. In order to use sentiment analysis systems trained on English data exclusively, special steps must be taken that either involve costly translation of all relevant news articles and social media posts, or complex, state-of-the-art methods that allow the trained system to transfer what it has learned from one language to another. As a consequence, most reputation management systems support sentiment analysis of sources only in a small number of languages.

# **Overcoming the Limitations of Sentiment Analysis for Reputation Management**

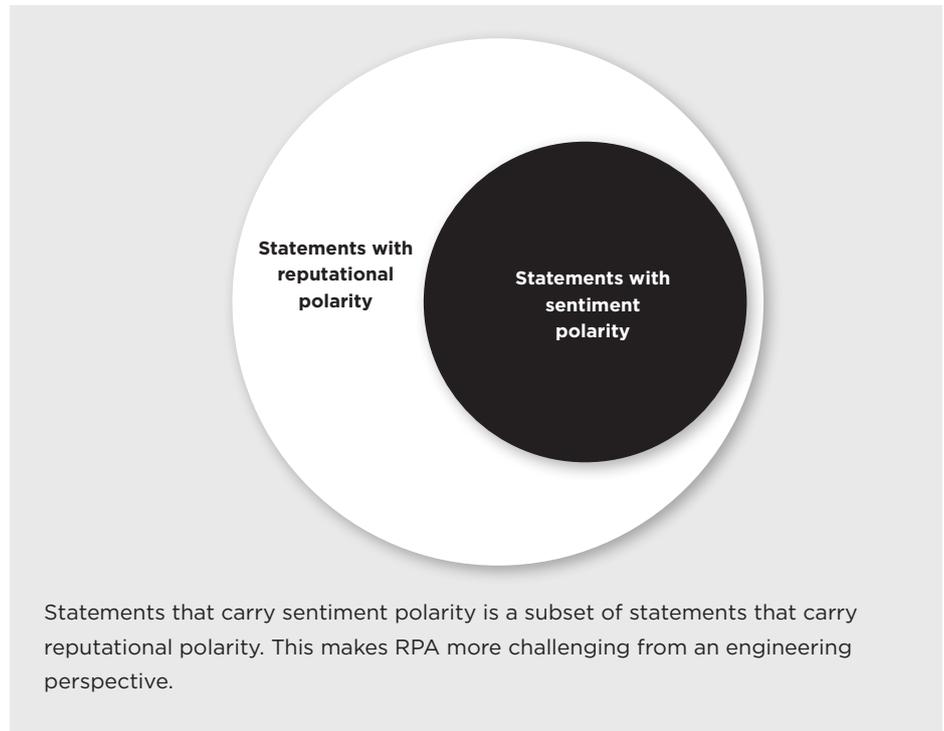
Reputation polarity analysis (RPA) is concerned with labelling text with the expected impact on the reader's perception of brand reputation in a negative, neutral or positive direction.

## **Reputation Polarity Analysis**

At Hypefactors, we are pushing the-state-of-the-art of AI assisted reputation management by going beyond the limitations of sentiment analysis. First and foremost, we are building our own curated datasets for training machine learning systems to label natural language, not with sentiment, but with what is generally known as reputation polarity.

Reputation polarity analysis (RPA) is concerned with labelling text with the expected impact on the reader's perception of brand reputation in a negative, neutral or positive direction. In contrast to sentiment analysis, RPA captures most of what you care about for the purpose of reputation management.

**Figure 4: Reputation polarity analysis**



A successful RPA system must learn to make these complicated distinctions as well, which requires a model of facts that are often not immediately expressed in the text itself.

We are using state-of-the-art techniques within a field of machine learning called transfer learning to provide reputation polarity analysis for 104 languages. This allows our customers to get the value of our AI systems even for truly global brands.

### **Challenges of Reputation Polarity Analysis**

From an engineering perspective, RPA is a much more difficult problem than sentiment analysis. An RPA system must be able to do everything a sentiment analysis system can do, since the expression of positive and negative feelings and opinions toward a brand is likely to impact the reader's perception of its reputation. In addition, an AI system for predicting reputation polarity must also be able to recognize events that are generally associated with negative changes in reputation perception such as product recalls, or events associated with positive changes in reputation perception such as charity contributions.

Finally, the impact of certain events on the reputation of your organization can depend on the type of your organization or brand. For example, if you are a tax-funded organization, throwing an extravagant party may influence your reputation negatively. If you are a private company on the other hand, the same event may not impact your reputation or may perhaps even impact it positively. A successful RPA system must learn to make these complicated distinctions as well, which requires a model of facts that are often not immediately expressed in the text itself.

### **Targeted Reputation Polarity Analysis**

In addition to tackling these difficult problems, our AI system is designed to make distinctions about different targets of reputation polarity within the same piece of text. This enables us to provide

correct labels for the United Airlines example in the previous section, depending on whether the target is considered to be United Airlines or Donald Trump.

Finally we are using state-of-the-art techniques within a field of machine learning called transfer learning to provide reputation polarity analysis for 104 languages. This allows our customers to get the value of our AI systems even for truly global brands.

## In summary

In summary, we have identified several shortcomings of traditional sentiment analysis systems offered by most PR and reputation management platforms in the market today:

- Sentiment analysis systems are fundamentally incapable of identifying statements that impact your reputation positively or negatively without expressing feelings and opinions.
- Many sentiment systems in use today does not take into account mentions of your brand that express sentiment towards something or someone else, and will in many cases imprecisely assign sentiment to the mention overall.
- The performance of many sentiment systems degrade when applied to news or social media data because these sources are linguistically different from the available training data sources.

In response to these shortcomings we have researched and redesigned an alternative approach which we believe will deliver much greater value than what is possible for other PR management products in the market. The core features of this AI system is:

- Ability to predict whether or not a particular statement will impact your brand's reputation negatively or positively even when the statement expresses no sentiment.
- Ability to distinguish between different targets of reputation polarity within the same text, and assign labels to them independently.
- Ability to assign reputation polarity labels to text sources in 104 different languages
- Trained on our in house curated datasets to provide accurate predictions on the media sources that are relevant to your brand.

Concretely, consider the following statement taken from a news article describing how Grundfos Lifelink is delivering technology that enables clean drinking water in africa:

*“There is plenty of water underground in most areas. The challenge is to make a system which is sustainable. There is no point setting up a system to supply people with water if they cannot afford to maintain it. It will simply fall into disrepair and very quickly be shut down. This is not a viable approach,” says Peter Todbjerg Hansen, Managing Director of Grundfos Lifelink.”*

Most conventional sentiment analysis systems labels the above statement as negative because of the presence of phrases such as *fall into disrepair*, *be shut down* and *not a viable approach*. Imagine that you are a PR manager at Grundfos Lifelink. The label you are provided by a state-of-the-art sentiment analysis system is technically correct, but its not useful for you to understand how the mention impacts the reputation of Grundfos Lifelink since this is the exact opposite of the sentiment analysis label, namely positive.

The Hypefactors PRA system is designed to label text with precisely this information. We believe that this improvement makes our AI system far more useful than existing solutions.

we will continue to improve and refine our AI systems for reputation management to deliver even better insights in the future.

## Conclusion

At Hypefactors we understand that reputation management is a crucial aspect of a successful PR management strategy. Our AI systems are capable of delivering insights into the public perception of your reputation at a level of detail and accuracy that, to our knowledge, is unmatched by any other solution in the market. Our extensive in-house experience with PR management enables us to identify shortcomings of existing solutions to common problems, and to deliver new solutions that are truly useful and innovative. Because AI innovation is part of our core strategy, we will continue to improve and refine our AI systems for reputation management to deliver even better insights in the future.

Hypefactors is a Copenhagen based tech company helping organizations all over the world to unlock the business potential of PR. We combine data, analytics, technology and tools to provide a unified and easy-to-use experience. We provide an AI-based PR automation platform, including all the tools to power the PR workflow and all the facts to document the PR results. All-in-one and beautifully simple.

Hypefactors deliver the transparency that companies have been missing in order to understand the value of their PR and communication investments. Incl. the financial value of every single piece of media coverage and a lot of other facts to prove and improve the Earned Media results.

[www.hypefactors.com](http://www.hypefactors.com)  
[info@hypefactors.com](mailto:info@hypefactors.com)

**hypefactors**   
Unlock the business potential of PR