



DEEP Factors: Unlocking hidden value in unstructured data

Leveraging Advanced AI to transform qualitative information into actionable investment insights

December 2024

About Entis

Entis is a data-driven investment intelligence company specializing in advanced AI combined with analyst expertise to deliver highly precise analyses of global corporate dynamics. By transforming complex qualitative data into actionable insights, Entis empowers more informed investment decisions. With a commitment to reliability, transparency, and precision, Entis provides data insights that deepen understanding of the investment landscape.

For more information, please visit: www.entis.ai

Authors



Machiel Westerdijk, PhD

Founder and Managing Director
of Entis



Marco Della Seta, PhD

Quant engineer and investment
manager



Olivera Rakic, PhD

Head of quant NLP research at
Entis

Executive summary

This whitepaper introduces “DEEP Factors” (Difficult-to-quantify Essential Elements of Performance), which utilize advanced technologies such as artificial intelligence (AI) and natural language processing (NLP) to analyze qualitative data often overlooked by traditional investment strategies. Drawing from diverse data sources, including corporate filings, patent information, and employee reviews, DEEP Factors uncover insights into a company’s long-term performance potential.

DEEP Factors are designed to assess qualitative and intangible attributes, offering a clearer understanding of a company’s core characteristics and market position. This framework enhances investment analyses by integrating complex, unstructured datasets into the decision-making process.

The paper discusses how DEEP Factors broaden traditional investment models, outlining the technologies and methodologies required to extract actionable insights. It explores their applications across various strategies, including active quant strategies, smart beta products, and ETFs.

Supported by academic research and case studies, this paper highlights how DEEP Factors capture critical insights often missed by conventional approaches. By incorporating these factors, investors gain a robust framework to achieve sustained performance in increasingly complex markets.

01 Introduction and Market Need

Investing successfully hinges on understanding market dynamics and identifying companies with strong potential for success. For decades, factor investing has served as a powerful tool for this purpose. It systematically links quantifiable metrics derived from financial statements—such as book-to-market ratios or other valuation indicators—to expected returns. Factors like these have allowed investors to identify patterns and allocate resources with greater confidence.

However, traditional factors have faced challenges in recent years. Their overexploitation through quantitative strategies and commoditization via exchange-traded funds (ETFs) have diluted their competitive edge. As more investors adopt these strategies, the once-powerful predictive signals embedded in traditional factors have weakened, leaving a gap in the ability to achieve differentiated performance.

To maintain an edge, investors must expand their focus beyond readily available metrics. The future lies in developing new factors that capture qualitative, intangible, and harder-to-quantify elements—insights that remain hidden in today’s increasingly complex, data-rich environment. These advanced factors offer an opportunity to uncover underutilized information and pave the way for stronger, more sustainable returns.

Introducing DEEP Factors: A New Paradigm in Factor Investing

DEEP Factors—Difficult-to-quantify Essential Elements of Performance—represent the next step in factor investing. They leverage cutting-edge technologies like artificial intelligence (AI) and natural language processing (NLP) to extract and analyze qualitative data often overlooked by traditional models. By focusing on sources such as company filings, patent documents, and employee reviews, DEEP Factors systematically identify and measure fundamental attributes like governance, innovation, and market positioning.

Unlike many non-traditional factors, such as signals derived from news sentiment or social media, DEEP Factors prioritize stability and long-term relevance. While short-lived factors can provide value for trading strategies, they often lack the durability needed for long-term growth. DEEP Factors, in contrast, target enduring company characteristics that are more predictive of sustained performance.

By transforming complex, unstructured data into actionable insights, DEEP Factors address a persistent gap in how markets process qualitative information. Key sources of valuable insights, such as annual reports or earnings call transcripts, remain underutilized due to their complexity. DEEP Factors bring these insights to light, equipping investors with tools to refine strategies and uncover hidden opportunities in an increasingly competitive landscape.

02 Application of DEEP Factors in Investment Strategies

DEEP Factors leverage qualitative and hard-to-measure data to provide insights into areas such as governance, innovation, and peer dynamics. This section explores the academic foundations supporting these factors and their role in enhancing modern investment strategies.

Academic Evidence Supporting DEEP Factors

DEEP Factors excel in capturing hard-to-measure, fundamental drivers critical for long-term performance. This section highlights key academic research that underscores their value, focusing on areas such as innovation, networks and peers, and governance and management, where DEEP Factors provide nuanced, lasting insights. Here we summarize select studies demonstrating the construction, logic, and effectiveness of these factors.

- **Innovation as a Long-term Growth Driver** - Quantifying Innovation's Impact on Market Leadership Through Patent Analysis: Innovation drives long-term growth and competitive advantage. For example, in "New Technology Sectoral Disruptions", Caskurlu, Hoberg, and Phillips (2023) analyzed patent portfolios using their Technology Sectoral Disruptions (TSD) metric, revealing its ability to predict positive abnormal stock returns over three years.
- **Networks and Peer Influence on Firm Performance** - Harnessing Network Connections for Competitive Advantage in Asset Pricing: Networks and peer connections influence a firm's competitive edge and market position. In "The effect of Innovation Similarity on Asset Prices: Evidence from Patents' Big Data", Bekkerman, Fich, and Khimich (2021) used patent data to measure technological connectedness, identifying high-performing firms embedded in strategic networks. Their findings demonstrated significant alpha generation, emphasizing the value of advanced network analysis.
- **Governance Quality and Firm Stability** - Governance quality impacts performance and risk management. In "Does Compensation Matter? Evidence from CD&A Disclosures", Martin, Xu, and Zhou, 2021, showed that stable governance practices, reflected in consistent CD&A language, predicted stock outperformance. This highlights how DEEP Factors can extract actionable insights from governance-related disclosures.

The table below summarizes insights derived from rigorous academic research, demonstrating the capabilities and potential of DEEP Factors across key investment domains. Each category reflects findings validated through peer-reviewed studies and credible working papers, providing strong evidence for the value of leveraging qualitative and intangible company characteristics in investment strategies. For readers interested in the specific studies underpinning these insights, Entis can provide detailed references upon request.

Category	Summary of Insights	Reported Alpha Evidence in academic papers
Peer and Industry Similarities	Highlights a company's position within its competitive landscape by analyzing similarities in business models, market focus, and risk profiles.	70–97 bps monthly
Governance Quality	Assesses management stability, compensation policies, and governance practices to evaluate operational resilience and risk management.	30–55 bps monthly
Company Relationships	Identifies strategic alliances, partnerships, and network connections that influence competitive advantage and market positioning.	70–90 bps monthly
Qualitative Changes Over Time	Tracks shifts in company disclosures, governance practices, or business strategies to uncover meaningful trends signaling evolving strengths or vulnerabilities.	48–70 bps monthly
Corporate Communication and Discrepancies	Examines gaps between qualitative statements and quantitative financial metrics to detect potential misalignments or hidden opportunities.	48–55 bps monthly
Technology and Innovation	Measures a company's focus on emerging or sustainable technologies and its ability to innovate, indicating future growth potential and industry leadership.	20–80 bps monthly
Corporate Culture	Analyzes employee reviews and internal dynamics to evaluate teamwork, leadership, and adaptability as key drivers of long-term performance.	30–76 bps monthly

Practical Implementation in Real-World Investment Strategies

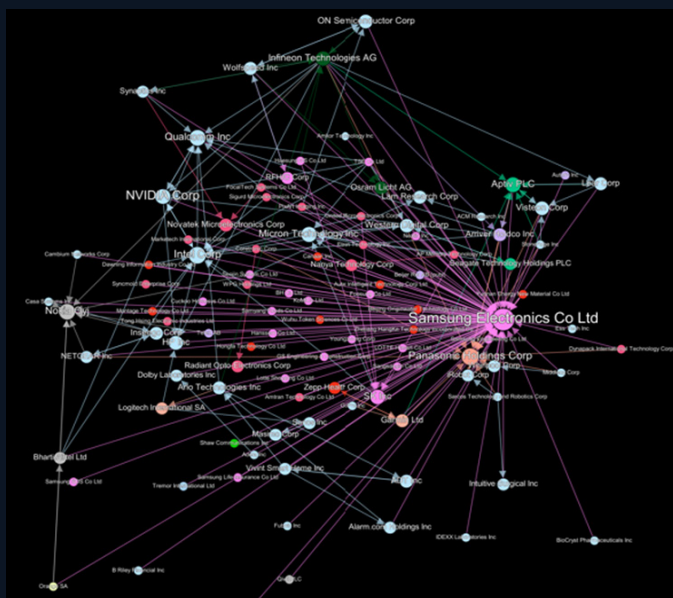
The value of DEEP Factors is evidenced not only by academic research and backtesting results but also by their potential to address critical gaps in investment analysis. These factors demonstrate stable performance over extended periods, providing a strong foundation for systematic investment strategies. Additionally, DEEP Factors offer structured, actionable metrics for previously hard-to-quantify domains such as governance, innovation, and industry dynamics, making them invaluable for modern portfolio management. Below is an overview of key DEEP Factors developed by Entis and their practical applications, illustrating how they enable investors to leverage fundamental insights in systematic and scalable ways.

DEEP Factors applications by Entis

DEEP Factors offer a significant advantage in active quantitative strategies by uncovering nuanced insights into company fundamentals and industry dynamics. Leveraging over 20 years of historical data, Entis has developed a comprehensive suite of DEEP Factors with global coverage to support more informed investment decision-making. Key factors include:

- **Corporate Digital Innovation Focus:** Analyzed by evaluating the emphasis companies place on digital innovation within their annual filings. This factor helps identify firms that are prioritizing technological advancements and integrating them into their business models.
- **Corporate Culture:** Assessed through extensive analysis of employee reviews, focusing on dimensions such as teamwork, leadership, and innovation. This factor provides insights into the internal dynamics of companies, which can be indicative of long-term success and employee satisfaction.
- **Competitiveness:** Constructed by creating a network of competitors based on mentions within annual reports. This unique approach allows for a nuanced understanding of a company's position within its industry and its direct competitors. Metrics such as network centrality or PageRank can quantify a company's influence and competitiveness within this network, offering actionable insights for investors.

Figure 1: This network visualizes the competitive landscape centered around Samsung Electronics. Each node represents an individual company, and the arrows indicate competition links based on one company naming another as a competitor in its annual report. The diagram shows only a small subset of a dynamic network of over 10,000 companies, focusing on Samsung's immediate neighborhood. The network evolves daily with the inclusion of new annual reports, offering continuously updated insights into industry dynamics.



- **Emerging Technology Exposure:** Measured by the proportion of a company's patent portfolio that consists of new and trending technologies. This factor is crucial for identifying firms that are not just current leaders but also poised to shape future industry standards.
- **UN SDG exposure:** This factor assesses the alignment of a company's activities with the United Nations Sustainable Development Goals (SDGs) through two distinct datasets. The first evaluates the proportion of a company's revenue derived from predefined product categories directly linked to specific SDGs, offering a clear view of how business operations contribute to sustainability objectives. The second focuses on patent portfolios, measuring their alignment with a granular framework of technologies deemed critical for achieving SDG targets. Together, these datasets provide a comprehensive understanding of how companies support global sustainability goals through both current operations and forward-looking innovation strategies.
- **Industry Peer momentum:** redefining the concept of industry momentum. Traditionally based on market capitalization or price movements within an industry, momentum can now be understood through a more nuanced lens—examining the qualitative similarities in business models, product development strategies, and market expansion activities among peer companies. This approach allows investors to predict price movements based on much more accurate peer group detection.

These DEEP Factors are designed for integration into various investment strategies, providing asset managers and investors with advanced tools to enhance quantitative methods. By leveraging these factors, professionals can better understand the qualitative drivers of company and market performance, positioning their portfolios to benefit from long-term trends and shifts.

Backtesting Results: Entis Industry Peer Momentum Factor

The figure below illustrates the performance of the Entis Industry Momentum factor, which differs from traditional industry momentum by leveraging industry groups identified through business description similarities extracted from company reports, rather than relying on standard industry classification. This approach captures peers that are not identified through conventional classifications and more effectively uncovers meaningful linkages across companies.

Applied to a U.S. equity universe of approximately 2,500 companies (equal-weighted), the Entis Industry Momentum factor demonstrates robust long-term performance. Backtesting results indicate cumulative growth exceeding 200% over the period 2001 to 2024, almost 5% in annualized terms. Moreover, the factor exhibits smaller and less extreme drawdowns compared to other trend-following strategies, making it a more stable option in dynamic market environments.

Quantitative analysis using the Fama-French three-factor model reveals a monthly alpha of 88 basis points (t-statistic = 2.78), providing strong evidence of both statistical and economic significance. Additionally, internal research highlights the factor's positive exposure to the value factor, which is in contrast to most momentum-based factors that typically exhibit negative value exposure. This unique characteristic enhances its diversification potential within the trend-following space.

The cumulative performance graph further highlights the factor's ability to deliver consistent performance as a systematic alpha-generating investment strategy.



Figure 2: Cumulative Long-Short Performance of the Entis Industry Peer Momentum Factor (2001–2024): Demonstrating stable returns exceeding 200% over the backtest period.

Transforming Investment Strategies with DEEP Factors

DEEP Factors provide a transformative approach to investment strategies by incorporating qualitative and hard-to-measure data. This section explores their application across active strategies, passive vehicles, and thematic investing.

Enhancing Active Quantitative Strategies with DEEP Factors

Active quantitative strategies traditionally rely on structured financial metrics. DEEP Factors enhance these approaches by incorporating unstructured and qualitative data, such as ESG reports, patent filings, and management communications. DEEP Factors redefine traditional metrics like industry momentum and quality while introducing new dimensions tied to governance and innovation. They also offer uncorrelated sources of return, delivering diversification beyond conventional models. By identifying signals for growth and competitive strength, DEEP Factors expand return potential and support robust, data-driven alpha generation.

Expanding Indices and ETFs with DEEP Factors

The growth of ETFs, particularly smart beta ETFs, has reshaped passive investing. Smart beta products aim to exploit inefficiencies in traditional factor investing, and DEEP Factors represent the next step in this evolution. Incorporating DEEP Factors into passive vehicles introduces advanced insights into underexplored areas such as governance and management quality. These factors address inefficiencies that conventional metrics miss, enabling passive vehicles to achieve superior risk-adjusted returns.

Enhancing Thematic Investing Through DEEP Factors

Thematic ETFs have traditionally grouped companies by industry or revenue-based classifications. While effective, these methods are inherently limited. DEEP Factors revolutionize thematic investing by uncovering nuanced interconnections and networks among companies, identifying themes that standard metrics overlook.

Through techniques like network analysis, DEEP Factors reveal clusters of firms aligned by innovative practices, corporate culture, or governance standards tied to sustainability. This dynamic approach transcends traditional classifications and captures emerging themes like digital transformation and sustainable innovation with greater precision and agility.

By expanding the scope of thematic indices, DEEP Factors provide investors with targeted exposures to long-term megatrends and tactical opportunities. The flexibility and granularity offered by these factors enable a wealth of new strategies.

This innovation not only enhances portfolio construction for investors but positions asset managers at the forefront of passive investing. DEEP Factors pave the way for a richer and more impactful future in thematic investing.

03 Technology and Implementation

Advancements in artificial intelligence, particularly through Large Language Models (LLMs), have transformed how insights are extracted from unstructured data. These technologies enable the conversion of qualitative, textual information into structured, actionable insights, forming the foundation for constructing DEEP Factors.

LLMs convert unstructured text into meaningful numerical representations, capturing qualitative attributes like innovation strength or governance quality. By analyzing descriptions in company reports or patent filings, these models systematically score or rank companies based on features relevant to investors, enabling nuanced comparisons across complex qualitative dimensions. Effectively utilizing AI for DEEP Factors requires a sophisticated data infrastructure capable of managing large volumes of text data from diverse sources such as annual reports, patents, and regulatory filings. This involves multiple stages of processing, including:

- **Document Linking and Extraction:** Ensuring precise document-to-company connections using advanced matching algorithms, and extracting relevant text, tables, and figures to provide comprehensive input data.
- **Data Preparation:** Translating and preprocessing information into formats that can be efficiently analyzed, ensuring that only meaningful content is passed to the AI.

These preprocessing steps ensure the reliability and relevance of the data analyzed, laying a strong foundation for accurate insights.

The next step involves crafting targeted inputs for LLMs. This phase focuses on determining what information to present to the AI, rather than merely guiding its analysis. Well-constructed prompts ensure that outputs are relevant, specific, and actionable, minimizing the risk of irrelevant or erroneous results.

AI-generated insights must undergo rigorous validation to maintain reliability. This includes: Benchmarking against trusted reference data to verify accuracy and involving human experts in the review process to mitigate risks such as model hallucinations or biases.

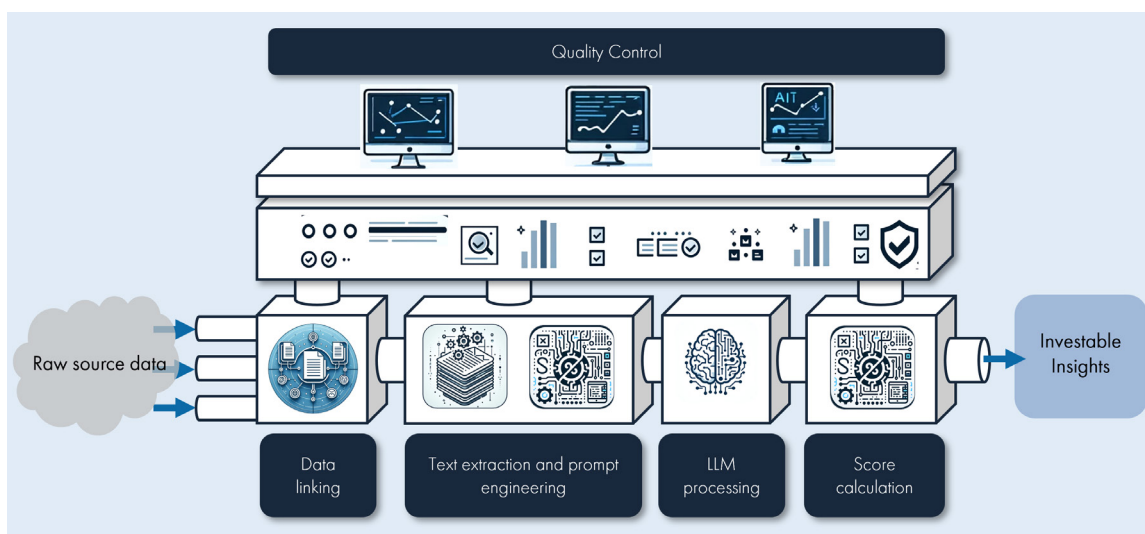


Figure 3: From raw source data to investable insights - Data processing infrastructure for utilizing AI models such as LLMs.

Conclusion

Leveraging AI, especially LLMs, for developing and implementing DEEP Factors requires a robust, technically advanced infrastructure for processing text data. From precise document extraction to effective input design and thorough validation, this approach enables asset managers to uncover long-term, fundamental investment opportunities that conventional methods often miss.

A critical strategic decision in this process is whether to build proprietary AI systems tailored to specific needs or adopt pre-made solutions. This choice involves weighing factors such as cost, competitive advantage, and alignment with the firm's broader investment goals. Firms that invest in the right infrastructure and decision-making frameworks will be well-positioned to capitalize on the transformative potential of DEEP Factors, gaining a significant edge in today's increasingly data-driven investment landscape.

Conclusion and Call to Action

Traditional financial metrics tell only part of the story. A wealth of qualitative and intangible factors—such as innovation, governance, and strategic direction—remain beyond their scope. These often-overlooked dimensions present a compelling opportunity for investors willing to explore deeper insights.

DEEP Factors address this gap by systematically uncovering hard-to-measure but critical drivers of long-term performance. By extracting and quantifying complex qualitative information, DEEP Factors provide a robust toolset for identifying underappreciated opportunities and gaining a strategic edge in an increasingly competitive market.

As AI tools and data-processing capabilities continue to evolve, the adoption of advanced factors like these will likely grow. However, the challenge of building and maintaining the infrastructure required to process clean, reliable information remains substantial. For many, the resources and expertise needed for such initiatives may be out of reach.

Investors and product developers looking to stay ahead must act now. Leveraging existing platforms, such as those supporting DEEP Factors, enables access to these insights without the need for costly, resource-intensive development. Partnering with established solutions provides a practical and scalable path to unlock the potential of advanced qualitative analysis, ensuring a competitive edge in the data-driven future of investing.

Contact Information

Entis b.v.

Building B
Papendorpseweg 97
3528 BJ Utrecht
The Netherlands

info@entis.ai

+31 (0)30 7991140

Visit the website for more information: www.entis.ai

