

Chattermill Lab

# The Science Behind Customer Experience Intelligence

Chattermill's Approach to Feedback Analytics

Aji Ghose, PhD, VP Data & Research

Dmitry Isupov, Chief Strategy & Insights Officer

# TABLE OF CONTENTS

Introduction ..... 03

What should a good feedback analytics solution look like? ..... 03

Different approaches to feedback analysis ..... 04

Extracting insights from customer feedback:  
What's unique about Chattermill's Lyra algorithm? ..... 12

About the authors ..... 20

About Chattermill ..... 21

## Introduction

Today customers can leave feedback through any channel – including surveys, app reviews, support tickets, and social media. But it is also buried in customer conversations, including chats and calls.

Feedback from all these sources helps businesses improve their products and services. When taken further, it can drive revenue growth and profitability. However, converting this mass of feedback into actionable insights is extremely challenging.

There is also risk. If the insights derived from this feedback are not accurate, businesses may end up wasting time and resources on misguided investments based on assumptions rather than solid evidence. This can make it difficult for customer experience (CX) leaders to achieve their objectives, demonstrate the impact of CX on the overall organization and maintain a competitive advantage.

Therefore, the quality and accuracy of customer insights are of utmost importance. To fully leverage the potential of customer feedback, businesses require a comprehensive and precise feedback analytics solution that delivers highly accurate and actionable insights.

### What should a good feedback analytics solution look like?

#### **It provides transformational insights**

Collecting a large volume of surveys, chat messages, and NPS data is not sufficient to drive meaningful changes for customers. To make a significant impact, it is crucial to identify the key drivers of feedback and analyze pain points throughout the entire customer journey. An effective solution should uncover these insights, enabling businesses to influence long-term strategy and vision, ensuring sustained growth and success.

#### **It unifies customer data in one platform for everyone**

A good feedback analytics solution should allow all the stakeholders across the organization to build insightful reports and collaboratively make informed decisions with unprecedented speed. That's why it must be intuitive, easy to use, and flexible to help everyone facilitate strategic decision-making with confidence.

#### **It's built for enterprise ecosystem and agility**

With the expanding capabilities of AI and its growing integration into software, a robust feedback analytics solution should seamlessly integrate with other CX, support, product, or marketing tools while maintaining enterprise-level security and compliance standards. It must prioritize the confidentiality and security of personally identifiable information (PII), while processing large volumes of customer feedback.

# Different Approaches to Feedback Analysis

## 1. Rule-based lexicon

The rule-based lexicon is a traditional text analysis technique that employs predefined rules to categorize textual data. For example, a manually-curated lexicon assigns words and phrases with positive and negative sentiments. It scans the text, matches words with its sentiment lexicon, and calculates overall sentiment based on positive and negative word counts. Companies like Medallia and Clarabridge use rule-based approaches.

Rule-based Definition for "Price / Value for Money"

- 1. cheap OR affordable → Price [Positive]
- 2. expensive OR costly → Price [Negative]
- 3. value NEAR money within 3 words → Price [Positive]
- 4. worth NEAR price within 2 words → Price [Positive]
- 5. bad NEAR (deal OR price) within 2 words → Price [Negative]

Customer Feedback	Rule-Based Classification	Correct?
"The software subscription is affordable, and the features are excellent!"	Positive Price Perception	
"Even though the meal was costly, I left feeling hungry."	Negative Price Perception	
"The product is expensive considering the lack of innovative features."	Negative Price Perception	
"For such an affordable price, I was surprised by the high quality."	Positive Price Perception	
"Why would anyone buy such cheap rubbish - no value whatsoever!"	Positive Price Perception	
"The item is priced so low, you'd think they're giving it away, but after all the repair costs, your bank account begs to differ."	Miscellaneous	
"Paying for this feels like buying a first-class ticket for a train that's always late."	Miscellaneous	



### Advantages:

1. **Simplicity:** Rule-based lexicons offer an uncomplicated and straightforward approach to text analysis.
2. **Consistency:** Since the rules are predefined, the sentiment analysis remains consistent across various texts.

### Disadvantages:

1. **Limited Accuracy:** The main drawback of rule-based lexicons is their inability to accurately understand language nuances or contexts and extract in-depth insights on specific themes and sentiments. Any word or expression outside the system's rules will not be recognized or categorized accurately. For instance, the sentence: "The new product is not bad." A rule-based system may incorrectly categorize it as negative because the word "bad" fails to account for the negation introduced by "not."
2. **Resource-Intensive Maintenance:** Rule-based systems require constant manual input and supervision from trained linguists, analysts, or engineers to stay relevant. These systems need extensive rules to manage the unstructured nature of customer feedback, and aligning it with corresponding business processes is a laborious task. As data volume increases, scalability becomes a challenge.

Although rule-based systems had their uses in the past, they are extremely limited in capturing the subtleties of language and require substantial resources to implement and maintain. While they provide a simple and consistent approach to topic and sentiment analysis, their inability to scale effectively and understand nuanced expressions often means they need to be supplemented with more advanced methods.

## 2. Thematic Analysis and Neural Embeddings

Thematic Analysis (TA) is a conventional text analysis method rooted in Good Old-Fashioned Artificial Intelligence (GOFAI). Its origins trace back to qualitative research practices of the 1980s. This method, predominantly used in psychology, sociology, and anthropology, focuses on identifying and interpreting emergent themes or patterns in data.

Thematic Analysis (TA) includes data coding (identifying and tagging themes) and categorizing these themes accordingly. It provides flexibility, allowing themes to be spontaneously developed and adapted.



The steps involved in the integrated methodology of Thematic Analysis (TA) and Neural Embeddings.

### Advantages:

1. **Flexibility:** Thematic Analysis (TA) provides a dynamic approach to text analysis due to its on-the-fly theme creation and adaptability.
2. **Authenticity:** TA allows for unique themes and patterns to surface organically from the data, preserving its authenticity.

### Disadvantages:

1. **Time-Consuming:** Thematic Analysis (TA) heavily relies on human interpretation, necessitating manual data inspection and validating key phrases defining a theme. This process is often laborious and subject to personal bias.
2. **Limited Robustness:** Thematic Analysis (TA) is somewhat rudimentary compared to more advanced NLP techniques which use supervised fine-tuning of Large Language Models (LLMs)<sup>1</sup> and capture the context of the analyzed text.

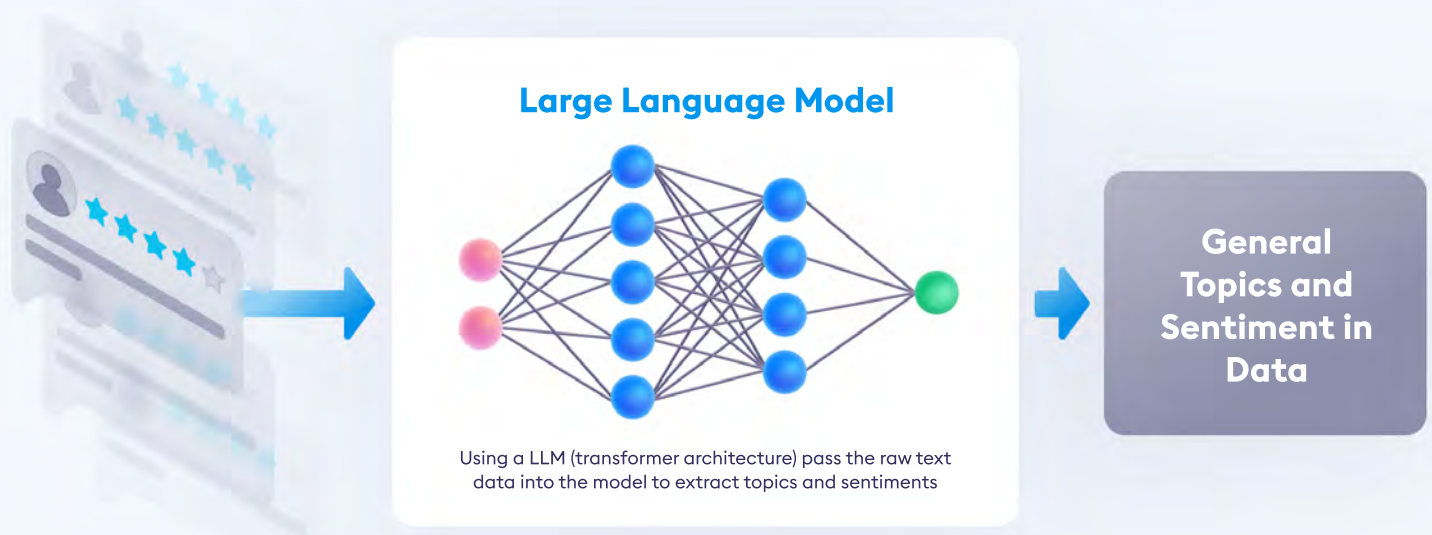
Recently, Thematic Analysis (TA) has been integrated with LLM-based neural embeddings to mitigate some of its limitations. However, this blended method still falls short compared to the performance of supervised fine-tuned LLMs for topic and sentiment categorization.

---

<sup>1</sup> LLMs (Language Model Transformers): A category of ML models used for natural language processing tasks utilizing the transformer architecture. LLMs can generate human-like text, understand the sentiment, answer questions, and perform translation, among other tasks. They are pre-trained on large corpora of text data and can be fine-tuned, using supervised or self-supervised learning, for specific applications, making them highly versatile tools in NLP.

### 3. Large Language Models (LLMs)

Large Language Models (LLMs) leverage extensive training data and capitalize on deep learning principles to generate human-like text. However, despite their sophisticated design and capabilities, LLMs do exhibit certain constraints when employed to analyze customer feedback or textual data.



Steps involved in the integrated methodology of Large Language Models (LLM).

#### Advantages:

1. **Sophistication:** LLMs stand out with their advanced deep learning techniques, enabling the generation of remarkably human-like text.

#### Disadvantages:

1. **Unreliability:** LLMs can produce inconsistent and creatively diverse responses, which may lead to untrustworthy insights.
2. **CX Contextual Inaccuracy:** These models may lack in providing accurate customer-centric context and domain-specific knowledge, yielding potentially incorrect insights.
3. **Output Variability:** The outputs generated by an LLM can differ with each execution, thereby complicating the assurance of reliable and consistent analysis of customer feedback.

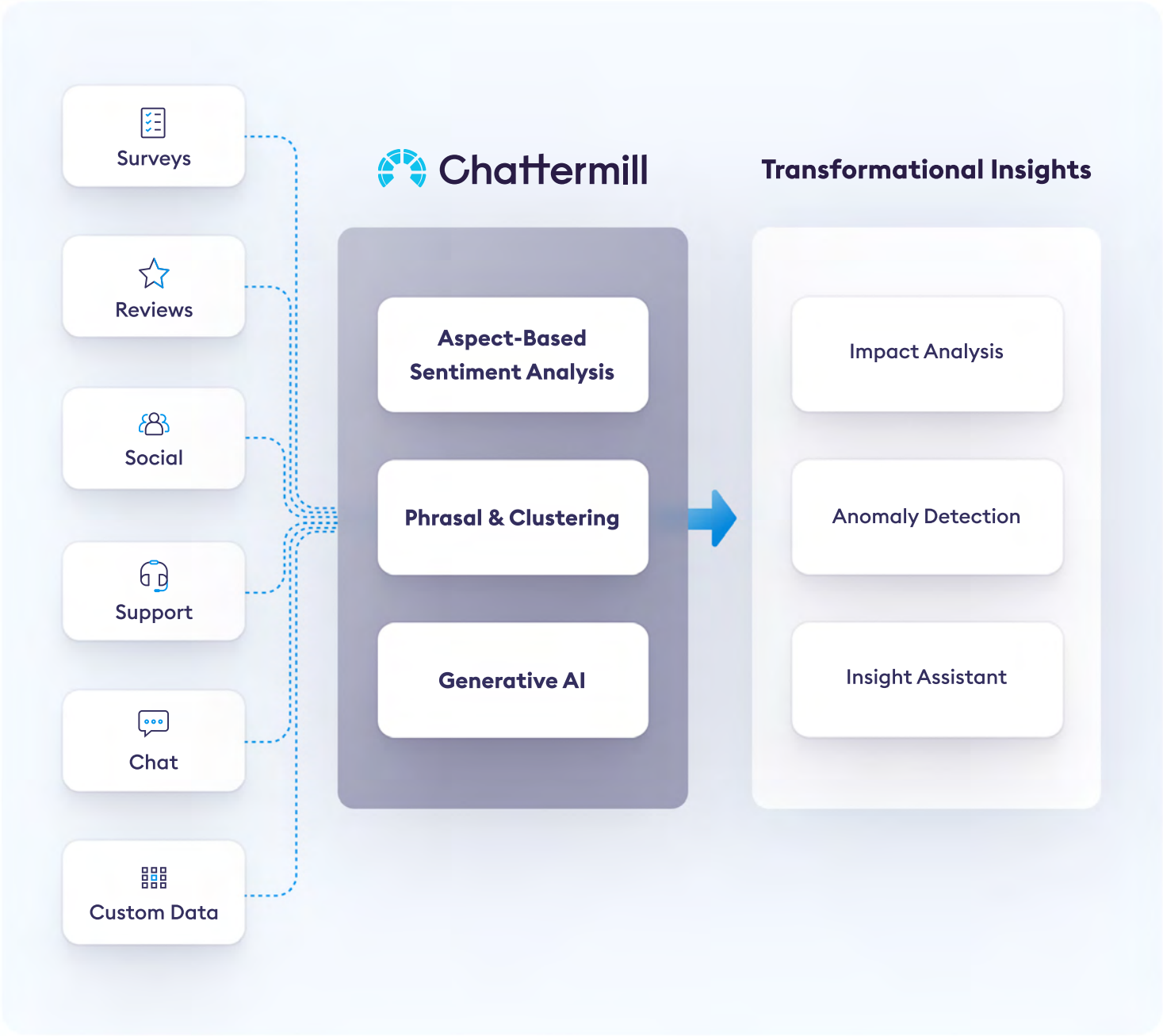
While LLMs trained for generative applications (e.g. GPT-4) succeed in generating human-like text, they do not stand as the ideal choice for generating precise and accurate summaries of customer feedback analysis.



## 4. Chattermill’s Approach to Text Analysis

Chattermill's in-house Natural Language Processing (NLP) platform, is an advanced AI system developed from over a decade of experience working alongside customer-centric businesses.

Lyra AI continually strives for high accuracy and precision in analyzing customer feedback.



### Advantages:

1. **Integration of Best Practices:** Lyra AI integrates the strengths of Aspect Based Sentiment Analysis (ABSA) with supervised and self-supervised learning. This combination allows for the nuanced capture of themes and sentiments harmonized with top-tier open-source Large Language Models (LLMs) that are tailored to handle a variety of data sources like emails, chats, feedback, and social media.
2. **Contextually Intelligent:** The system excels in pinpointing contextually intelligent and granular keyphrases, facilitating the generation of actionable insights that can drive business outcomes.
3. **Generative AI Incorporation:** Lyra AI leverages the latest Generative AI (GAI) technology to create human-readable insights, delivering a state-of-the-art approach to Customer Experience Intelligence.
4. **Customizable and Efficient:** Lyra AI provides high customization, enabling businesses to align the analysis to their specific needs and domain nuances. This reduces the maintenance burden and delivers an efficient user experience.

### Disadvantages:

1. **Dependence on Quality of Input Data:** Like any AI-based system, the effectiveness and accuracy of Lyra AI relies heavily on the quality of the input data. Poorly structured or sparse data can lead to less reliable or meaningful insights. But denoising and data cleaning can counter this.

Chattermill's Lyra AI stands out for integrating sophisticated techniques and technologies—offering a unique approach to deriving actionable, accurate and contextually-rich insights from customer feedback. As with any advanced technology, continuous improvement and updates is required to maintain its leading position in the market.

# Chattermill's Lyra AI: Multilevel NLP for Customer Experience Intelligence

	Rule-Based Lexicon	Thematic Analysis & Neural Embeddings	Large Language Models	Chattermill's Lyra AI
<p><b>Topic &amp; Sentiment Accuracy:</b> The accuracy level in classifying sentiments related to specific aspects or topics and identifying themes within customer comments respectively. Both topic and sentiment accuracy are evaluated using the metrics <i>precision</i> and <i>recall</i>. Together, they measure the model's effectiveness in sentiment and topic classification.</p> <ul style="list-style-type: none"><li>• <b>Precision</b> refers to the percentage of correct positive predictions out of all positive predictions made by the model.</li><li>• <b>Recall</b> is the percentage of correct positive predictions out of all actual positive cases in the dataset.</li></ul>	<div></div> <div></div>	<div></div> <div></div>	<div></div> <div></div>	<div></div> <div></div>
<p><b>Mutually Exclusive:</b> The ability of the model to accurately identify and represent unique CX topics without significant overlap. This ensures that similar themes are not double-counted, providing a clearer and more accurate understanding of customer feedback, allowing organizations to address customer needs more effectively and efficiently.</p>	<div></div>	<div></div>	<div></div>	<div></div>
<p><b>Collectively Exhaustive:</b> The model's ability to comprehensively capture and represent all relevant CX topics present in customer feedback. This ensures that no significant theme or issue is overlooked, providing a complete understanding of the customer's perspective and a holistic view of customer experience.</p>	<div></div>	<div></div>	<div></div>	<div></div>
<p><b>Robustness Over Time:</b> Ability to generate consistent and objective data-driven insights and ensuring they are meaningful and accurate. Additionally, this supports system maintenance by not requiring ongoing manual interventions.</p>	<div></div>	<div></div>	<div></div>	<div></div>
<p><b>Descriptive Insights:</b> The degree to which the information extracted by the AI describes the insight in detail vs needing additional manual examination of comments.</p>	<div></div>	<div></div>	<div></div>	<div></div>
<p><b>Customisation:</b> Easily being able to edit or manage tracked themes and can be tailored to any industry.</p>	<div></div>	<div></div>	<div></div>	<div></div>

Image Description: Key capabilities of different approaches to feedback analytics.

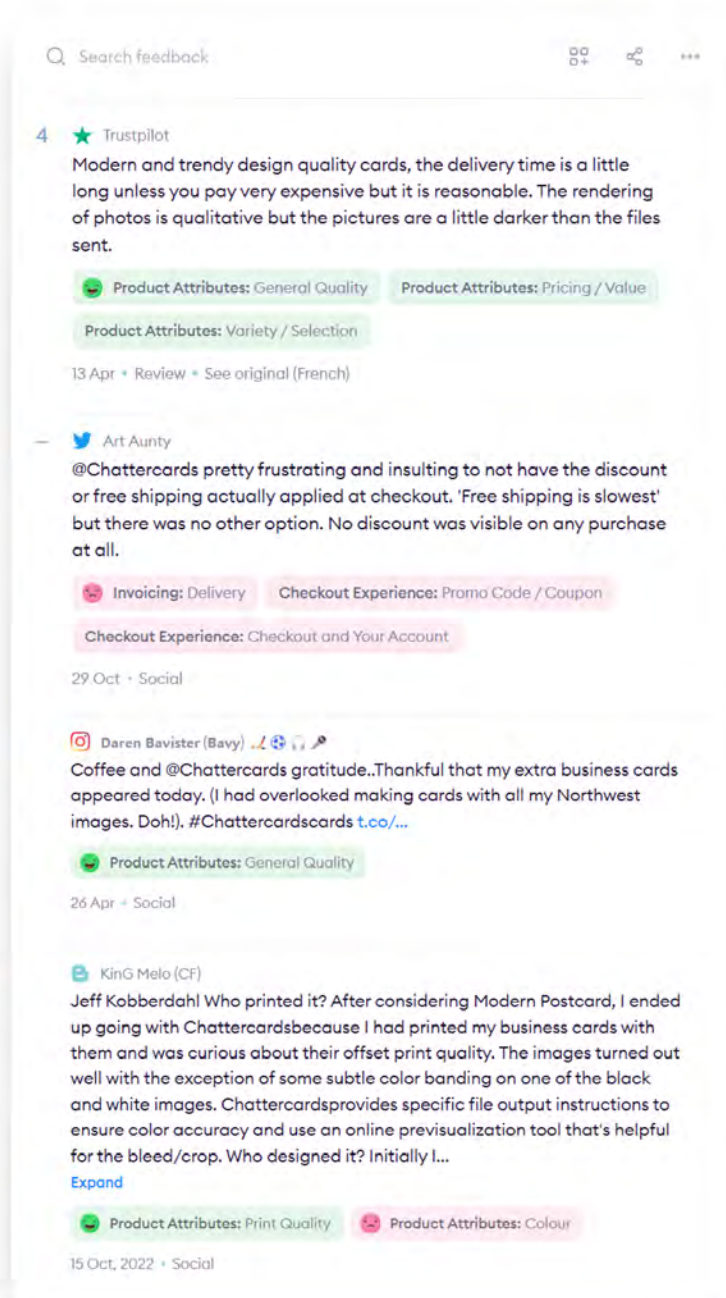
# Extracting Insights from Customer Feedback: What's Unique About Chattermill's Lyra AI algorithm?

Chattermill's Lyra AI sets itself apart from alternative approaches through the following distinctive capabilities:

## TOPIC & SENTIMENT ACCURACY

In text analysis, topics or themes refer to specific subjects or aspects discussed in the text. For instance, a restaurant review may encompass themes like food quality, service, or price. Topic accuracy evaluates how precisely an algorithm identifies and assigns topics to the given text data, assessing how well the generated topics align with the underlying themes in the text.

On the other hand, sentiment accuracy measures how effectively a sentiment analysis model can classify the sentiment and topic expressed in a piece of text. This evaluation gauges the model's ability to correctly identify and assign the appropriate sentiment label, such as positive, negative, or neutral, based on the expressed opinions, emotions, or attitudes.





## The Science Behind Customer Experience Intelligence

### Rule-Based Lexicon

Rule-based lexicons have limitations in aspect-level sentiment analysis due to their inflexibility, struggle in capturing nuances, language variations, and ambiguity. As a result, they often produce inaccuracies and misclassifications in sentiment assignments.

### Thematic Analysis & Neural Embeddings

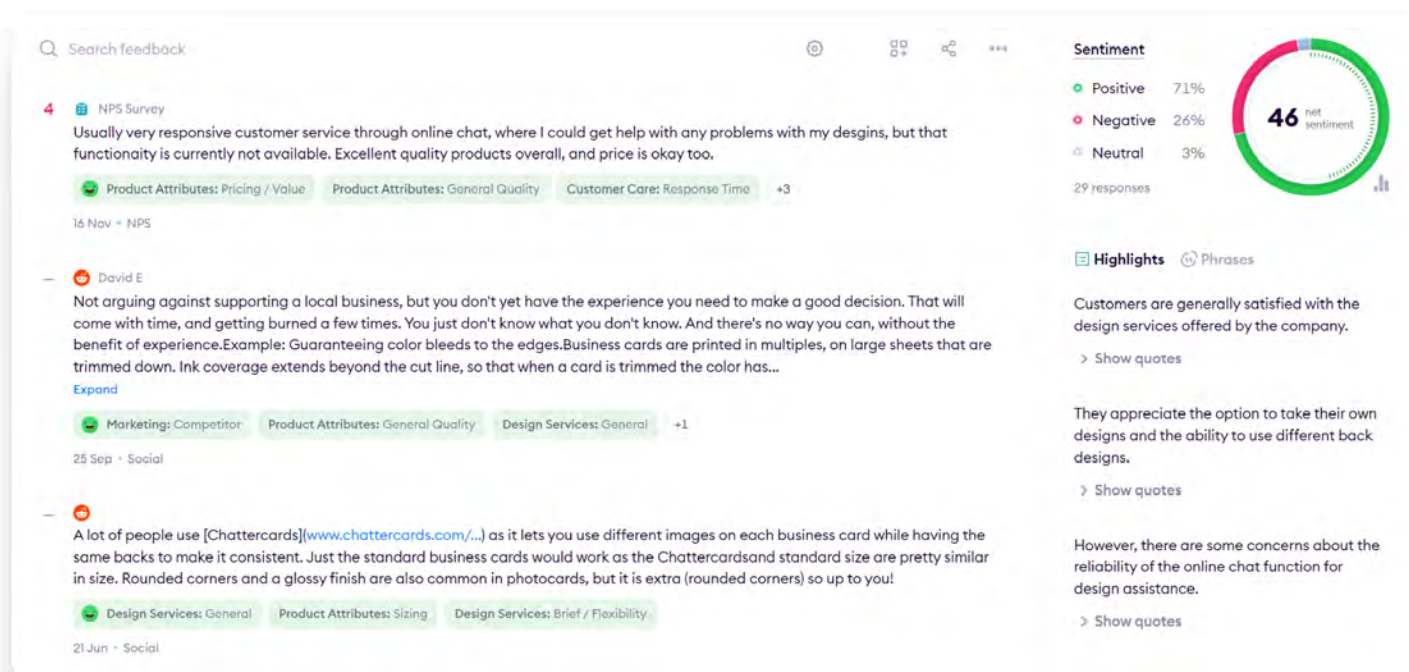
Thematic Analysis heavily relies on human interpretation. It involves researchers identifying themes based on their subjective understanding and judgement of the content. This subjective element introduces the potential for bias and inconsistency, leading to variations in topic and sentiment identification and reduced accuracy.

### Pure Generative AI (i.e. GPT-4)

While generative AI models can generate coherent and contextually relevant text, they lack fine-grained control over sentiment expression and were not designed for reliably identifying themes. They may generate text that includes mixed sentiments, ambiguous expressions, or inconsistent sentiment predictions, leading to lower sentiment accuracy. The models may also introduce unrelated or tangential information, leading to a dilution of topic accuracy.

### Chattermill's Lyra AI

Lyra AI can identify multiple themes in a single comment and assign sentiment to each one of them—ensuring high accuracy and relevance of insights. Chattermill leverages Aspect Based Sentiment Analysis (ABSA) which identifies and analyses the sentiment expressed towards specific aspects within a piece of text. It goes beyond the overall sentiment, providing a more precise and targeted understanding of sentiment at a granular level. Combined with in-depth phrasal analysis, it allows for even higher topic accuracy.





## MUTUALLY EXCLUSIVE TOPICS

In topic analysis, mutually exclusive topics refer to distinct, non-overlapping themes identified within a piece of feedback. Assigning mutually exclusive topics is essential for quickly identifying the correct root-causes of an issue identified in the feedback rather than having to investigate a group of themes which moved together due to their overlapping nature.

<b>Rule-Based Lexicon</b>	Struggle to handle cases where multiple topics overlap or when sentiments expressed in the text are associated with more than one topic. The overlapping topics can affect the interpretability and usability of the results, leading to potentially wrong business decisions.
<b>Thematic Analysis &amp; Neural Embeddings</b>	Thematic Analysis allows for flexibility in identifying themes, which can be advantageous for exploring diverse aspects of the text. However, this flexibility can also lead to the creation of multiple themes that overlap or share common elements. Analysts may introduce new themes as they uncover additional insights or connections, potentially resulting in less strict mutual exclusivity.
<b>Pure Generative AI (i.e. GPT-4)</b>	GPT models are not specifically designed for topic modelling. They excel at understanding and generating contextually coherent text, but may not offer consistency or precise control over the specific topics that are generated.
<b>Chattermill's Lyra AI</b>	Chattermill's Lyra AI excels in capturing sentiment expressions tied to specific aspects by providing fine-grained sentiment analysis within specific aspects. By ensuring that similar themes are not double-counted, Chattermill's Lyra AI provides a precise and more accurate understanding of customer feedback, allowing organizations to address customer needs more effectively and efficiently.

### Mutually Exclusive



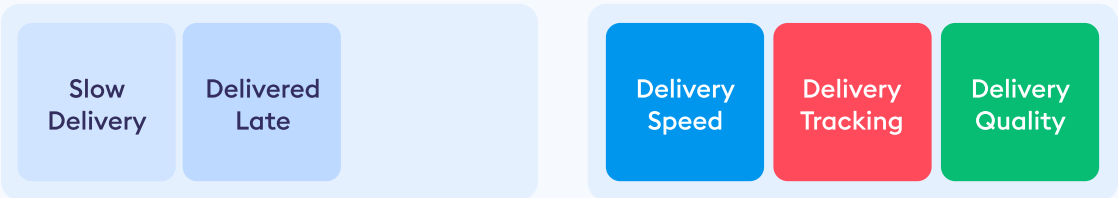
The MECE Principle (Mutually Exclusive, Collectively Exhausting).

## COLLECTIVELY EXHAUSTIVE TOPICS

Collectively exhaustive topics refer to a set of topics or categories that encompass all important aspects within a given context. In other words, if a set of topics is collectively exhaustive, it means that no significant theme or issue is overlooked, providing a comprehensive understanding of the customer's perspective.

<b>Rule-Based Lexicon</b>	Rule-based approaches rely on predefined rules which are not flexible enough to capture the full range of topics or handle emerging or nuanced aspects that were not explicitly considered during rule creation. Creating and maintaining additional rules for all possible topics is complex, brittle, and time-consuming.
<b>Thematic Analysis &amp; Neural Embeddings</b>	Although this approach provides more flexibility in adding more topics to the analysis, allowing unlimited creation of new themes without sufficient control or guidelines results in a lack of standardization. Different team members may create their own set of themes, leading to inconsistency and difficulty in aggregating or comparing results.
<b>Pure Generative AI (i.e. GPT-4)</b>	GPT models are primarily designed for language generation, not specifically for topic identification or classification. While they can generate text on a wide range of topics, they may not ensure comprehensive coverage of all relevant aspects.
<b>Chattermill's Lyra AI</b>	Chattermill's Lyra AI has the ability to comprehensively capture and represent all relevant CX topics present in customer feedback. This ensures that no significant theme or issue is overlooked, providing a holistic view of customer experience.

### Collectively Exhausting



*The MECE Principle (Mutually Exclusive, Collectively Exhausting).*

## ROBUSTNESS OVER TIME

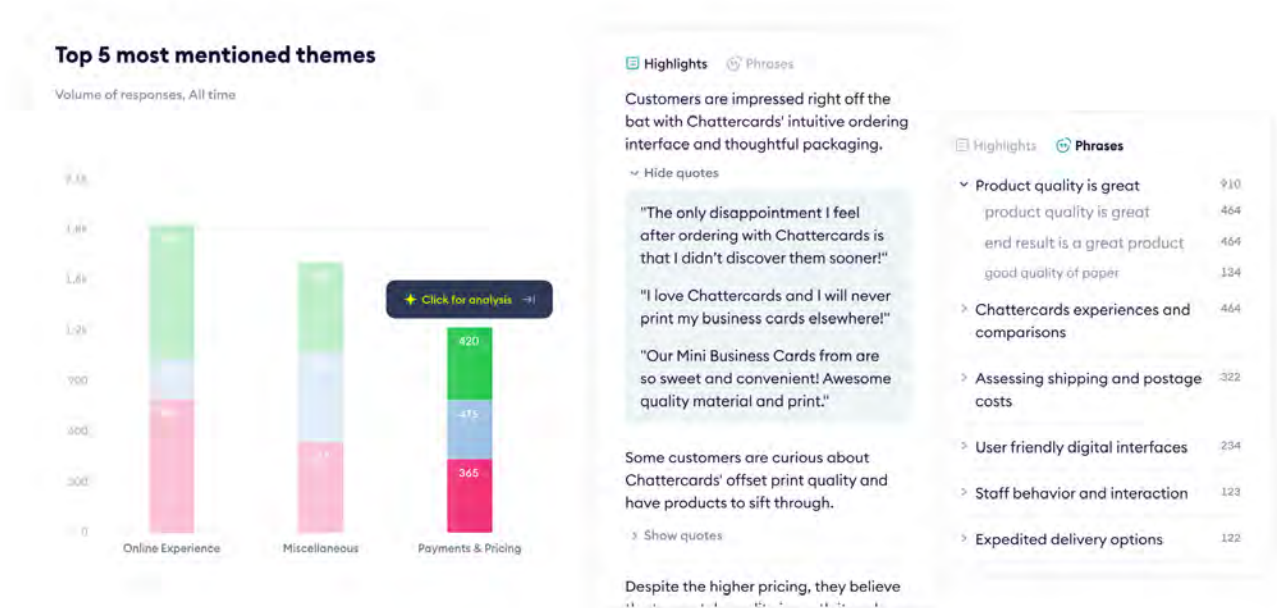
Machine learning (ML) model robustness over time is the ongoing process of monitoring, updating, and optimizing the topic and sentiment analysis model/s to ensure continued accuracy and relevance over time.

As businesses evolve, customer preferences change, and new data becomes available. Model maintenance is crucial for guaranteeing the performance and effectiveness of text analysis.

<b>Rule-Based Lexicon</b>	Rule-based systems struggle with generating descriptive insights because they strictly follow predefined rules and don't adapt to new information or context. They do not capture sentiment effectively since this requires understanding subtleties and nuances in language. Additionally, they may miss frequently mentioned topics that don't match their predefined patterns.
<b>Thematic Analysis &amp; Neural Embeddings</b>	Thematic approaches do not inherently possess the ability to learn over time. They rely on predefined themes or patterns, and any updates or changes to the themes require manual intervention. Therefore, they require a great deal of human maintenance and trial-and-error curation.
<b>Pure Generative AI (i.e. GPT-4)</b>	Generative AIs do not actively learn from new data or adapt their knowledge base without additional training of the base model from external providers such as OpenAI or Google. Moreover, this training is general and not CX-centric. Lastly, as most generative approaches require paying for each and every API call, this can become prohibitively expensive for larger datasets.
<b>Chattermill's Lyra AI</b>	Chattermill's Lyra AI model is highly scalable due to its pre-trained and self-supervised nature. It can handle large volumes of text data efficiently by leveraging its existing knowledge and linguistic understanding—without the need to maintain brittle rules as the amount of data increases. As a result, it can be applied to various applications at scale without increasing the workload or resource requirements.

## DESCRIPTIVE INSIGHTS

The ability to generate descriptive insights refers to various techniques of summarizing the sentiment distribution, identifying frequently mentioned topics or aspects, and highlighting the most common sentiment expressions associated with those topics.



Descriptive insights generated by Chattermill's Insights Assistant.

### Rule-Based Lexicon

While rule-based approaches have their own strengths in exploring and categorizing data, they are not the optimal choices for providing summaries of data due to their primary focus on identifying themes or patterns rather than summarization.

### Thematic Analysis & Neural Embeddings

While thematic analysis can contribute to understanding a dataset and provide some qualitative insights, it does not directly generate concise summaries or descriptive insights. Thematic analysis can miss nuances or misunderstand context, while neural embeddings don't naturally identify sentiment or topics. Despite combining these methods, gaps may remain, like misinterpreting sentiment tied to a topic.

### Pure Generative AI (i.e. GPT-4)

GPT models can be unreliable in providing consistent summaries of data due to their generative nature. They generate text based on learned patterns and probabilistic sampling. As a result, even with the same input, the model can produce different outputs on different runs.

### Chattermill's Lyra AI

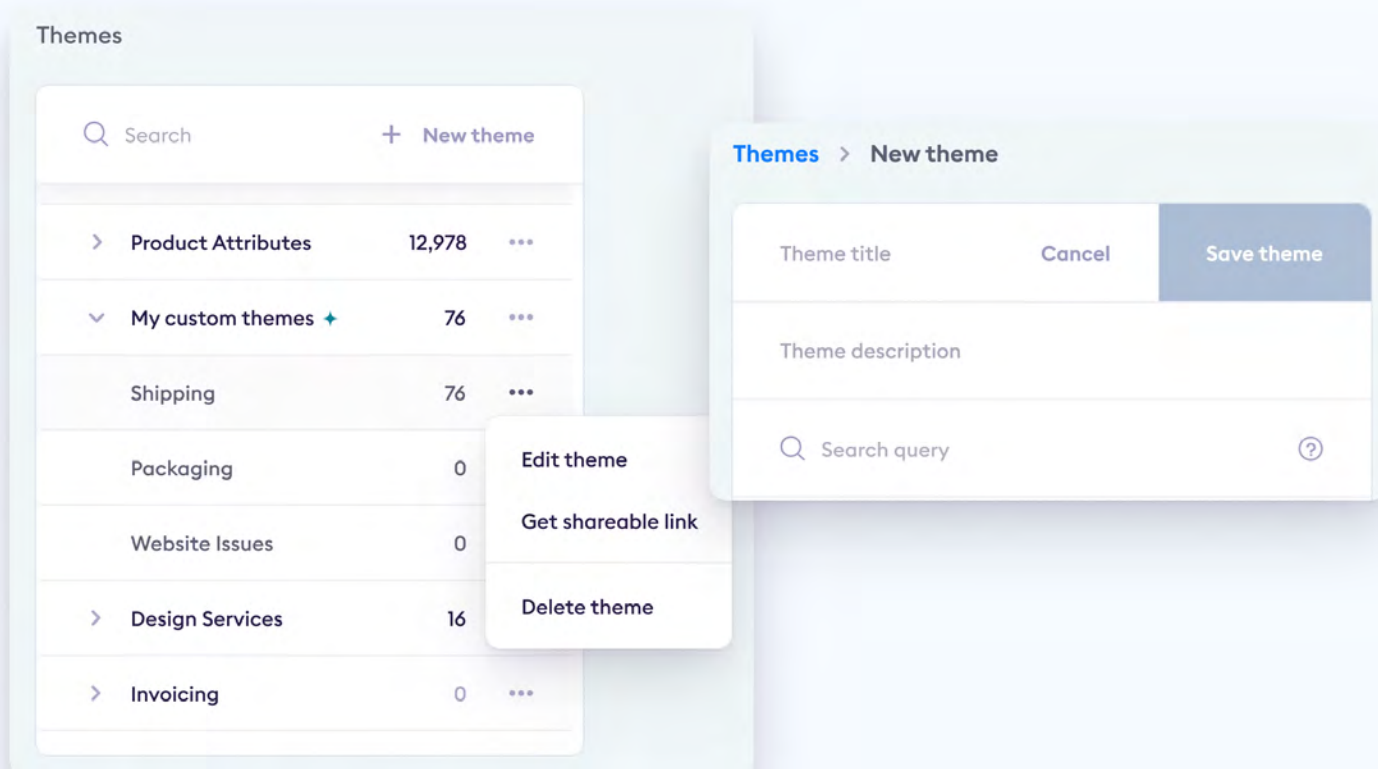
Chattermill's Lyra AI combines state-of-the-art GPT models with our proprietary technology to generate concise summaries of large datasets. This powerful combination ensures that the extracted information is relevant and described in detail, making it highly actionable.

Furthermore, our approach includes an in-depth phrasal analysis targeting longer and context-rich phrases. This means there is no need to manually read through each comment to gather additional information. Our technology efficiently captures the essence of the data, providing comprehensive insights without the need for extensive manual review.

## CUSTOMIZATION

In the context of sentiment analysis, customization means adapting and fine-tuning the approach to fit specific business needs and domain characteristics. This involves defining and refining the tracked themes to accurately capture sentiment related to particular subjects, entities, or language patterns.

Customization ensures that the sentiment analysis is better suited to the specific requirements, making it more precise and relevant for the given context.



Chattermill users have the flexibility to add custom themes to monitor emerging trends and topics relevant to their business.



## The Science Behind Customer Experience Intelligence

### Rule-Based Lexicon

Rule-based approaches offer a high degree of customization. Domain experts or linguists can manually create and refine the rules based on specific requirements and domain knowledge, but these mainly revolve around the researcher's subjective decisions during the analytical process.

### Thematic Analysis & Neural Embeddings

Customizing thematic analysis requires making subjective decisions about the themes and categories to be analyzed. Researchers need to define and establish the criteria for theme identification, which can vary based on personal interpretation and expertise.

### Pure Generative AI (i.e. GPT-4)

While prompt engineering allows some control over the generated text, there can still be limitations in guiding the model's output precisely. The model might exhibit unpredictable behaviour, producing outputs that may not align with the desired customizations or contexts.

### Chattermill's Lyra AI

Customization in Aspect-Based Sentiment (ABSA) primarily involves training the model to improve its performance and adapt it to a specific context. In addition, companies have the flexibility to manually add and manage custom themes or topics of interest to further monitor emerging trends and topics relevant to their business. Adding custom themes in Chattermill does not require additional model training and can be performed by account users at any time. This allows businesses to easily incorporate new themes or topics of interest without the need for extensive technical expertise or retraining the underlying Lyra AI model.

## About the Authors



**Dr Aji Ghose**

**Dr Aji Ghose** serves as the Vice President of Data & Research at Chattermill, overseeing Data Science, MLOps, and MLEng, focused on unearthing actionable insights from customer feedback. Previously, he was the Head of Research & Analytics at Sky, managing a large group of data scientists, researchers, and analysts, working on product and marketing optimization, content recommendation, and machine learning research. Aji earned his PhD in Computational Cognitive Science, specializing in multimodal deep learning, from Birkbeck, University of London. He holds an MSc in Computer Science with Artificial Intelligence from the University of York. Aji also serves as a chair and lecturer for Data Science, Statistics, and AI courses at the Market Research Society.



**Dmitry Isupov**

**Dmitry** has a background in customer insights and data analysis. Prior to founding Chattermill he worked in Market Research building technology to extract insights from customer feedback for some of the worlds biggest brands including Jaguar Land Rover and BT. In 2015 he co-founded Chattermill to bring a new way of analyzing customer feedback through the use of neural networks. Today Dmitry heads up the Insights & Strategy Team at Chattermill and is passionate about helping brands adopt the Customer Experience-Led Growth methodology. Dmitry holds a bachelors degree in Mathematics from the University of Bristol.

## About Chattermill

The Chattermill Customer Experience Intelligence Platform helps businesses unlock their customer reality. Using Chattermill, companies can unify their customer feedback data across reviews, support tickets, conversations, and social media to uncover what customers want, need, and expect from their products and services.

Chattermill unifies customer feedback, customer support, and product feedback into a single platform and uses deep learning Artificial Intelligence (AI) to analyze customer data at scale and provide actionable insights.

The world's leading brands like Amazon, Uber, Virgin Mobile, Zendesk, Tesco, JustEat, and H&M use Chattermill to transform their customer experiences and drive growth for their businesses. Chattermill is a private company headquartered in London, UK.



### Follow us:



[chattermill.com](https://chattermill.com)



[blog.chattermill.com](https://blog.chattermill.com)



[/chattermill](https://www.linkedin.com/company/chattermill)



[hello@chattermill.com](mailto:hello@chattermill.com)