MASTERS IN MANAGEMENT (MIM)

MASTERS FINAL WORK

DISSERTATION

FROM PROMISE TO PERCEPTION: EXPLORING EMPLOYER BRANDING SIGNALS THROUGH SENTIMENT ANALYSIS

SOFIA RAMALHO BASTOS SERRA

FEBRUARY - 2026

MASTERS IN MANAGEMENT (MIM)

MASTERS FINAL WORK

DISSERTATION

FROM PROMISE TO PERCEPTION: EXPLORING EMPLOYER BRANDING SIGNALS THROUGH SENTIMENT ANALYSIS

SOFIA RAMALHO BASTOS SERRA

JURY:

PRESIDENT: PROF. JOSÉ MANUEL CRISTÓVÃO VERÍSSIMO

RAPPORTEUR: PROF. MARIA EDUARDA MARIANO AGOSTINHO SOARES

SUPERVISION: PROF. SUMITA AMITAVA DATTA

FEBRUARY - 2026

ACKNOWLEDGMENTS

First and foremost, I would like to express my genuine appreciation to my supervisor, Professor Sumita Datta, for all her guidance and continuous support over these past few months. Her feedback and confidence placed in me were crucial to the development of this dissertation.

I am also truly grateful to my family for their unconditional support. This milestone would not have been possible without their constant encouragement and unwavering belief in my abilities, even in the most demanding moments.

My appreciation extends to my friends and colleagues for their kind words and motivation. Their presence made the challenges lighter and the achievements more meaningful.

I would also like to thank my boyfriend, Diogo, for his support. His understanding, reassurance and constant encouragement were particularly important during the most difficult phases of writing this dissertation.

Finally, a very special acknowledgment goes to Inês Vicente, a friend who has accompanied me since the beginning of my undergraduate years at ISEG. She was, without doubt, an essential part of this achievement, and I remain deeply grateful for the support and valuable insights she provided along the way.

ABSTRACT

This dissertation examines employee experience of employer branding (EB) by employing sentiment analysis of online employee reviews. More specifically, the study analyses whether EB dimensions are reflected in online reviews, how these dimensions differ in their associations with employee sentiment, and whether formal signals of EB, such as the Great Place to Work (GPTW) ranking, align with employee sentiment. Furthermore, it examines both formal and informal EB signals in explaining employee sentiment.

The research adopts a socio-computational method involving both qualitative and quantitative techniques. A total of 779 employee reviews from Glassdoor and Indeed, covering 10 companies in Portugal, were analysed using text-mining techniques, including sentiment analysis and thematic classification. Regression analysis was used to examine the relationship between EB dimensions and employee sentiment, while the association with GPTW ranking was analysed at a descriptive level.

The results indicate that (1) established EB dimensions are clearly reflected in online employee reviews; (2) these dimensions differ in their association with employee sentiment, with healthy work atmosphere emerging as the strongest driver of positive sentiment; and (3) GPTW ranking shows a positive, although descriptive, association with employee sentiment.

Academically speaking, this study contributes to EB and signaling literature by integrating formal certifications and online employee reviews within a unified empirical framework. From a business perspective, the results highlight the importance of monitoring both informal and formal signals to better understand and manage employer reputation in digital platforms.

Key-words: Employer Branding, Employer Reputation, Online Employee Reviews, Signaling Theory, Text Mining, Great Place to Work (GPTW).

RESUMO

Esta dissertação examina a experiência dos colaboradores relativamente a e*mployer branding* (EB), através da análise de sentimentos em avaliações *online* de colaboradores. Mais especificamente, o estudo analisa se as dimensões da EB são refletidas nas avaliações *online*, como estas dimensões diferem nas suas associações com os sentimentos dos colaboradores e se os sinais formais de EB, como a classificação Great Place to Work (GPTW), estão alinhados com os sentimentos dos colaboradores. Além disso, analise o papel dos sinais formais e informais de EB na explicação dos sentimentos dos colaboradores.

A investigação adota um método sóciocomputacional, envolvendo técnicas tanto qualitativas quanto quantitativas. Foi analisado um total de 779 avaliações de colaboradores da Glassdoor e Indeed, abrangendo 10 empresas em Portugal, através de técnicas de mineração de texto, incluindo análise de sentimento e classificação temática. Foi utlizado um modelo de regressão para examinar as relações entre as dimensões da EB e o sentimento dos colaboradores, enquanto a associação com a classificação GPTW foi analisada de forma descritiva.

Os resultados indicam que (1) as dimensões estabelecidas da EB estão claramente refletidas nas avaliações *online* dos colaboradores; (2) estas dimensões diferem na sua associação com os sentimentos dos colaboradores, surgindo um ambiente de trabalho saudável como o fator mais forte para o sentimento positivo; e (3) a classificação GPTW apresenta uma associação positiva, embora de natureza descritiva, com o sentimento dos colaboradores.

Do ponto de vista académico, este estudo contribui para a literatura sobre EB e a sinalização, integrando as certificações formais e as avaliações *online* dos colaboradores numa estrutura empírica unificada. Numa perspetiva empresarial, os resultados destacam a importância de monitorizar os sinais formais e informais para melhor compreender e gerir a reputação do empregador nas plataformas digitais.

Palavras-chave: *Employer Branding*, Reputação do Empregador, Reviews *Online* de Colaboradores, Teoria da Sinalização, Mineração de Texto, Great Place to Work (GPTW).

ABBREVIATIONS

BART – Bidirectional and Auto-Regressive Transformers.

BERT – Bidirectional Encoder Representations from Transformers.

CSR – Corporate Social Responsibility.

EB – Employer Branding.

EVP – Employer Value Proposition.

eWOM – Electronic Word-of-Mouth.

GPTW – Great Place to Work® ranking.

H – Hypothesis.

NLP – Natural Language Processing.

RoBERTa – Robustly Optimized BERT Approach

RQ – Research Question.

VIF – Variance Inflation Factor.

WOM – Word-of-Mouth.

ZSL – Zero-Shot Learning.

INDEX

ACKNOWLEDGMENTS .... **III**

ABSTRACT .... **IV**

RESUMO .... **V**

ABBREVIATIONS .... **VI**

CHAPTER 1 – INTRODUCTION **1**

1.1. Theoretical context 1

1.2. Study objectives 2

1.3. Relevance of the study 2

1.4. Dissertation structure 3

CHAPTER 2 - LITERATURE REVIEW **4**

2.1. Employer branding and employer reputation 4

2.2. Information asymmetry and signaling theory in the labour market 5

2.3. Rankings and certifications as formal signals of quality of employer 7

2.3.1. GPTW as a formal signal 7

2.4. Digital media, eWOM and perceptions of employees as informal signals 8

2.4.1. From WOM to eWOM 8

2.4.2. Online employee reviews as an expression of employee voice 9

2.4.3. Dimensions of experience of the employee in online employee reviews ... 10

2.5. Consistency and inconsistency between signals: mismatch 11

CHAPTER 3 - CONCEPTUAL MODEL AND HYPOTHESIS DEVELOPMENT **13**

CHAPTER 4 – METHODOLOGY **14**

4.1. Type of study 14

4.2. Population and sampling 14

4.3. Data collection 14

4.4. Data preprocessing 15

4.5. Text mining measures 16

4.5.1. Sentiment analysis (BERT) 16

4.5.2. Thematic classification (Zero-shot BART) 17

4.6. Exploratory regression analysis 18

CHAPTER 5 - DATA ANALYSIS AND RESULTS **19**

5.1. Sample characterisation 19

5.2. Descriptive and exploratory analysis of employee reviews 19

5.2.1. Overall employee sentiment 19
5.2.2. Employer branding dimensions reflected in online employee reviews 20
5.2.3. Employer branding dimensions and employee sentiment 21
5.2.4. Employee sentiment and formal employer branding signals 22
5.3. Exploratory regression analysis 23
5.3.1. Employer branding dimensions and employee sentiment 23
5.4. Summary of hypothesis testing results 24
CHAPTER 6 – DISCUSSION AND CONCLUSION 25
6.1. Discussion of research questions 25
6.1.1. Discussion of RQ1 and H1 25
6.1.2. Discussion of RQ2 and H2 26
6.1.3. Discussion of RQ3 and H3 27
6.2. Academic contributions 27
6.3. Contribution to management 28
6.4. Study limitations and future research 29
REFERENCES 30
DECLARATION 38
APPENDICES 39
Appendix A - Web scraping setup and configuration – Python code 39
Appendix B - Translation and concatenation (Data preprocessing) 41
Appendix C - Sentiment analysis (Transformer-based RoBERTa model) 41
Appendix D - Thematic analysis (Zero-shot BART) 42
Appendix E - Company-level statistics 43
Appendix F - Statistical analysis for the regression analysis 44

TABLE INDEX

Table I - Top 10 GPTW® Portuguese companies (2025) 14
Table II - Distribution of employee reviews by company and platform 19
Table III - Descriptive statistics of employee sentiment 20
Table IV- Employer Branding dimensions reflected in online employee reviews......... 21
Table V – Average employee sentiment by employer branding dimension 22
Table VI - Average employee sentiment and GPTW ranking by company 23
Table VII - OLS Regression Results: Impact of Employer Branding Dimensions on Employee Sentiment.. 24
Table X - Summary of hypothesis testing results............**Error! Bookmark not defined.**

FIGURES INDEX

Figure 1 - Conceptual Model.. 13
Figure 2 – Histogram of employee sentiment 20

CHAPTER 1 – INTRODUCTION

1.1. Theoretical context

The concept of employer branding (EB) was first introduced by Ambler and Barrow (1996), who defined it as the package of functional, economic and psychological benefits provided by employment. Since its first conceptualisation, research has shown that EB plays an important role in shaping how organisations are perceived by both potential and current employees, influencing their attractiveness and reputation (Backhaus & Tikoo, 2004; Theurer et al., 2018).

EB is especially relevant in contexts with information asymmetry. In labour markets, individuals rarely have complete knowledge of what it is like to work for a company before joining it and must rely on available signals to form expectations. Signalling theory, introduced in economics by Michael Spence in 1973, helps explain this process by suggesting that organisations communicate information about employer quality through observable signals (Spence, 1973; Connelly et al., 2011).

These signals can vary in credibility, visibility and costliness (Carson & Westerman, 2023) and can take different forms. On one hand, organisations use formal EB signals, such as rankings, awards and certifications, which are controlled by external institutions or the organisations themselves. On the other hand, informal signals such as employee narratives, word-of-mouth (WOM) and online employee reviews, are more subjective and emerge from lived experiences (Willems et al., 2017). These informal signals provide valuable information about organisations and strongly influence perceptions of employer quality (Symitsi et al., 2021). The reviews also act as a form of digital employee voice (Holland et al., 2016), due to their voluntary and relatively anonymous nature (Symitsi et al., 2021). As a result, EB is no longer only shaped by formal signals, such as corporate and official communication, awards or other certifications, but also by employee-generated evaluations shared online.

This form of electronic word-of-mouth (eWOM), despite their unstructured textual nature, consistently reflects recurring EB dimensions (Berthon et al., 2005; Tanwar & Prasad, 2017). By examining these dimensions, it becomes possible to identify which aspects of the employment experience employees discuss most and where possible areas for improvement may exist. Understanding how these signals are related to employee sentiment globally and across dimensions is essential for a complete view of EB.

1.2. Study objectives

The main objective of this dissertation is to understand how EB is expressed in online employee reviews and how it is related to employee sentiment. Particularly, it intends to assess whether the image companies formally communicate as employers is consistent with the perceptions expressed by employees on digital platforms.

Firstly, the study aims to identify which EB dimensions are actually present in online employee reviews. Secondly, it analyses how these dimensions are associated with employee sentiment, examining whether some are more strongly related than others. Finally, it investigates whether formal EB signals, such as the GPTW ranking, are associated with employee sentiment expressed in online reviews.

To be specific, this dissertation addresses the following research questions:

- **RQ1:** Which EB dimensions are reflected in online employee reviews?
- **RQ2:** How are EB dimensions reflected in online employee reviews associated with employee sentiment, and do these associations differ across dimensions?
- **RQ3:** Is employee sentiment associated with formal EB signals, such as the GPTW ranking?

1.3. Relevance of the study

EB has become an important function of organisations. Companies invest heavily in EB strategies (Wilden et al., 2010), such as employer rankings, to present themselves as attractive places to work. However, these formal signals do not always reflect how employees experience their jobs (Dabirian et al., 2017).

At the same time, online employee reviews have emerged as a widely used source of information about the working conditions of a company. Prior research highlights that such reviews provide good insights about employee perceptions (Symitsi et al., 2021); however, despite their growing relevance, employee reviews are often studied separately from traditional EB frameworks.

This study is relevant because it brings these perspectives together. By examining EB dimensions in online employee reviews and relating them to employee sentiment, the study helps clarify how different aspects of work are discussed by employees (Tanwar & Prasad, 2017). Besides that, comparing informal signals derived from employees with formal EB signals allows a more complete understanding of how EB is formed and interpreted (Dineen & Allen, 2016). Furthermore, this study contributes to Sustainable

Development Goal 8 (Decent Work and Economic Growth), as it promotes greater transparency regarding working conditions, emphasising the role of employee voice.

1.4. Dissertation structure

This dissertation is organised into six chapters. Chapter 1 introduces the study, presenting its theoretical context, objectives and relevance. Chapter 2 is the review of relevant literature related to EB, signaling theory and online employee reviews. Chapter 3 presents the conceptual model and develops the research hypotheses. Chapter 4 addresses the methodology, including the type of study, data collection process, sample and population characteristics, text mining measures and analytical procedures adopted. Chapter 5 presents the study's findings, including descriptive analysis and regression results. Finally, Chapter 6 discusses the main results regarding research questions and hypotheses. This chapter also includes the study's academic and managerial contributions, highlights its limitations and suggests directions for future research.

CHAPTER 2 - LITERATURE REVIEW

2.1. Employer branding and employer reputation

In the last few decades, the labour market has become increasingly competitive, intensifying the so-called “war for talent” and leading organisations to find ways to differentiate themselves as employers. The difficulty of retaining and attracting talent reinforced the need for a strong brand to build and demonstrate trust with internal and external audiences (Reis et al., 2021). In this context, EB emerges as a strategic response to talent shortages, helping organisations position themselves as “employer of choice” (Berthon et al., 2005). Some of them even created specific functions in Human Resources departments to focus on the employer brand (Joglekar & Tan, 2022).

The phenomenon of EB has evolved significantly since its initial formulation by Ambler and Barrow in 1996. The authors defined the employer brand as “the package of functional, economic, and psychological benefits provided by employment, and identified with the employing company” (Ambler & Barrow, 1996, p. 187). According to the authors, functional benefits are useful activities and development opportunities, economic benefits are material or monetary rewards, and psychological benefits are the feeling of belonging, purpose and direction. Overall, this view allows the employer brand to be understood not just as communication, but as a value promise that shapes the expectations of both current and future employees (Alves et al., 2020).

This conceptualization was later consolidated by Backhaus and Tikoo (2004), who proposed the first theoretical model linking the internal and external dimensions of the employer brand. These authors showcased the importance of organizational culture and the Employer Value Proposition (EVP), defined as the equalization of the rewards and benefits employees receive in return for their performance at work (Pawar, 2016), combining functional, economic and psychological benefits that are communicated through organizational messages, symbols and everyday practices (Berthon et al., 2005). While Backhaus and Tikoo (2004) focused on structuring the EB process, Lievens and Highhouse (2003) and later Lievens and Slaughter (2016), expanded the concept’s comprehension, distinguishing instrumental (tangible job characteristics) and symbolical attributes (intangible meanings, such as values, identity and organizational prestige). These attributes define the EVP and the employer image organisations seek to communicate as employers, but they are not always interpreted as intended. As a result, the image organisations aim to project may differ from how they are perceived (Theurer et al., 2022).

In contrast to EB, corporate reputation is referred to as a global, temporally stable, evaluative judgement about an organisation that is shared by the general public (Highhouse et al., 2009). These evaluations correspond to the image, perceived and evaluated by stakeholders, resulting from a social construction over time (Lievens & Slaughter, 2016). While EB represents an attempt to shape perceptions, employer reputation reflects how these messages are assimilated. This is crucial, since employer reputation is not directly controlled by the organisation and it can diverge from its intended employer image (Junça-Silva & Dias, 2022).

This distinction provides a useful framework for differentiating between intended and perceived employer image. On one hand, EB, a top-down process, enables organisations to communicate their desired employer image, that is, the way they want to be perceived by current and potential employees (Dell et al., 2001). On the other hand, employer reputation develops bottom-up, as it accumulates stakeholder experiences and external signals. Cable and Turban's study (2003) show that job seekers apply to companies based on employer reputation, which develops over time through external signals rather than isolated recruitment messages. As a result, EB operates indirectly by shaping employer reputation.

This way, EB can be understood as a controlled organisational signal that firms use to communicate who they are as employers (Wilden et al., 2010). When there is information asymmetry (Connelly et al., 2011) and job seekers want to know what it is like to work at an organization (Turban, 2001), candidates resort to observable signals, such as career websites and social media, EVP statements, awards or certifications to infer on less visible aspects of the employment experience (Carpentier et al., 2019). These signals, combined with direct experience and third-party information, contribute to the creation of employer reputation (Stockman et al., 2020). When organisational messages align with lived experiences, EB reinforces employer reputation. When there are inconsistencies, reputational damage may occur (Junça-Silva & Dias, 2022).

2.2. Information asymmetry and signaling theory in the labour market

Labour markets are characterised by information asymmetry, whenever employers and candidates make decisions without being able to directly observe all the qualities of the other party (Connelly et al., 2011). For instance, candidates fully understand employer quality only after joining the company, and during recruitment, employers can only rely on observable signals to assess candidates' productivity or future behavior.

In this context, the Signaling Theory emerges. This theory introduced in economics in 1973 by Michael Spence who analysed the labour market as a context in which employers and candidates have unequal information and resort to observable signals to infer on characteristics that are not observable by others (Spence, 1973; Drover et al., 2018). The theory was later extended to management, marketing, and entrepreneurship to explain how organizations and individuals show their quality when there is uncertainty (Choudhury, 2024).

Within the signaling theory framework, organisational signals can be defined as observable actions, attributes, or forms of communications that, either intentionally or unintentionally, shape external audiences' beliefs about less visible organisational characteristics, such as reputation, reliability, or the quality of the work atmosphere (Spence, 1973; Connelly et al., 2011).

Research distinguishes formal and informal organisational signals. Formal signals are structured mechanisms such as rankings, certifications, and official reports, often based on third-party validation. In contrast, informal signals, such as employee narratives, word-of-mouth (WOM), and online reviews, are less clearly defined, more subjective, and emerge from lived experiences and judgements of stakeholders (Willems et al., 2017). These informal signals may either support or contradict the formal messages, as audiences consider multiple signals at the same time (Drover et al., 2018).

The effectiveness of a signal depends on several factors according to Signaling Theory. Connelly et al. (2011, 2025) argue that organisational signals depend on their credibility, costliness, and visibility. Credible and costly signals are harder to imitate, discouraging low-quality organizations from sending misleading messages. At the same time, signals can only influence perceptions if they are visible to receivers and understood as credible representations of quality (Carson & Westerman, 2023). As a result, signals that combine credibility, costliness, and visibility tend to be interpreted as more reliable.

In practice, candidates do not interpret a single isolated signal when forming perceptions of an employer. Instead, they interpret a portfolio of signals, including job advertisements, online reviews, rankings, certifications, sustainability reports, and career websites. Recent research refers to this multitude of signals as "signal sets" (Drover et al., 2018). When they are aligned, they tend to reinforce positive perceptions. When signals are contradictory, they may create skepticism, even if some individual signals are positive.

2.3. Rankings and certifications as formal signals of quality of employer

From an institutional perspective, rankings and certifications serve as forms of third-party employment branding, outsourcing the evaluation of employer attractiveness to independent, credible organisations (Love & Singh, 2011). In signaling theory, these certifications reduce information asymmetries between employers and candidates, reducing information to a ranking position (Saini & Jawahar, 2019). Green certifications, for example, have been gaining relevance as they operate as ethical signals, increasing prestige and person-organization fit (Guillot-Soulez et al., 2021).

From an EB perspective, research shows that the use of rankings and certifications is strongly correlated to employer attractiveness and reputation. Studies on the topic identify that, to be "employers of choice", organizations use external distinctions to reinforce their image and differentiate themselves from competitors (Dineen & Allen, 2016), contributing to visibility among candidates and the media. Evidence also indicates that "best employer" rankings are associated with positive outcomes, such as lower turnover intentions and more applicants (Zhang et al., 2022).

There are many advantages to employer certification; however, in the scope of this study, the main advantage is the comparability across organisations, as rankings allow candidates to quickly assess the attractiveness of different employers, reducing uncertainty and effort (Saini & Jawahar, 2019). Plus, the fact that these rankings are public amplifies the notoriety of well-ranked companies, which is associated with greater perceived desirability and better reputation (Vidaver-Cohen, 2007).

Despite this, several authors identify important limitations, noting that the nature of employer rankings and certifications is to simplify complex organisational realities within the same organisation (Dineen & Allen, 2016). Specifically, research shows that formal signals, such as rankings, do not always align with assessments produced on digital platforms. This highlights the risk of oversimplification and potential mismatches with employees' lived experiences.

2.3.1. GPTW as a formal signal

The GPTW certification is part of a group of certifications known as "Best Place to Work Certification," and it is frequently associated with a form of EB promoted by an external entity. The GPTW organisation gained relevance from conducting studies on trust and quality in work relationships and has since expanded to multiple countries and sectors, awarding seals and producing national rankings and sector-based "Best Workplaces" lists (Uribe & Ríos, 2024).

The GPTW methodology is a process in which organisations voluntarily apply, combining anonymous employee surveys with an analysis of people management practices, such as development, benefits, recognition and well-being. The survey, commonly known as the Trust Index, gathers perceptions of the work environment across five central organisational dimensions: credibility, respect, fairness/impartiality, pride, and camaraderie/partnership (Love & Singh, 2011), which are aggregated into an overall score.

From a signaling theory perspective, certification is a costly, observable signal because it requires paying fees, allocating organisational resources, and maintaining sustained employee trust, making it difficult for low-quality employers to imitate. Consequently, GPTW certification can help improve employee trust, increase retention rates, and enhance employer brand awareness (Great Place to Work, n.d.). Besides that, prior research shows that organizations with the GPTW seal "have been shown to positively affect employee attitudes, applicant pool quality and financial performance" (Dineen & Allen, 2016, p. 3). In practice, the GPTW certification is often used as an external EB signal, positioning the organisation as an "employer of choice" (Mohiya, 2024). However, despite its relevant role, a GPTW-certified organisation may experience negative outcomes in certain departments or countries (Dabirian et al., 2017).

2.4. Digital media, eWOM and perceptions of employees as informal signals

2.4.1. From WOM to eWOM

Van Hoye and Lievens (2005) showed that, when assessing an organisation's attractiveness, job seekers rely not only on corporate communication but also on external and interpersonal sources. This phenomenon is called word-of-mouth (WOM). It is typically defined as informal interpersonal communication about products, services, or organisations between people with social ties (family, friends, colleagues) and independent of the company (Stockman et at., 2020). Earlier WOM research, conducted before the widespread use of online review platforms, indicates the impact of WOM depends on the credibility and proximity of the source, with recommendations from closer or more trusted sources having stronger effects (Van Hoye & Lievens, 2007).

Moreover, with the emergence of Web 2.0 and new media channels, WOM has evolved into electronic word-of-mouth (eWOM), defined as any positive or negative statement about a company made available online through digital platforms (Verma & Yadav, 2022). Specifically, eWOM is distinguished by being decentralized (any user can produce content), persistent (messages remain accessible over time), and highly accessible (it can be consulted by a wide audience) (Donthu et al., 2021). In the labour market, eWOM about employers includes reviews on specialised websites, comments on social networks, and discussions in forums (Symitsi et al., 2021).

In the digital context and from a signaling perspective, the credibility of eWOM is often high because messages are perceived as coming from “insiders”, that is, current or former employees, and, therefore, are seen as more authentic than corporate communications (Gimpl, 2025). Despite this, there are risks of bias, such as reviews that may reflect extreme experiences (very positive or very negative), specific motivations, or subjective perceptions (Sainju et al., 2021).

Research further indicates that negative eWOM tends to attract more attention and has a stronger impact on employer attractiveness than positive reviews, confirming a slight negativity bias typical of digital communication (Yu et al., 2022). However, the effect is moderate through employer brand equity, meaning strong brands are less affected by isolated negative comments, whereas less well-known companies are more vulnerable, as online criticism may become the main signal of quality available, with greater persuasive power (Stockman et al., 2020). Overall, the growing relevance of informal signals, particularly eWOM, makes employer reputation more complex to manage, even if it simplifies information search and comparison for job seekers and other stakeholders (Gimpl, 2025).

2.4.2. Online employee reviews as an expression of employee voice

Online employee reviews are comments and evaluations published by both current and former employees on specialized platforms such as Glassdoor or Indeed, including quantitative ratings and text about pros, cons, and recommendations (Sainju et al., 2021). For example, Glassdoor, a platform with tens of millions of users globally and over 2.5 million company profiles, combines different work experience factors, such as satisfaction, culture, and relationships, that are difficult to see in official communications (Glassdoor, 2023).

The reviews are characterised by relative anonymity and voluntariness (Symitsi et al., 2021), allowing them to function as a form of digital employee voice (Holland et al., 2016). The experiences reported in reviews have a significant impact on external perceptions, making employees act as brand ambassadors (Kanwal & Van Hoye, 2024). In some ways, review platforms shift control of communication from organizations to employees. Unlike institutional communication, reviews are a form of internal eWOM that job seekers actively see and consult (Symitsi et al., 2021).

Research treats online employee reviews as a proxy for perceptions. Studies that include many reviews show that employees discuss topics such as management, compensation, career opportunities, work-life balance, and culture, revealing drivers of satisfaction and turnover. These reviews, published on platforms, such as Glassdoor and Indeed, are generally valid and correlate with objective indicators, such as financial performance, turnover, and organisational climate metrics (Morgan & Chapman, 2024).

These characteristics legitimize treating reviews as research data and justify their use in studies seeking to capture informal signals about employer quality, allowing comparisons of perceptions based on lived experiences with formal signals.

2.4.3. Dimensions of experience of the employee in online employee reviews

Studies applying text mining techniques to large collections of online reviews show that employees' perceptions tend to cluster around a limited set of recurring themes or dimensions related to everyday work experiences (Guo et al., 2017). On platforms such as Glassdoor, reviews are not fully structured around predefined themes. Instead, employees provide open-text reviews alongside general ratings. Despite this unstructured format, research consistently shows that recurring themes emerge when large volumes of Glassdoor reviews are analysed (Morgan & Chapman, 2024).

Several models have been developed to conceptualise these dimensions. For example, Berthon et al. (2005) proposed the EmpAt scale which operationalises employer attractiveness through five dimensions: interest value (challenging and innovating work), social value (team environment and positive relationships), economic value (competitive salary and benefits), development value (growth and learning opportunities) and application value (possibility to apply and expand skills). This scale has become an international reference, allowing comparisons of perceptions of attractiveness and identifying the most valued factors across generations (Benraïss-Noailles & Viot, 2018).

More recently, Tanwar and Prasad (2017) developed and validated a multidimensional framework based on five core dimensions of the EVP. Although the terminology differs, these models largely overlap and point to similar aspects of the work experience. Nevertheless, the five dimensions are: healthy work atmosphere, training and development, work-life balance, ethics and corporate social responsibility (CSR) and compensation and benefits. It is worth noting that, while the EmpAt scale (Berthon et al., 2005) does not consider the sustainability of the employer brand, Tanwar and Prasad address this with a CSR dimension. These dimensions reflect the conditions under which work is experienced, rather than outcomes such as satisfaction or turnover. Such outcomes are understood as responses to these underlying conditions.

By dividing online employee reviews into distinct dimensions, it becomes possible to understand more accurately how employees assess various traits of their work experience (Guo et al., 2017). These evaluations collectively originate employee sentiment, reflecting positive or negative tone with which employees feel about their employer and their employment experience. Research using text mining measures on platforms, such as Glassdoor, shows that textual sentiment and attribute ratings are strongly related to overall job satisfaction and perceptions of the organization (Jung & Suh, 2019).

2.5. Consistency and inconsistency between signals: mismatch

In today's digital labour market, information about employers reaches candidates and job seekers through different channels. Rankings, awards, and certifications provide formal signals about employer quality, whereas online reviews and eWOM offer informal signals regarding employees' everyday experiences (Connelly et al., 2011; Dineen & Allen, 2016). Consistency between these signals refers to the extent to which the positive image promoted through distinctions such as "best employer" matches what employees are saying about working conditions, culture and leadership. When there is an alignment across signals, uncertainty decreases and the perceptions of the employer tend to be more credible. However, if there is no alignment, the mixed signals can weaken employer attractiveness and raise doubts about the organization's true quality (Drover et al., 2018).

Within this framework, a mismatch can be defined as a significant misalignment between formal and informal signals about employer quality. Research on "mixed signals" shows that when candidates are confronted with inconsistent information from institutional websites and online employer reviews, perceived attractiveness decreases compared to scenarios with aligned and congruent signals (Pernkopf et al., 2020).

There are many reasons that can explain this misalignment. First, difference in criteria: rankings and certifications tend to be based on macro indicators such as formal policies, benefits or CSR programs (Wang et al., 2022), which may not capture the full experience reflected in the reviews. Second, timing issues can create discrepancies: recent organisational changes, including restructurings, transition of leadership or new policies may take time to be incorporated into rankings, while employee reviews react more quickly. Third, communication strategies: organisations may highlight only positive aspects in official communications and on social media, while problems such as negative climates tend to appear more on employees' accounts (Dabirian et al., 2017).

From the perspective of information recipients, the alignment between formal and informal signals is a critical indicator of employer quality. When official communication and employee reviews align, the employer brand is more likely to be perceived as transparent and reliable (Zhang et al., 2022). In contrast, a marked mismatch may raise doubts about honesty and openness to employee voice (Pernkopf et al., 2020).

CHAPTER 3 - CONCEPTUAL MODEL AND HYPOTHESIS DEVELOPMENT

Drawing from signaling theory, the model assumes that employee perceptions are shaped by observable signals that organisations use to communicate unobservable characteristics under information asymmetry (Connelly et al., 2011). In this context, online employee reviews function as informal EB signals. The model therefore assumes that EB dimensions commonly discussed in the literature, specifically those proposed by Tanwar and Prasad (2017), can be empirically identified in online employee reviews.

H1: Online employee reviews reflect EB dimensions proposed in the literature.

Once they are identified, these dimensions are expected to be related to employee sentiment, which is derived from the same employee-generated textual content. Different dimensions capture distinct aspects of employment and are, therefore, likely to be associated with employee sentiment to varying degrees (Dineen & Allen, 2016).

H2: EB dimensions reflected in online employee reviews are associated with employee sentiment, differing in strength across dimensions.

In addition to informal employee-generated signals, the model incorporates a formal EB signal in the form of the GPTW certification, functioning as an externally visible indicator intended to communicate employer quality (Dineen & Allen, 2016). As a result, the model assumes that formal EB signals are positively associated with employee sentiment.

H3: Formal EB signals are positively associated with employee sentiment.

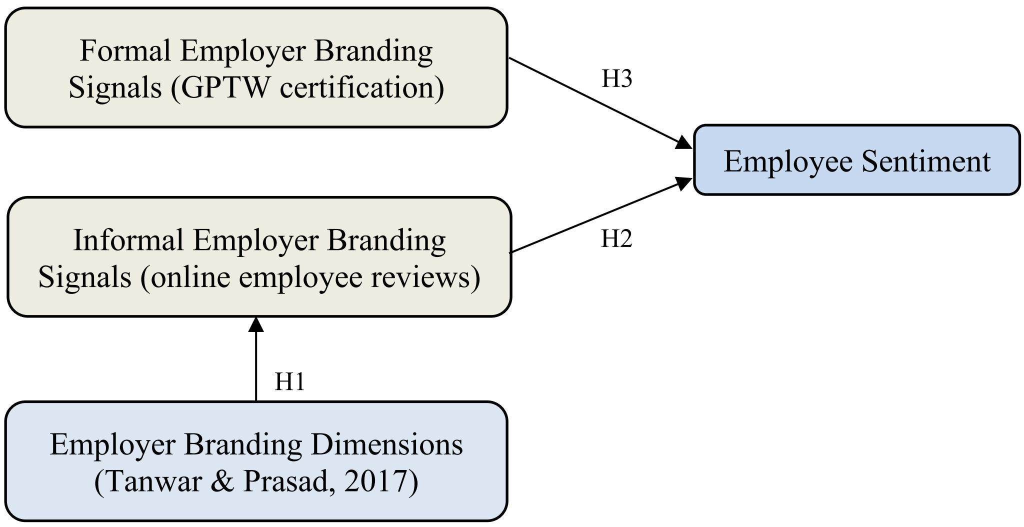

FIGURE 1 - CONCEPTUAL MODEL.

CHAPTER 4 – METHODOLOGY

4.1. Type of study

The study adopts a socio-computational research design, combining qualitative and quantitative techniques, applying computational text analysis to unstructured text. Although the data consists of qualitative employee narratives, these are systematically transformed into quantitative analysis. This approach allows consistent analysis of subjective employee perceptions, making it possible to identify patterns across a large number of online reviews (Gimpl, 2025).

4.2. Population and sampling

The population of this study consists of online reviews published by current or former employees on digital job platforms (e.g., Glassdoor, Indeed). Prior research has demonstrated that such employee reviews constitute a valid and informative source of information about the employee experience, as they reflect real perceptions of satisfaction, work conditions, and organisational practices (Symitsi et al., 2021).

As for the sample, the study comprises 779 reviews collected from two platforms: Glassdoor and Indeed. The sample focuses on the ten highest-ranked companies located in Portugal, according to the GPTW ranking for 2025, restricted to organisations with more than 500 employees (Table I). Similar sampling strategies targeting top-ranked employers are common in the literature to capture perceptions of notorious organisations (Sainju et al., 2021). Glassdoor and Indeed were chosen for their widespread recognition and international use, and because they collect a significant volume of comments from former and current employees.

TABLE I - TOP 10 GPTW® PORTUGUESE COMPANIES (2025)

1	Noesis	**6**	DHL Supply Chain
2	Allianz	**7**	Credibom Bank
3	Cisco	**8**	Solvay
4	ERA Imobiliária	**9**	EY
5	Vestas	**10**	Synopsys Inc.

Source: Great Place to Work Portugal® (2025).

4.3. Data collection

The use of automatically collected data from digital employment platforms has increased in recent research to analyse employee experiences. As Swain et al. (2020)

states, “anonymized platforms such as Glassdoor provide “safe spaces” for employees to share and assess their workplace experiences” (p. 3), proving the importance of taking online data collection into account. From this perspective, the data collection process in this study relied on secondary data from online employee reviews on Glassdoor and Indeed, as well as institutional information on the GPTW ranking. All data used in the study focused exclusively on publicly accessible content and was intended for academic purposes only.

The reviews from the Glassdoor platform were collected through a web scraping process, carried out by using Python (a programming tool), which was later introduced into Microsoft Excel (a data analysis tool). This procedure is consistent with other studies that provide code for collecting and structuring Glassdoor data (Zhou et al., 2024). The Python code used for the web scraping process is provided in Appendix A. In contrast, the reviews collected from the Indeed platform were obtained manually and organized directly in Excel files for further analysis. This was decided because there were fewer reviews available on this platform directed to the analysed companies. Therefore, manual collection was more appropriate and time efficient. In total, the dataset combines 779 employee reviews from both platforms.

The collected and analysed reviews correspond to a period between the 1st of January of 2020 and the date of the data collection (December 2025), covering approximately the last five years. The option for defining 5 years as the time interval was intended to capture recent employee perceptions during a period of significant organisational and workplace changes, while also accounting for the growing importance and increased use of online employee review platforms in recent years (Morgan & Chapman, 2024).

For each review, the following elements were considered for collection: title of the review, rating given by the current or former employee, the date of publication and, finally, the textual content of the review, organized into pros and cons. On the Glassdoor platform, it was also possible to collect information about the employee's position and tenure with the organization, variables not available in Indeed reviews. For the purpose of sentiment analysis, only three components were used as input, namely the title, pros and cons. The remaining variables were retained to provide contextual information.

4.4. Data preprocessing

After data collection, the dataset was prepared for the analysis defined in the study. In the first stage, preprocessing was applied to reduce “noise” in the text without compromising semantic information (Chai, 2023). As noted earlier, the initial dataset combined two different collection processes: reviews automatically scraped from Glassdoor using Python and reviews collected manually from Indeed. In both cases, the information was consolidated and organized in ten Microsoft Excel files, one per company, to facilitate further analysis.

An initial preprocessing step was performed in Excel, focusing on simple normalisation operations, specifically converting text to lowercase and removing special characters, such as emojis. These steps aim to avoid inconsistencies that could affect the interpretation of results. Aggressive traditional preprocessing techniques, such as stemming, lemmatisation, or heavy normalisation, were not applied because they may not yield benefits and could even harm model performance (Kurniasih & Manik, 2022).

Subsequently, additional preprocessing steps were implemented in Python. As the sentiment analysis model operates on English-language input, all reviews were translated into English (Devlin et al., 2019). Additionally, the main components of each review (title, pros, and cons) were concatenated into a single text field, separated by a vertical bar (|), to create a unified text. This preprocessing strategy was designed to be compatible with transformer-based language models, such as those used in this study. The Python code used for translation and concatenation is found in Appendix B.

4.5. Text mining measures

Employee-generated data collected from digital job platforms has been gaining relevance in organisational research. In this context, the use of text mining, defined by Kobayashi et al. (2017, p. 734) as "the discovery and extraction of interesting, non-trivial knowledge from free or unstructured text", is becoming important, particularly for large volumes of data, namely the ones collected for the current research. Within this framework, two text-mining techniques were used to convert employee reviews into quantitative measures. First, sentiment analysis was used to quantify the overall sentiment in reviews. Second, thematic classification was used to assess the extent to which reviews address specific dimensions. The development of these measures and the associated models is described in the following sections.

4.5.1. Sentiment analysis (BERT)

Sentiment analysis, or Opinion analysis, is a text-mining technique in natural language processing (NLP) that involves collecting and analyzing people's opinions, thoughts, and impressions about different topics, products, subjects, and services (Wankhade et al., 2022). In this study, sentiment analysis is used to quantify the overall sentiment expressed in employee reviews. And, to do this task, a pre-trained BERT (Bidirectional Encoder Representations from Transformers) model was used, more specifically RoBERTa (Robustly Optimized BERT Approach), a variant of the original BERT architecture.

BERT, originally developed by Google AI Language, is a transformer-based language model designed to "pretrain deep bidirectional representations from unlabeled text jointly conditioning on both left and right context in all layers" (Devlin et al., 2019, p. 4171), meaning BERT can analyse the context of words, making it extremely adequate for sentiment analysis. Building on this architecture, RoBERTa, proposed by Facebook AI, retains the same transformer-based structure as BERT; however, it introduces improvements in the pre-training process, including larger datasets and more processing capacity (Pipalia et al., 2020). As a result, RoBERTa outperforms BERT on a variety of NLP tasks, making it more adequate for this study (Liu et al., 2019).

After defining the sentiment analysis model adopted in this study, the implementation of the model was carried out by using Python (Appendix C). Regarding the information obtained, the model estimates the probability of each employee review expressing negative, neutral or positive sentiment. Based on these probabilities, a continuous sentiment score was calculated for each review. This score, which ranges from -1 to +1, is calculated as the difference between the predicted probability of positive sentiment and the predicted probability of negative sentiment $p_{\{positive\}} - p_{\{negative\}}$. Positive values indicate favourable sentiment, negative values indicate unfavourable sentiment, and finally, values close to zero indicate neutral sentiment. This continuous score is the primary sentiment measure used in empirical analysis, enabling comparisons across reviews, companies and EB and employer reputation signals.

4.5.2. Thematic classification (Zero-shot BART)

In addition to overall sentiment, thematic classification, a text-mining technique, was used to identify the extent to which reviews refer to specific dimensions. This analysis, also conducted in Python (Appendix D), complements the sentiment analysis by identifying what employees discuss, regardless of sentiment.

This thematic analysis, grounded in NLP, used a zero-shot learning (ZSL) approach based on the BART (Bidirectional and Auto-Regressive Transformers) language model (Kyritsis et al., 2023). In ZSL, large pre-trained models such as BART, allow textual data to be classified into predefined dimensions without requiring task-specific labelled training data. Instead, the model relies on its pre-trained linguistic knowledge to assess the semantic similarity between each review's content and a set of predefined thematic descriptions (Yin et al., 2019).

The thematic dimensions were defined as a priori based on the literature, following the framework proposed by Tanwar and Prasad (2017). The dimensions considered are: healthy work atmosphere, compensation and benefits, work-life balance, training and development, ethics, CSR, and management. This framework was chosen because it is well-suited for capturing employee perceptions expressed in online employee reviews, rather than survey-based institutional measures such as the GPTW model.

For each review, the model generated a continuous relevance score for each thematic dimension. These scores range from 0 to 1 and indicate how well a review addresses a specific theme. Higher values indicate greater thematic relevance, whereas lower values suggest the theme is weakly present. These scores, unlike sentiment analysis, measure the presence of specific themes in each review rather than the overall sentiment expressed. These thematic scores, together with the sentiment measure, were used in the subsequent empirical analysis.

4.6. Exploratory regression analysis

To examine the relationships identified in the descriptive analysis, an exploratory regression analysis is conducted to assess whether the association between employee sentiment and the EB dimensions holds when multiple variables are considered simultaneously (Tkalac Verčič, 2021).

The regression analysis focuses on informal EB signals, operationalised through EB dimensions derived from online employee reviews. All analyses are performed at the review-level and estimated using ordinary least squares (Shikhman & Müller, 2021), using the continuous sentiment score derived from employee reviews as the dependent variable. Given the limited number of companies in the sample, the regression analysis is exploratory, so the results should be interpreted as indicative associations rather than causal effects. To ensure robustness, multicollinearity diagnostics were examined using variance inflation factors (VIF). All VIF values remained close to 1 and well below conventional thresholds, suggesting that collinearity among predictors is low and unlikely to substantially inflate coefficient variances (Kim, 2019). All statistical analyses are conducted using IBM SPSS Statistics.

CHAPTER 5 - DATA ANALYSIS AND RESULTS

5.1. Sample characterisation

This section provides an overview of the empirical data, analysed in the study, and examines the overall sentiment expressed in online employee reviews. Specifically, the dataset comprises 779 employee reviews from 10 companies, located in Portugal, sourced from two major online job platforms: Glassdoor and Indeed. Table II presents the distribution of reviews by company and platform. The number of observations varies considerably among companies, reflecting differences in organisational size, engagement with review platforms and platform-specific usage patterns. Additionally, the distribution across platforms is uneven: Glassdoor accounts for the majority of observations, while Indeed represents a smaller share.

TABLE II - DISTRIBUTION OF EMPLOYEE REVIEWS BY COMPANY AND PLATFORM

	Glassdoor	**Indeed**	**Total**
Allianz	10	10	20
Cisco	20	6	26
Credibom Bank	10	0	10
DHL	5	19	24
ERA Imobiliária	2	28	30
EY	244	12	256
Noesis	114	4	118
Solvay	64	5	69
Synopsys Inc.	100	1	101
Vestas	124	1	125
Total	**693**	**86**	**779**

Source: Author's own elaboration based on data collected from Glassdoor and Indeed.

5.2. Descriptive and exploratory analysis of employee reviews

5.2.1. Overall employee sentiment

Employee sentiment was quantified using a continuous sentiment score ranging from -1 to +1, derived from sentiment analysis, where higher values indicate more positive evaluations and lower values indicate more negative evaluations. Descriptive statistics for the sentiment score are reported in Table III. The results show that the average sentiment score is positive (mean = 0.43), suggesting that employee reviews tend to express favourable evaluations. The median score (0.77) is notably higher than the mean, indicating that the typical review expresses a clearly positive tone, while a smaller number of reviews of strong negative reviews pull the mean downward. The relatively high standard deviation (0.63) further points to substantial dispersion in sentiment values, indicating heterogeneity in perceptions.

TABLE III - DESCRIPTIVE STATISTICS OF EMPLOYEE SENTIMENT

Variable	N	Mean	Median	Std. Dev.	Min	Max
Sentiment score	779	0.43	0.77	0.63	-0.94	0.99

Source: Author's own elaboration.

To contextualize the sentiment distribution, 65.6% of employee reviews express positive sentiment, 16.8% are neutral, and 17.6% are negative. Figure 2 complements by illustrating the distribution of sentiment scores across the sample. The histogram reveals a concentration of observations on the positive side of the scale, while still displaying a wide dispersion across the full range from -1 to +1.

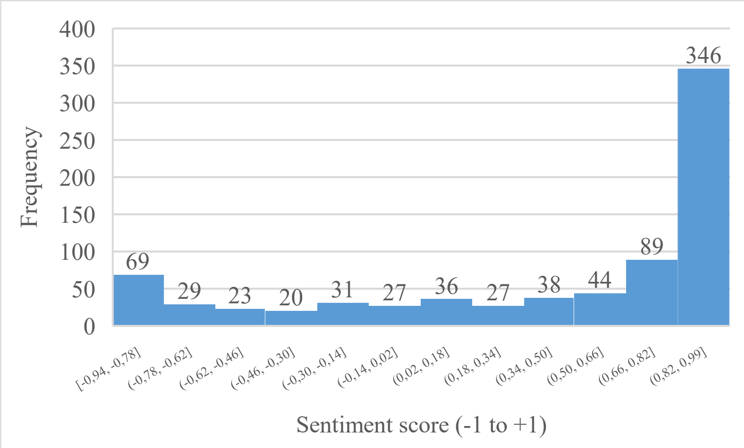

FIGURE 2 – HISTOGRAM OF EMPLOYEE SENTIMENT
Source: Author's own elaboration.

5.2.2. Employer branding dimensions reflected in online employee reviews

This section addresses H1, which proposes that the EB dimensions discussed in the literature are empirically observable in online employee reviews. In line with Tanwar and Prasad (2017), five dimensions are examined: healthy work atmosphere, work-life balance, compensation and benefits, training and development and ethics, CSR and management. The presence of these dimensions in employee reviews was analysed through a thematic analysis and for each review, a continuous relevance score (0 to 1) was assigned. For example, comments about colleagues or workplace atmosphere were treated as healthy work atmosphere, references to flexibility or long hours as work-life balance, mentions of pay or benefits as compensation, statements about learning or career growth as training and development and, finally, comments about leadership or ethical conduct as ethics, CSR and management.

Table IV reports the frequency with which each dimension appears in the dataset, expressed both as the number and percentages of reviews in which the dimension is identified. Results indicate that all five dimensions are reflected in online employee reviews, although they vary across dimensions. Healthy work atmosphere (74.5%) and compensation and benefits (58.7%) emerge most frequently whereas ethics, CSR and management appear less often (30.2%). Besides the frequency analysis, the table also presents the average relevance score for each EB dimension across the full sample. This table gives the intensity with which each dimension is addressed in the reviews.

TABLE IV- EMPLOYER BRANDING DIMENSIONS REFLECTED IN ONLINE EMPLOYEE REVIEWS

EB dimensions	Nº of reviews	% of reviews	Mean relevance score [0;1]
Healthy work atmosphere	580	74.5%	0.69
Work-life balance	448	57.5%	0.48
Compensation and benefits	457	58.7%	0.51
Training and development	293	37.6%	0.32
Ethics, CSR and Management	235	30.2%	0.21

Source: Author's own elaboration.
Note: A review is considered to reflect a given dimension when the thematic relevance score is at least 0.30.

5.2.3. Employer branding dimensions and employee sentiment

Employee sentiment, measured by the sentiment score, varies not only across companies but also depending on which aspects of the employment experience employees discuss in their reviews. For this reason, this section addresses H2 by focusing on identifying whether employee sentiment varies across EB dimensions and whether systematic differences emerge across dimensions.

To assess these relationships, employee sentiment is analysed using thematic relevance scores for the EB dimensions. For each dimension, reviews are classified as relevant if the relevance score exceeds the threshold (0.30), indicating that the dimension is substantially reflected in the review.

Table V reports the average sentiment score for each dimension, distinguishing between reviews with and without their dimension. The results reveal clear differences across dimensions. Reviews that emphasise a healthy work atmosphere are associated with substantially higher average sentiment, with the sentiment score increasing by 0.89 when this dimension is present. There are positive, though moderate, differences regarding training and development, compensation and benefits, and work-life balance. In these cases, reviews that address these dimensions tend to express more favourable sentiment than reviews that do not. This suggests that these aspects contribute positively to employees' overall evaluations, to a lesser extent than a healthy work atmosphere. The dimension related to ethics, CSR, and management exhibits a marginal difference in average sentiment (**Δ** = 0.04), suggesting that references to this dimension are not systematically associated with stronger positive sentiment and may reflect mixed employee perceptions.

TABLE V – AVERAGE EMPLOYEE SENTIMENT BY EMPLOYER BRANDING DIMENSION

	Average sentiment [-1:1]		
EB dimensions	**Theme present**	**Theme absent**	**Δ**
Healthy work atmosphere	0.66	-0.23	0.89
Work-life balance	0.48	0.36	0.12
Compensation and benefits	0.49	0.35	0.14
Training and development	0.54	0.36	0.18
Ethics, CSR and Management	0.46	0.42	0.04

Source: Author's own elaboration.

5.2.4. Employee sentiment and formal employer branding signals

This section presents descriptive evidence for H3, which examines the association between formal EB signals (GPTW ranking) and employee sentiment (derived from individual reviews and aggregated at the company level). Table VI reports the number of reviews, average employee sentiment score, and the GPTW ranking position for each company. The results show noticeable variation across companies. While some companies with higher GPTW rankings exhibit relatively positive employee sentiment, the overall pattern is mixed with substantial variation. For instance, certain organisations with lower GPTW ranking display comparatively high sentiment scores, suggesting that positive employee experiences are not exclusively associated with top-ranked companies.

These results should be interpreted within the context of the sample, but, overall, the evidence points to a heterogeneous association between formal EB signals and employee sentiment, suggesting that formal rankings do not fully capture employees' experiences. Therefore, the results provide limited and descriptive support for H3. Detailed company-level statistics are reported in the Appendix E.

TABLE VI - AVERAGE EMPLOYEE SENTIMENT AND GPTW RANKING BY COMPANY

	Nº of reviews	Average sentiment score [-1;1]	GPTW ranking position
Noesis	118	0.33	1
Allianz	20	0.37	2
Cisco	26	0.55	3
ERA Imobiliária	30	0.28	4
Vestas	125	0.48	5
DHL Supply Chain	24	0.28	6
Credibom bank	10	-0.10	7
Solvay	69	0.53	8
EY	256	0.40	9
Synopsys Inc.	101	0.60	10

Source: Author's own elaboration.

5.3. Exploratory regression analysis

The following regression analysis provides a formal test of the relationship between EB dimensions and employee sentiment. While the previous sections provided descriptive and exploratory insights, this regression analysis assesses whether this relationship holds when the different dimensions are jointly examined and statistically controlled. The dependent variable used is the continuous sentiment score derived from employee reviews.

5.3.1. Employer branding dimensions and employee sentiment

This regression analysis formally tests H2 by examining the association between EB dimensions and employee sentiment, with all dimensions considered simultaneously. To be precise, all five dimensions are included as independent variables to assess their relative association with sentiment, while controlling the other dimensions.

The coefficient estimates are presented in Table VII. The regression results are statistically significant overall ($F = 133.533$; $p < 0.001$), indicating that the EB dimensions together explain variation in employee sentiment. The model explains a substantial proportion of the variance in employee sentiment ($R^2 = 0.464$; Adjusted $R^2 = 0.460$), indicating a moderate explanatory power. Detailed model-fit statistics and diagnostic tests for the regression analysis are reported in Appendix F.

As presented in Table VII, reviews emphasizing a healthy work atmosphere are strongly and positively associated with employee sentiment. This dimension shows a strong positive coefficient ($B = 1.079$) with a small standard error ($SE = 0.044$), resulting in a high t-value ($t = 24.783$) and strong statistical significance ($p < 0.001$). It also shows the largest standardised effect in the model ($\beta = 0.685$), indicating that references to a positive work atmosphere are closely associated with more favourable sentiment in employee reviews. Training and development also show a positive, statistically significant association with sentiment, though with a smaller magnitude. This dimension has a coefficient of $B = 0.203$ ($SE = 0.058$), corresponding to a t-value of 3.520 and a p-value below 0.001. The standardized coefficient ($\beta = 0.097$) suggests a more modest, yet meaningful contribution to sentiment when controlling the other dimensions. In contrast, the other three dimensions do not show statistically significant associations in this model ($p > 0.05$), indicating that their effects weaken when other EB dimensions are taken into account. These results are consistent with the earlier descriptive findings and suggest that not all EB dimensions contribute equally to employee sentiment when considered together.

TABLE VII - OLS REGRESSION RESULTS: IMPACT OF EMPLOYER BRANDING DIMENSIONS ON EMPLOYEE SENTIMENT

Variable	B	Std. Error	Beta	t	p-value
Constant	-0.298	0.042	-	-7.090	<0.001
Healthy work atmosphere	1.079	0.044	0.685	24.783	<0.001
Work-life balance	-0.076	0.046	-0.045	-1.643	0.101
Compensation and benefits	-0.056	0.045	-0.035	-1.259	0.208
Training and development	0.203	0.058	0.097	3.520	<0.001
Ethics, CSR and management	-0.097	0.087	-0.031	-1.113	0.266

Source: Author's own elaboration.

Note: The dependent variable is the employee sentiment score, and the independent variables include the EB dimensions derived from online employee reviews.

5.4. Summary of hypothesis testing results

The analyses presented in this chapter make it possible to assess whether the proposed hypotheses are supported or not. H1 is supported since the EB dimensions identified in the literature are clearly present in online employee reviews, although some appear more frequently than others. H2 is also supported. Employee sentiment differs across EB dimensions. Reviews that focus on certain dimensions tend to express more positive sentiment. Finally, H3 is partially supported, since descriptive evidence suggests a positive but inconsistent association between the GPTW ranking and employee sentiment.

CHAPTER 6 – DISCUSSION AND CONCLUSION

6.1. Discussion of research questions

6.1.1. Discussion of RQ1 and H1

The first research question examined which EB dimensions are reflected in online employee reviews. The results suggest that employees naturally describe their work experiences using dimensions that closely align with established EB frameworks, namely Tanwar and Prasad’s (2017). This reinforces the view that EB is a multidimensional concept, not only communicated by organizations but also shaped by employees’ narratives (Lievens & Slaughter, 2016).

From a theoretical perspective, the presence of these dimensions in online reviews highlights the relationship between EB and employer reputation. While EB traditionally refers to intended and controlled organisational messages (Ambler & Barrow, 1996; Backhaus & Tikoo, 2004), employer reputation reflects how organisations are perceived in practice (Highhouse et al., 2009). The fact that employees discuss the same dimensions in their online reviews suggests that the aspects organisations use to promote themselves are also the criteria employees use to evaluate them. This means EB and employer reputation are structured around the same core dimensions of the employment experience. Online employee reviews make this connection visible, as they show whether the communicated employer brand is affirmed or challenged by employee voice.

Regarding frequency, some dimensions occur more often than others. Healthy work atmosphere, compensation and benefits, and work-life balance are the most discussed dimensions, meaning that these themes are especially relevant in employees' lived experiences. This aligns with prior EB research, which highlights these aspects as important elements of the EVP (Berthon et al., 2005; Tanwar & Prasad, 2017).

The analysis also highlights the value of using text mining measures to capture EB dimensions from large volumes of unstructured textual data. By applying thematic relevance scores, the analysis enables the systematic identification and comparison of EB dimensions derived from employees’ own words (Morgan & Chapman, 2024). The discussion of RQ1 shows that informal EB signals, such as online employee reviews, provide a reliable basis for identifying established EB frameworks (Symitsi et al., 2021).

6.1.2. Discussion of RQ2 and H2

The second research question focuses on the relationship between EB dimensions and employee sentiment. The findings suggest that these dimensions are associated with employee sentiment in uneven ways, with some having a stronger influence than others.

Theoretically, this pattern aligns with the distinction between instrumental and symbolic attributes, as dimensions related to daily interactions and the work environment tend to evoke stronger emotional responses than more transactional aspects of employment (Lievens & Highhouse, 2003; Lievens & Slaughter, 2016). This aligns with the finding that healthy work atmosphere is the dimension most strongly associated with positive sentiment, whereas compensation and benefits, ethics, CSR, and management show weaker associations. In Tanwar and Prasad's (2017) study, healthy work atmosphere, training and development, and work-life balance are identified as the core dimensions of a strong employer brand. Regarding its association with employee sentiment, this study partially supports this view. Healthy work atmosphere and training and development show clear and significant associations with employee sentiment. While work-life balance does not reach statistical significance in the model, it remains comparatively stronger than the other non-significant dimensions, indicating some consistency with Tanwar and Prasad's study.

An important insight emerging from this is that EB dimensions with higher relevance scores in online reviews are not necessarily the ones most strongly associated with employee sentiment. For example, compensation and benefits and work-life balance are widely discussed by employees, but they appear to play a more limited role in overall employee sentiment. In contrast, training and development appear less prominent in employee narratives, yet it shows a comparatively stronger association with employee sentiment. This means that what employees talk about most is not necessarily what matters most for their emotional evaluations. These results are consistent with the literature on eWOM, which emphasises that online reviews tend to amplify what people feel and experience most strongly at work (Verma & Yadav, 2022), explaining why experiential dimensions are more strongly associated with sentiment.

Both the descriptive and regression analyses indicate that EB dimensions differ in the strength of their association with employee sentiment. While the descriptive results show uneven sentiment differences across dimensions, the regression analysis further clarifies which dimensions retain an independent, strong, and statistically significant influence when considered jointly, providing clear support for H2.

6.1.3. Discussion of RQ3 and H3

The third research question examined whether employee sentiment is associated with formal EB signals, namely the GPTW ranking. The findings indicate a positive, although inconsistent, association between GPTW ranking and employee sentiment. While some higher-ranked companies exhibit more positive employee sentiment, this pattern is not consistently observed across the sample.

In theory, these results are consistent with signaling theory. Certifications and rankings, such as GPTW, serve as positive signals of employer quality (Dineen & Allen, 2016; Zhang et al., 2022), reducing information asymmetry (Connelly et al., 2011). In this sense, GPTW can be understood as a public signal of employer attractiveness. However, the descriptive nature of the analysis and the inconsistency observed in the results suggest that formal rankings capture only part of employees' perceptions. While GPTW may signal overall quality or public reputation, it does not fully reflect employees' lived experiences expressed in online reviews (Dabirian et al., 2017).

The results highlight the distinction between formal EB signals and informal employee-generated signals. GPTW provides a general and externally validated indicator of employer quality, whereas employee sentiment reflects subjective detailed perceptions.

This suggests that employees do not rely on a single source of information when forming perceptions of employers, but rather consider multiple signals simultaneously (Drover et al., 2018). In this context, formal rankings and online employee reviews are jointly interpreted, reinforcing the idea that employer perceptions are shaped by complementary sources of information.

6.2. Academic contributions

This study contributes to the EB and signaling literature in several important ways. First, it extends EB research by showing that the dimensions proposed in the literature can be empirically identified in unstructured online employee reviews using text-mining techniques. While prior studies largely relied on surveys or organization-controlled measures (e.g., Ambler & Barrow, 1996; Backhaus & Tikoo, 2004; Berthon et al., 2005; Tanwar & Prasad, 2017), this study demonstrates how unstructured textual data can be used to operationalize EB in a data-driven manner.

Second, the study advances research on online employee reviews by linking EB dimensions to employee sentiment. In line with research treating online reviews as signals of perceptions (Dabirian et al., 2017; Symitsi et al., 2021; Morgan & Chapman, 2024), the results show that not all dimensions contribute equally to sentiment. Importantly, the results highlight a distinction between the salience of dimensions in employee narratives and their actual impact, adding value to existing work on eWOM and employee voice.

Third, the research contributes to signaling theory by examining both formal and informal EB signals. Prior studies have emphasized the role of third-party certifications and rankings as credible signals of employer quality (Connelly et al., 2011; Dineen & Allen, 2016), as well as the informational value of reviews (Willems et al., 2017; Zhang et al., 2022). By examining both types of signals within the same research context, the study shows that formal certifications, namely the GPTW one, and online employee reviews appear to be broadly aligned in direction, although they capture different aspects of the employment experience.

Finally, this study contributes to the body of research that uses employee-generated data on online platforms to examine organizational phenomena (Gimpl, 2025; Jung & Suh, 2019). It illustrates how computational text analysis can be combined with established theoretical frameworks in EB and signaling research, helping link employee narratives to quantitative analysis.

6.3. Contribution to management

This study provides insights for organizations seeking to strengthen their employer brand. Online employee reviews emerge as a valuable source of information about how employees perceive their workplace (Dabirian et al., 2017), complementing more formal feedback mechanisms.

The results also show that EB dimensions do not influence sentiment equally, so organizations may benefit from rethinking their EB priorities. While compensation and work-life balance are highly visible and frequently discussed, improvements in the work atmosphere and training and development opportunities appear to have a stronger effect on how employees feel about their employer.

In addition, formal EB signals, such as the GPTW certification, show a positive, although descriptive, association with employee sentiment, but do not fully reflect employees' lived experiences (Dineen & Allen, 2016). Together, the findings suggest that organizations benefit from combining formal initiatives with employee-generated feedback in order to build a more authentic employer brand. The practical implications discussed are especially applicable to organizations operating in Portugal, where the empirical analysis was conducted.

6.4. Study limitations and future research

The first limitation of the study lies in the analysis of formal and informal EB signals, as differences in the number of online employee reviews across companies may reduce the stability and comparability of average sentiment measures. Second, the empirical analysis is limited to a small set of companies within Portugal, which may limit the generalization of the findings, as EB practices and the visibility of formal certification vary across countries and industries (Dineen & Allen, 2016). Third, online employee reviews are used as a proxy for employee sentiment. Although their informational value is well established (Symitsi et al., 2021), these reviews are voluntary and may overrepresent more extreme experiences. Finally, although the analysis includes employee reviews from multiple years, the data are aggregated. Therefore, the study identifies associations between EB signals and employee sentiment but cannot determine cause-and-effect relationships.

Future research could address these limitations. For example, extending the analysis to larger, more diverse samples, including organizations from different countries and industries, would help assess the robustness of the findings. Besides that, combining online employee reviews with internal survey data or other organizational indicators could provide a more balanced view of employee perceptions. In addition, longitudinal research designs could examine how changes in EB signals, such as the acquisition of formal certifications, are associated with changes in employee sentiment.

REFERENCES

Ambler, T., & Barrow, S. (1996). The employer brand. *Journal of Brand Management*, *4*. https://doi.org/10.1057/bm.1996.42

Alves, P., Santos, V., Reis, I., Martinho, F., Martinho, D., Correia Sampaio, M., José Sousa, M., & Au-Yong-Oliveira, M. (2020). Strategic Talent Management: The Impact of Employer Branding on the Affective Commitment of Employees. *Sustainability*, *12*(23), 9993. https://doi.org/10.3390/su12239993

Backhaus, K., & Tikoo, S. (2004). Conceptualizing and researching employer branding. *Career Development International*, *9*(5), 501–517. https://doi.org/10.1108/13620430410550754

Benraïss-Noailles, L., & Viot, C. (2021). Employer brand equity effects on employees well-being and loyalty. *Journal of Business Research*, *126*, 605–613. https://doi.org/10.1016/j.jbusres.2020.02.002

Berthon, P., Ewing, M., & Hah, L. L. (2005). Captivating company: Dimensions of attractiveness in employer branding. *International Journal of Advertising: The Quarterly Review of Marketing Communications,* *24*(2), 151–172. https://doi.org/10.1080/02650487.2005.11072912

Cable, D. M., & Turban, D. B. (2003). The value of organizational reputation in the recruitment context: A brand-equity perspective. *Journal of Applied Social Psychology, 33*(11), 2244–2266. https://doi.org/10.1111/j.1559-1816.2003.tb01883.x

Carpentier, M., Van Hoye, G., & Weijters, B. (2019). Attracting Applicants Through the Organization's Social Media Page: Signaling Employer Brand Personality. *Journal of Vocational Behavior*, *115*, 103326. https://doi.org/10.1016/j.jvb.2019.103326

Carson, J. E., & Westerman, J. W. (2023). The Effectiveness of Organizational Sustainability Messaging to New Hires: An Exploratory Analysis of Signal Cost, Perceived Credibility, and Involvement Intention. Sustainability, 15(2), 1167. https://doi.org/10.3390/su15021167

Chai, C. P. (2023). Comparison of text preprocessing methods. *Natural Language Engineering*, *29*(3), 509–553. https://doi.org/10.1017/S1351324922000213

Choudhury, M. (2024). Signaling Theory: An Approach to Organizational Behavior Research. *Journal of Accounting Business and Management (JABM)*, *31*, 98–120. https://doi.org/10.31966/jabminternational.v31i2.1199

Connelly, B. L., Certo, S. T., Ireland, R. D., & Reutzel, C. R. (2011). Signaling theory: A review and assessment. Journal of Management, 37, 39–67. https://doi.org/10.1177/0149206310388419

Connelly, B. L., Certo, S. T., Reutzel, C. R., DesJardine, M. R., & Zhou, Y. S. (2025). Signaling Theory: State of the Theory and Its Future. *Journal of Management*, *51*(1), 24–61. https://doi.org/10.1177/01492063241268459

Dabirian, A., Kietzmann, J., & Diba, H. (2017). A great place to work!?: Understanding crowdsourced employer branding. *Business Horizons*, 60, 197-205. https://doi.org/10.1016/j.bushor.2016.11.005

Dell, D., Ainspan, N., T., B., Troy, K. & Hickey, J. (2001), Engaging Employees Through Your Brand, *Research Report 1288-01-RR*, New York.

Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. *Proceedings of the 2019 Conference of the North*, 4171–4186. https://doi.org/10.18653/v1/N19-1423

Dineen, B., & Allen, D. (2015). Third Party Employment Branding: Human Capital Inflows and Outflows Following «Best Places to Work» Certifications. *Academy of Management Journal*, *59*. https://doi.org/10.5465/amj.2013.1091

Donthu, N., Kumar, S., Pandey, N., Pandey, N., & Mishra, A. (2021). Mapping the electronic word-of-mouth (eWOM) research: A systematic review and bibliometric analysis. *Journal of Business Research*, *135*, 758–773. https://doi.org/10.1016/j.jbusres.2021.07.015

Drover, W., Wood, M. S., & Corbett, A. C. (2018). Toward a Cognitive View of Signalling Theory: Individual Attention and Signal Set Interpretation. *Journal of Management Studies*, *55*(2), 209–231. https://doi.org/10.1111/joms.12282

Glassdoor. (n.d.). *Glassdoor*. [Online] https://www.glassdoor.com/ [Retrieved: 2025/10/02]

Great Place To Work. (n.d.). *Great Place To Work*. [Online] https://www.greatplacetowork.com/ [Retrieved: 2025/10/07]

Gimpl, N. (2025). The Wisdom of Electronic Employee Crowds – Employee Reviews as a Data Source in Finance, Accounting, Economics, and Management Research: A Systematic Literature Review. *Intelligent Systems in Accounting, Finance and Management*, *32*(2), e70007. https://doi.org/10.1002/isaf.70007

Guillot-Soulez, C., Saint-Onge, S., & Soulez, S. (2022). Green certification and organizational attractiveness: The moderating role of firm ownership. *Corporate Social Responsibility and Environmental Management*, *29*(1), 189–199. https://doi.org/10.1002/csr.2194

Guo, Y., Barnes, S. J., & Jia, Q. (2017). Mining meaning from online ratings and reviews: Tourist satisfaction analysis using latent dirichlet allocation. *Tourism Management*, *59*, 467–483. https://doi.org/10.1016/j.tourman.2016.09.009

Highhouse, S., Broadfoot, A., Yugo, J., & Devendorf, S. (2009). Examining Corporate Reputation Judgments With Generalizability Theory. *Journal of Applied Psychology*, *94*, 782–789. https://doi.org/10.1037/a0013934

Holland, P., Cooper, B. K., & Hecker, R. (2016). Use of social media at work: A new form of employee voice? *The International Journal of Human Resource Management*, *27*(21), 2621–2634. https://doi.org/10.1080/09585192.2016.1227867

Joglekar, J., & Tan, C. S. L. (2022). The impact of LinkedIn posts on employer brand perception and the mediating effects of employer attractiveness and corporate reputation. *Journal of Advances in Management Research*, *19*. https://doi.org/10.1108/JAMR-10-2021-0343

Junça Silva, A., & Dias, H. (2023). The relationship between employer branding, corporate reputation and intention to apply to a job offer. *International Journal of Organizational Analysis*, *31*(8), 1–16. https://doi.org/10.1108/IJOA-01-2022-3129

Kanwal, H., & Van Hoye, G. (2024). Beyond employer brand content: The role of employer brand process attributes in understanding employees' reactions toward their employer. *Human Resource Management*, *63*(1), 67–86. https://doi.org/10.1002/hrm.22187

Kim, J. H. (2019). Multicollinearity and misleading statistical results. *Korean Journal of Anesthesiology*, *72*(6), 558–569. https://doi.org/10.4097/kja.19087

Kobayashi, V. B., Mol, S. T., Berkers, H. A., Kismihók, G., & Den Hartog, D. N. (2018). Text Mining in Organizational Research. *Organizational Research Methods*, *21*(3), 733–765. https://doi.org/10.1177/1094428117722619

Kurniasih, A., & Manik, L. P. (2022). On the Role of Text Preprocessing in BERT Embedding-based DNNs for Classifying Informal Texts. *International Journal of Advanced Computer Science and Applications*, *13*(6). https://doi.org/10.14569/IJACSA.2022.01306109

Kyritsis, K., Perikos, I., & Paraskevas, M. (2023). Zero-Shot Sentiment Analysis Exploring BART Models. *2023 IEEE/ACIS 8th International Conference on Big Data, Cloud Computing, and Data Science (BCD)*, 192–197. https://doi.org/10.1109/BCD57833.2023.10466289

Lievens, F., & Highhouse, S. (2003). The Relation of Instrumental and Symbolic Attributes to a Company's Attractiveness as an Employer. *Personnel Psychology*, *56*, 75–102. https://doi.org/10.1111/j.1744-6570.2003.tb00144.x

Lievens, F., Van Hoye, G., & Schreurs, B. (2005). Examining the relationship between employer knowledge dimensions and organizational attractiveness: An application in a military context. *Journal of Occupational and Organizational Psychology*, *78*(4), 553–572. https://doi.org/10.1348/09631790X26688

Lievens, F., & Slaughter, J. E. (2016). Employer Image and Employer Branding: What We Know and What We Need to Know. *Annual Review of Organizational Psychology and Organizational Behavior*, *3*, 407-440. https://doi.org/10.1146/annurev-orgpsych-041015-062501

Liu, Y., Ott, M., Goyal, N., Du, J., Joshi, M., Chen, D., Levy, O., Lewis, M., Zettlemoyer, L., & Stoyanov, V. (2019). *RoBERTa: A Robustly Optimized BERT Pretraining Approach* (Version 1). arXiv. https://doi.org/10.48550/ARXIV.1907.11692

Love, L., & Singh, P. (2011). Workplace Branding: Leveraging Human Resources Management Practices for Competitive Advantage Through "Best Employer" Surveys. *Journal of Business and Psychology*, 26, 175-181. https://doi.org/10.1007/s10869-011-9226-5

Mohiya, M. (2024). What constitutes an employer of choice? A qualitative triangulation investigation. *Human Resources for Health*, *22*(1), 41. https://doi.org/10.1186/s12960-024-00928-7

Morgan, J. A., & Chapman, D. S. (2024). Examining the characteristics and effectiveness of online employee reviews. *Computers in Human Behavior Reports*, *16*, 100471. https://doi.org/10.1016/j.chbr.2024.100471

Pawar, A. (2016). *Employee Value Proposition: A Collaborative Methodology for Strengthening Employer Brand Strategy*.

Pernkopf, K., Latzke, M., & Mayrhofer, W. (2021). Effects of mixed signals on employer attractiveness: A mixed-method study based on signalling and convention theory. *Human Resource Management Journal*, *31*(2), 392–413. https://doi.org/10.1111/1748-8583.12313

Pipalia, K., Bhadja, R., & Shukla, M. (2020). Comparative Analysis of Different Transformer Based Architectures Used in Sentiment Analysis. *2020 9th International Conference System Modeling and Advancement in Research Trends (SMART)*, 411–415. https://doi.org/10.1109/SMART50582.2020.9337081

Reis, I., Sousa, M. J., & Dionísio, A. (2021). Employer Branding as a Talent Management Tool: A Systematic Literature Revision. Sustainability, 13(19), 10698. https://doi.org/10.3390/su131910698

Saini, G. K., & Jawahar, I. M. (2019). The influence of employer rankings, employment experience, and employee characteristics on employer branding as an employer of choice.

Career Development International, *24*(7), 636–657. https://doi.org/10.1108/CDI-11-2018-0290

Sainju, B., Hartwell, C., & Edwards, J. (2021). Job satisfaction and employee turnover determinants in Fortune 50 companies: Insights from employee reviews from Indeed.com. *Decision Support Systems*, *148*, 113582. https://doi.org/10.1016/j.dss.2021.113582

Spence, M. (1973). Job Market Signaling. *The Quarterly Journal of Economics*, *87*(3), 355. https://doi.org/10.2307/1882010

Stockman, S., Van Hoye, G., & Veiga, S. (2020). Negative word-of-mouth and applicant attraction: The role of employer brand equity. *Journal of Vocational Behavior*, *118*, 103368. https://doi.org/10.1016/j.jvb.2019.103368

Symitsi, E., Stamolampros, P., Daskalakis, G., & Korfiatis, N. (2021). The informational value of employee online reviews. *European Journal of Operational Research*, *288*(2), 605–619. https://doi.org/10.1016/j.ejor.2020.06.001

Tanwar, K., & Prasad, A. (2017). Employer brand scale development and validation: A second-order factor approach. *Personnel Review*, *46*(2), 389–409. https://doi.org/10.1108/PR-03-2015-0065

Theurer, C., Tumasjan, A., Welpe, I., & Lievens, F. (2016). Employer Branding: A Brand Equity-based Literature Review and Research Agenda: Brand Equity-Based Employer Branding. *International Journal of Management Reviews*, *20*. https://doi.org/10.1111/ijmr.12121

Turban, D. B. (2001). Organizational attractiveness as an employer on college campuses: An examination of the applicant population. Journal of Vocational Behavior, 58, 293–312. https://doi.org/10.1006/jvbe.2000.1765

Uribe, A., & Ríos, M. (2024). The "great place to work" certification as an employer branding promotional startegy in the recruitment of colombian SME´s. *Tendencias*. https://doi.org/10.22267/rtend.242501.245

Van Hoye, G., & Lievens, F. (2007). Social Influences on Organizational Attractiveness: Investigating If and When Word of Mouth Matters[1]. *Journal of Applied Social Psychology*, *37*(9), 2024–2047. https://doi.org/10.1111/j.1559-1816.2007.00249.x

Verma, S., & Yadav, N. (2021). Past, Present, and Future of Electronic Word of Mouth (EWOM). *Journal of Interactive Marketing*, *53*(1), 111–128. https://doi.org/10.1016/j.intmar.2020.07.001

Wang, L., Han, J., Ramasamy, B., & Peng, S. (2022). Incongruous employer brand signals and organizational attractiveness: Evidence from multinational companies in China. *Human Resource Management*, *61*(5), 563–584. https://doi.org/10.1002/hrm.22121

Wankhade, M., Rao, A. C. S., & Kulkarni, C. (2022). A survey on sentiment analysis methods, applications, and challenges. *Artificial Intelligence Review*, *55*(7), 5731–5780. https://doi.org/10.1007/s10462-022-10144-1

Wilden, R., Gudergan, S., & Lings, I. (2010). Employer branding: strategic implications for staff recruitment. *Journal of Marketing Management*, *26*(1–2), 56–73. https://doi.org/10.1080/02672570903577091

Willems, J., Waldner, C. J., Dere, Y. I., Matsuo, Y., & Högy, K. (2017). The Role of Formal Third-Party Endorsements and Informal Self-Proclaiming Signals in Nonprofit Reputation Building. *Nonprofit and Voluntary Sector Quarterly*, *46*(5), 1092–1105. https://doi.org/10.1177/0899764017720770

Yin, W., Hay, J., & Roth, D. (2019). Benchmarking Zero-shot Text Classification: Datasets, Evaluation and Entailment Approach. *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, 3912–3921. https://doi.org/10.18653/v1/D19-1404

Yu, K. Y. T., Goh, K. H., & Kawasaki, S. (2022). Investigating electronic word-of-mouth on social media: An eye-tracking approach. *Human Resource Management*, *61*(5), 599–616. https://doi.org/10.1002/hrm.22099

Zhang, Y. (Carrie), Shum, C., & Belarmino, A. (2022). "Best Employers": The Impacts of Employee Reviews and Employer Awards on Job Seekers' Application Intentions.

Cornell Hospitality Quarterly, *64*(3), 298–306. https://doi.org/10.1177/19389655221130741

Zhou, M., Li, Y., Qiao, Z., & Shi, B. (2024). Employee Ratings and Reviews Data from Glassdoor. *Journal of Information Systems*, *38*(3), 93–105. https://doi.org/10.2308/ISYS-2023-003

DECLARATION

During the preparation of this dissertation, the author used AI-based tools at different stages of the work. First, Grammarly was used for language and grammar revision, Deepl for translation when needed and Consensus to identify relevant academic literature. Finally, ChatGPT was used to support the organization of specific sections. All content was critically reviewed and edited by the author. I understand the importance of maintaining academic integrity and take full responsibility for the content, analysis and originality of the work.

APPENDICES

Appendix A - Web scraping setup and configuration – Python code

```
#------------------ Libraries ------------------
from selenium import webdriver
from selenium.webdriver.common.by import By
from selenium.webdriver.chrome.options import Options
from selenium.webdriver.support.ui import WebDriverWait
from selenium.webdriver.support import expected_conditions as EC
import pandas as pd
import time

options = Options()
options.add_argument("--start-maximized")

driver = webdriver.Chrome(options=options)

url = https://www.glassdoor.com/Reviews/NOESIS-Portugal-Reviews-
E593391.htm?filter.iso3Language=eng #for each company
driver.get(url)
time.sleep(3)

wait = WebDriverWait(driver, 10)

def extrair_reviews_pagina():
    """Extrai todas as reviews da página actual e devolve uma lista
de dicts."""
    reviews_data = []

    reviews = wait.until(
        EC.presence_of_all_elements_located(
            (By.XPATH,
"//div[@id='ReviewsFeed']//ol[contains(@class,'ReviewsList_reviewsLi
st')]/li")
        )
    )

    print("Reviews nesta página:", len(reviews))

    for idx, rev in enumerate(reviews, start=1):
        try:
            titulo = rev.find_element(
                By.XPATH,
                "./div/div/div[2]/div[1]/a"
            ).text

            rating = rev.find_element(
                By.XPATH,
                "./div/div/div[1]/div[1]/div/div[2]/span"
            ).text

            job_title = rev.find_element(
                By.XPATH,
                "./div/div/div[2]/div[1]/div[1]/div[2]/span"
            ).text

            employment_time = rev.find_element(
            By.XPATH,
            "./div/div/div[2]/div[1]/div[2]/div/div/div"
```

```
            ).text

            pros = rev.find_element(
                By.XPATH,
                "./div/div/div[2]/div[2]/div/div[2]/p[2]/span"
            ).text

            cons = rev.find_element(
                By.XPATH,
                "./div/div/div[2]/div[2]/div/div[3]/p[2]/span"
            ).text

            date = rev.find_element(
            By.XPATH,
            "./div/div/div[1]/div[2]/span"
            ).text

            reviews_data.append({
                "titulo": titulo,
                "rating": rating,
                "job_title": job_title,
                "employment_time": employment_time,
                "pros": pros,
                "cons": cons,
                "date": date
            })

        except Exception as e:
            print(f"Erro ao extrair review {idx}:", e)

    return reviews_data

todas_reviews = []
page_num = 1

while True:
    print(f"\n=== PÁGINA {page_num} ===")
    todas_reviews.extend(extrair_reviews_pagina())

    try:
        next_btn = wait.until(
            EC.element_to_be_clickable(
                (By.XPATH, "//button[@data-test='next-page' and
not(@disabled)]")
            )
        )

        driver.execute_script("arguments[0].scrollIntoView({block:
'center'});", next_btn)
        time.sleep(1)

        next_btn.click()
        time.sleep(2)
        page_num += 1

    except TimeoutException:
        print("Não há mais botão 'Next' — fim da paginação.")
        break

print(f"\nTOTAL DE REVIEWS RECOLHIDAS: {len(todas_reviews)}")
```

```
# ----------------- Criar DataFrame -----------------
df = pd.DataFrame(todas_reviews, columns=["titulo", "rating",
"job_title", "employment_time", "pros", "cons", "date"])

print(df.head())
```

Appendix B - Translation and concatenation (Data preprocessing)

```
#-------------- Libraries ----------------
import re
import numpy as np
import pandas as pd
from tqdm.auto import tqdm
from transformers import pipeline
from deep_translator import GoogleTranslator

# For each company dataset
translator = GoogleTranslator(source='auto', target='en')

def traduzir(texto):
    if pd.isna(texto) or texto.strip() == "":
        return ""
    try:
        return translator.translate(texto)
    except:
        return texto

TITLE_COL = "titulo"
PROS_COL  = "pros"
CONS_COL  = "cons"

def clean_text(x):
    if x is None or (isinstance(x, float) and np.isnan(x)):
        return ""
    x = str(x)
    x = re.sub(r"\s+", " ", x).strip()
    return x

def build_full_review(row):
    parts = [
        clean_text(row[TITLE_COL]),
        clean_text(row[PROS_COL]),
        clean_text(row[CONS_COL]),
    ]
    text = " | ".join([p for p in parts if p])
    return text[:MAX_CHARS]
```

Appendix C - Sentiment analysis (Transformer-based RoBERTa model)

```
MODEL_NAME = "cardiffnlp/twitter-roberta-base-sentiment-latest"
BATCH_SIZE = 32
MAX_CHARS = 6000


# ---------- SENTIMENT MODEL ----------
sentiment_pipe = pipeline(
    "sentiment-analysis",
```

```
    model=MODEL_NAME,
    device=-1
)

# Get probabilities for ALL labels
all_scores = sentiment_pipe(
    df["full_review_text"].tolist(),
    truncation=True,
    return_all_scores=True
)

# Convert list-of-lists into a DataFrame of probs
probs = pd.DataFrame([{d["label"].lower(): d["score"] for d in row}
for row in all_scores])

# Ensure expected columns exist (in case a label is missing in some
outputs)
for c in ["negative", "neutral", "positive"]:
    if c not in probs.columns:
        probs[c] = 0.0

df["p_negative"] = probs["negative"]
df["p_neutral"]  = probs["neutral"]
df["p_positive"] = probs["positive"]

# Continuous sentiment in [-1, 1]
df["sentiment_score_pm1"] = df["p_positive"] - df["p_negative"]

# Predicted label
df["sentiment_label"]                                              =
probs[["negative","neutral","positive"]].idxmax(axis=1)
df["sentiment_prob"] = probs.max(axis=1)  # confidence of top class

# ---------- QUICK SUMMARY ----------
print("\nSentiment distribution:")
print(df["sentiment_label"].value_counts())
```

Appendix D - Thematic analysis (Zero-shot BART)

```
# Load zero-shot classifier
theme_classifier = pipeline(
    "zero-shot-classification",
    model="facebook/bart-large-mnli",
    device=0
)

THEMES = [
    "compensation and benefits",
    "training and development",
    "work life balance",
    "healthy work atmosphere",
    "ethics, csr and management"
]

texts = df["full_review_text"].tolist()

results = []
```

```
for text in tqdm(texts, desc="Assigning themes"):
    output = theme_classifier(
        text,
        candidate_labels=THEMES,
        multi_label=True
    )
    results.append(
        {label: score for label, score in zip(output["labels"],
output["scores"])}
    )

themes_df = pd.DataFrame(results)
df          =          pd.concat([df.reset_index(drop=True),
themes_df.reset_index(drop=True)], axis=1)
```

Appendix E - Company-level statistics

TABLE E1 - DESCRIPTIVE STATISTICS OF EMPLOYEE SENTIMENT

	Nº Reviews	Avg. sentiment score	Std. Deviation	% Negative	% Positive
Allianz	20	0.37	0.76	25%	65%
Cisco	26	0.55	0.58	77%	12%
Credibom	10	-0.10	0.70	60%	40%
DHL	24	0.28	0.79	33%	58%
ERA	30	0.28	0.72	30%	53%
EY	256	0.40	0.60	15%	60%
Noesis	118	0.33	0.72	27%	63%
Solvay	69	0.53	0.55	9%	70%
Synopsys	101	0.60	0.54	12%	80%
Vestas	125	0.48	0.56	14%	70%

Source: Author’s own elaboration.

TABLE E2 - Δ = SENTIMENT (THEME PRESENT) − SENTIMENT (THEME ABSENT)

	Healthy work atmosphere	Work-life balance	Compensation and benefits	Training and development	Ethics, CSR and management
Allianz	1.66	0.00	0.16	0.13	0.07
Cisco	1.06	-0.04	0.08	0.18	-0.51
Credibom	0.89	0.53	-0.03	1.21	0.10
DHL	1.23	0.71	0.05	0.28	0.20
ERA	1.25	0.63	1.03	0.44	-0.16
EY	0.80	0.11	0.10	0.12	0.13
Noesis	1.24	0.15	-0.05	0.30	-0.31
Solvay	0.53	-0.08	-0.15	0.18	0.10
Synopsys	0.92	0.08	0.10	0.23	0.27
Vestas	0.65	-0.02	0.26	0.19	0.21

Source: Author’s own elaboration.

Appendix F - Statistical analysis for the regression analysis

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	,681[a]	,464	,460	,46334010832 0967

a. Predictors: (Constant), ethics_csr_management, healthy_work_environment, work_life_balance, training_development, compensation_benefits

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	143,337	5	28,667	133,533	<,001[b]
	Residual	165,736	772	,215		
	Total	309,073	777			

a. Dependent Variable: sentiment_score_pm1

b. Predictors: (Constant), ethics_csr_management, healthy_work_environment, work_life_balance, training_development, compensation_benefits

Coefficients[a]

Model		Sig.	Collinearity Statistics	
			Tolerance	VIF
1	(Constant)	<,001		
	healthy_work_environm ent	<,001	,910	1,099
	work_life_balance	,101	,917	1,091
	compensation_benefits	,208	,908	1,101
	training_development	<,001	,910	1,099
	ethics_csr_management	,266	,900	1,111

Collinearity Diagnostics[a]

Model	Dimension	Variance Proportions		
		compensation _benefits	training_devel opment	ethics_csr_ma nagement
1	1	,01	,01	,01
	2	,09	,37	,27
	3	,05	,48	,61
	4	,36	,08	,06
	5	,46	,03	,00
	6	,03	,03	,03

Collinearity Diagnostics[a]

Model	Dimension	Eigenvalu e	Condition Index	Variance Proportions		
				(Constant)	healthy_work_ environment	work_life_bala nce
1	1	4,535	1,000	,01	,01	,01
	2	,464	3,127	,01	,05	,06
	3	,351	3,594	,00	,00	,05
	4	,299	3,895	,00	,01	,66
	5	,233	4,412	,05	,44	,19
	6	,118	6,204	,93	,49	,03