WEP Warwick Evans Publishing

International Journal of Global Economics and Management

ISSN: 3005-9690 (Print), ISSN: 3005-8090 (Online)
Volume 10, Number 10, Year 2026
DOI: https://doi.org/10.62051/ijgem.v10n10.02



# The Impact of Natural Language Processing Technology on Enterprise New-Quality Productive Forces

Lina Guo *

School of Economics and Management, Huanghe Jiaotong University, Jiaozuo, 454000, China
*Corresponding Author: 2004080078@zjtu.edu.cn

## ABSTRACT

Natural language processing (NLP) technology is increasingly integrated into corporations, with important implications for enterprise productivity. Using Chinese A-share listed companies from 2015 to 2023, this study constructs an index of NLP technology adoption based on the annual report MD&A sections and examines its impact on enterprise new-quality productive forces and the underlying mechanisms. The results show that NLP technology significantly enhances enterprise new-quality productive forces. Mechanism tests indicate that NLP technology enhances new-quality productive forces by increasing innovation investment, reducing operating expense ratios, and improving the development and utilization of data as a factor of production.

## KEYWORDS

Natural Language Processing Technology; New-Quality Productive Forces; Innovation Investment; Operational Efficiency; Data as a Factor of Production

## 1. INTRODUCTION

With the rapid development of natural language processing (NLP) has evolved from rule-based and statistical methods to deep learning, pre-trained language models, and large language models. Che et al. (2023) [1] note that large language models are driving a paradigm shift in NLP. As these technologies advance, NLP applications have expanded from language recognition and text analysis to knowledge extraction, content generation, and decision support, facilitating their integration into corporate R&D, operations, and management.

At the policy level, China has increasingly emphasized the development and application of NLP technology. In 2025, the Ministry of Education, the National Language Commission, and the Cyberspace Administration of China jointly called for strengthening innovation in NLP and large language models and promoting the application of related research outcomes. The State Council's "Opinions on Deeply Implementing the 'AI+' Initiative" further emphasized the integration of AI across economic and social sectors to accelerate the development of new-quality productive forces. Against this background, the productivity effects of NLP technology warrant further investigation.

New-quality productive forces are primarily driven by scientific and technological innovation and are closely associated with technological progress, factor allocation, and production efficiency. Zhou and Xu (2023) [2] examined their connotations, characteristics, and key development areas. Using Chinese A-share listed companies, He et al. (2025) [3] found that AI applications can enhance enterprise new-quality productive forces, providing empirical support for examining the microeconomic effects of intelligent technologies.

Content from this work may be used under the terms of CC BY-NC 4.0 licence (https://creativecommons.org/licenses/by-nc/4.0/).
Published by Warwick Evans Publishing.

However, existing NLP research has largely focused on technological development, application scenarios, and text-processing capabilities, with limited attention to the productivity effects of enterprise-level NLP adoption. Meanwhile, studies on enterprise new-quality productive forces have mainly examined broad technological developments, such as AI and digital transformation, rather than specific technologies such as NLP. This leaves the enterprise-level effects and transmission mechanisms of NLP technology insufficiently understood.

To address this gap, this study uses Chinese A-share listed companies from 2015 to 2023 and constructs an NLP technology adoption index based on the Management's Discussion and Analysis (MD&A) sections of annual reports. It examines the impact of NLP technology on enterprise new-quality productive forces and investigates three potential mechanisms: innovation investment, operational efficiency, and data factor utilization. By focusing on NLP as a specific AI technology, this study provides micro-level empirical evidence on how AI technologies contribute to enterprise new-quality productive forces.

# 2. THEORETICAL ANALYSIS AND RESEARCH HYPOTHESES

## 2.1. Natural Language Processing Technologies and New-Quality Productive Forces in Enterprises

Scientific and technological innovation is a key driver of new-quality productive forces, which depend not only on technological progress but also on improvements in factor allocation and production and management practices. He et al. (2025) [3] found that AI applications can significantly enhance enterprise new-quality productive forces. Unlike general digital technologies, natural language processing (NLP) technology primarily processes unstructured information such as text. By improving the identification, comprehension, and utilization of such information, NLP technology can reduce the costs of acquiring and processing market, technological, and operational information. Its integration into R&D, production, and operational management can further facilitate innovation, improve operational efficiency, and optimize factor allocation, thereby enhancing enterprise new-quality productive forces. Accordingly, the following hypothesis is proposed:

H1: NLP technology can significantly promote the enhancement of enterprises' new-quality productive forces.

## 2.2. Innovative Investment Mechanisms

Innovation depends on the acquisition, integration, and application of knowledge, while high information-processing costs may constrain corporate R&D activities. NLP technology facilitates the retrieval and analysis of unstructured information, such as patents, technical reports, and R&D materials, thereby reducing knowledge acquisition costs and supporting corporate innovation. Ji et al. (2023) [4] found that digital transformation can increase corporate R&D investment and improve R&D efficiency. By improving the allocation of R&D resources and supporting the development of new technologies, products, and production methods, NLP technology can enhance enterprise new-quality productive forces through greater innovation investment. Accordingly:

H2: NLP technology can promote the enhancement of enterprises' new-quality productive forces by increasing corporate innovation investment.

## 2.3. Operational Efficiency Mechanism

Business operations involve extensive information collection, transmission, and processing, while information asymmetry and processing costs may reduce internal coordination efficiency. Ang et al. (2000) [5] examined corporate expenses from the perspective of agency costs, providing a theoretical

basis for analyzing operational efficiency. NLP technology can be applied to contract processing, business analysis, customer service, and internal management, improving information-processing efficiency while reducing repetitive labor and management costs. Consequently, lower operating expense ratios and more efficient resource allocation can contribute to enterprise new-quality productive forces. Accordingly:

H3: NLP technology can promote the enhancement of new-quality productive forces by reducing corporate operating expense ratios and improving operational efficiency.

### 2.4. Data Factor Empowerment Mechanism

Enterprises generate large amounts of unstructured data, including customer feedback, business records, and R&D materials, whose value depends on their effective development and utilization. Jones and Tonetti (2020) [6] highlight the non-rivalry and reusability of data, which support its sustained economic utilization. NLP technology can identify, extract, and analyze textual information, transforming dispersed information into usable data resources and improving firms' ability to develop and utilize data. The integration of these data resources with technology, capital, and labor can further improve resource allocation efficiency and enhance enterprise new-quality productive forces. Accordingly:

H4: NLP technology can promote the enhancement of enterprises' new-quality productive forces by strengthening the enabling role of data as a production factor.

## 3. RESEARCH DESIGN

### 3.1. Sample Selection and Data Sources

This study uses Chinese A-share listed companies from 2015 to 2023 as the initial sample. We exclude (1) financial firms and (2) firms designated as ST or *ST or with missing data on key variables. All continuous variables are winsorized at the 1st and 99th percentiles to mitigate the influence of outliers, yielding 18,738 firm-year observations. Data on natural language processing (NLP) technology are extracted from the Management's Discussion and Analysis (MD&A) sections of annual reports, while other firm-level data are obtained from the CSMAR database.

### 3.2. Variable Descriptions

(1) Dependent Variable: Enterprise New-Quality Productive Forces (NQP). Following Song et al. (2024) [7], we use the New-Quality Productivity Index (NQP) to measure enterprise new-quality productive forces. The index incorporates dimensions such as innovation-driven development, digital transformation, and production efficiency. After standardizing the indicators, the entropy weighting method is used to determine their weights and construct the composite index. A higher NQP value indicates a higher level of enterprise new-quality productive forces. For robustness, we replace NQP with the alternative index (NPRO) developed by Zhang et al. (2025) [8].

(2) Explanatory Variable: Natural Language Processing Technology (lnNLP). We construct a dictionary-based measure of corporate NLP technology adoption using the MD&A sections of annual reports. Drawing on the technological evolution of NLP from deep learning to pre-trained and large language models (Qiu et al., 2020) [9], the dictionary covers six categories: basic NLP technologies, NLP tasks and applications, language models, large language model technologies, corporate applications, and representative models. Representative terms include natural language processing, semantic analysis, sentiment analysis, machine translation, Transformer, large language models, prompt engineering, retrieval-augmented generation, agents, BERT, GPT, and T5. After cleaning the

MD&A texts, we count the annual frequency of NLP-related terms and construct the NLP technology adoption index (lnNLP) as:

$$\text{lnNLP} = \ln(1 + \text{NLP_freq})$$

Where NLP_freq denotes the frequency of NLP-related terms in a firm's MD&A text. For robustness, we construct an alternative measure, NLP_ratio, defined as the frequency of NLP-related terms relative to the total number of words in the annual report:

$$\text{NLP_ratio} = \text{NLPWord Frequency} / \text{Total Number of Words}$$

This alternative measure is used to test whether the results are sensitive to the measurement of the core explanatory variable.

(3) Mechanism Variables. We examine three potential mechanisms: innovation investment, operational efficiency, and data factor utilization.

Innovation Investment (RD). Following Xie and Fang (2011) [10], RD is measured as R&D expenditure divided by operating revenue. A higher value indicates greater R&D intensity.

Operational Efficiency (OER). Following Li and Yu (2015) [11], OER is measured as the sum of administrative and selling expenses divided by operating revenue. A higher OER indicates higher operating expenses and thus lower operational efficiency.

Data Factor Utilization (lndata). Following Xu et al. (2021) [12], we construct a dictionary of data-related keywords and use their frequency in annual reports to measure firms' development and utilization of data as a factor of production. A higher lndata value indicates a higher level of data factor utilization.

(4) Control variables. This study controls for firm characteristics such as firm size (Size), debt-to-asset ratio (Lev), time since listing (ListAge), profitability (ROA), growth (Growth), cash flow level (CashFlow), concentration of ownership (Top1), ownership structure (SOE), board size (Board), and the combination of executive and supervisory roles (Dual).

### 3.3. Model Setting

To examine the impact of NLP technology on enterprises' new-quality productive forces and the underlying mechanisms, we employed a multiple linear regression model, controlling for fixed effects of year (Year) and firm (Firm), and adjusted for firm-level clustering among listed companies during the regression analysis.

## 4. EMPIRICAL ANALYSIS

### 4.1. Descriptive Statistics

Table 1 presents the descriptive statistics. The mean value of NQP is 0.010, with a standard deviation of 0.010, indicating variation in enterprise new-quality productive forces across the sample. The mean value of lnNLP is 0.030, with a standard deviation of 0.170 and a range from 0 to 1.390, suggesting substantial variation in NLP technology adoption across firms.

**Table 1.** Descriptive Statistics

| Variable | N | Mean | SD | Min | Max |
|---|---|---|---|---|---|
| NQP | 18738 | 0.010 | 0.010 | 0.000 | 0.030 |
| lnNLP | 18738 | 0.030 | 0.170 | 0.000 | 1.390 |
| Size | 18738 | 22.230 | 1.120 | 20.230 | 25.640 |
| Lev | 18738 | 0.400 | 0.180 | 0.060 | 0.820 |
| ListAge | 18738 | 2.080 | 0.760 | 0.690 | 3.330 |
| ROA | 18738 | 0.040 | 0.060 | -0.190 | 0.200 |
| Growth | 18738 | 0.150 | 0.310 | -0.490 | 1.670 |
| CashFlow | 18738 | 0.060 | 0.060 | -0.110 | 0.240 |
| Top1 | 18738 | 0.330 | 0.140 | 0.080 | 0.730 |
| SOE | 18738 | 0.270 | 0.440 | 0.000 | 1.000 |
| Board | 18738 | 2.100 | 0.190 | 1.610 | 2.560 |
| Dual | 18738 | 0.330 | 0.470 | 0.000 | 1.000 |

## 4.2. Regression Against a Benchmark and Endogeneity Tests

Table 2 reports the baseline regression and endogeneity test results. The estimated coefficients of lnNLP in Columns (1)–(3) are 0.003, 0.004, and 0.001, respectively, all significant at the 1% level. The consistently positive coefficients indicate that NLP technology is positively associated with enterprise new-quality productive forces, supporting H1.

To address potential endogeneity, we further employ an instrumental variable (IV) approach. Following the external-shock and initial-exposure framework of Goldsmith-Pinkham et al. (2020) [13], the instrument is constructed as the interaction between a city's emphasis on digital economy policies and a firm's initial digital infrastructure. The first-stage result in Column (4) shows that the IV is significantly associated with lnNLP at the 1% level. In the second stage, the coefficient of lnNLP is 0.012 and significant at the 5% level, indicating that the positive effect remains after accounting for potential endogeneity.

**Table 2.** Regression Against a Benchmark and Endogeneity Tests

| | (1) | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|
| | NQP | NPRO | NQP | lnNLP | NQP |
| lnNLP | 0.003*** | 0.004*** | 0.001*** | | 0.012** |
| | (9.877) | (13.410) | (3.662) | | (2.452) |
| IV | | | | -2.104*** | |
| | | | | (-3.566) | |
| Size | | -0.001*** | -0.001*** | 0.024*** | -0.001*** |
| | | (-18.457) | (-3.295) | (3.854) | (-3.840) |
| Lev | | 0.001*** | -0.000 | -0.019 | -0.000 |
| | | (4.063) | (-0.648) | (-0.988) | (-0.278) |
| ListAge | | 0.000 | -0.000 | -0.004 | -0.000 |
| | | (1.039) | (-1.510) | (-0.457) | (-1.222) |
| ROA | | 0.005*** | -0.001 | -0.093** | -0.000 |
| | | (4.365) | (-1.644) | (-2.399) | (-0.379) |
| Growth | | 0.001*** | 0.000*** | -0.013*** | 0.001*** |
| | | (6.239) | (3.112) | (-2.891) | (3.537) |
| CashFlow | | 0.001 | -0.000 | 0.023 | -0.000 |
| | | (1.098) | (-0.067) | (0.919) | (-0.385) |
| Top1 | | 0.003*** | 0.002** | -0.062** | 0.003** |
| | | (6.716) | (2.040) | (-2.010) | (2.441) |
| SOE | | 0.000*** | -0.000 | -0.016** | -0.000 |
| | | (3.236) | (-1.339) | (-2.384) | (-0.760) |
| Board | | 0.003*** | 0.001*** | -0.038** | 0.002*** |
| | | (11.894) | (3.345) | (-2.234) | (3.642) |
| Dual | | -0.001*** | -0.000 | 0.009* | -0.000 |
| | | (-9.998) | (-0.755) | (1.689) | (-1.364) |
| _cons | 0.006*** | 0.022*** | 0.018*** | -0.372*** | |
| | (114.123) | (17.403) | (4.169) | (-2.806) | |
| N | 18738 | 18738 | 17587 | 16664 | 16439 |
| r2_a | 0.7606 | 0.9005 | 0.7677 | 0.7645 | 0.7496 |
| Firm/Year | Yes | Yes | Yes | Yes | Yes |

Note: *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively. The values in parentheses are t-statistics based on standard errors clustered at the firm level. The same convention applies to subsequent tables.

### 4.3. Robustness Tests

Table 3 presents the robustness test results. In Column (1), the core explanatory variable is replaced by NLP_ratio, yielding a positive coefficient of 7.7746 that is significant at the 1% level. In Column (2), NQP is replaced by the alternative dependent variable NPRO, and the coefficient of lnNLP is 0.0055 and significant at the 1% level. Columns (3)–(5) sequentially exclude observations from 2015, 2018, and 2020; the coefficients of lnNLP remain positive and statistically significant. Overall, the results are robust to alternative variable measures and sample adjustments.

**Table 3.** Robustness Test Results

| | (1) | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|
| | Replace the explanatory variable | Replace the interpreted variable | Excluding 2015 | Excluding 2018 | Excluding 2020 |
| | NQP | NPRO | NQP | NQP | NQP |
| NLP_ratio | 7.7746*** | | | | |
| | (3.185) | | | | |
| lnNLP | | 0.0055*** | 0.0005** | 0.0007*** | 0.0008*** |
| | | (9.481) | (2.563) | (3.366) | (3.313) |
| Controls | Yes | Yes | Yes | Yes | Yes |
| Firm/Year | Yes | Yes | Yes | Yes | Yes |
| _cons | 0.0174*** | 0.0019 | 0.0171*** | 0.0180*** | 0.0175*** |
| | (4.132) | (0.433) | (4.075) | (4.209) | (4.069) |
| N | 18738 | 18738 | 17587 | 16664 | 16439 |
| r2_a | 0.7606 | 0.9005 | 0.7677 | 0.7645 | 0.7496 |

## 4.4. Mechanism Testing

Table 4 examines three mechanisms through which NLP technology may affect enterprise new-quality productive forces: innovation investment, operational efficiency, and data factor utilization.

**Table 4.** Mechanism Test Results

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| | RD | NQP | OER | NQP | lndata | NQP |
| lnNLP | 0.6377*** | 0.0008*** | -0.0106* | 0.0008*** | 0.0230** | 0.0008*** |
| | (3.072) | (3.446) | (-1.814) | (3.488) | (2.176) | (3.629) |
| RD | | 0.0001*** | | | | |
| | | (5.804) | | | | |
| OER | | | | -0.0030** | | |
| | | | | (-2.136) | | |
| lndata | | | | | | 0.0004* |
| | | | | | | (1.840) |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes |
| Firm/Year | Yes | Yes | Yes | Yes | Yes | Yes |
| _cons | 8.5634** | 0.0168*** | 0.6004*** | 0.0193*** | 2.7582*** | 0.0165*** |
| | (2.489) | (4.008) | (8.587) | (4.506) | (12.700) | (3.865) |
| N | 18738 | 18738 | 18738 | 18738 | 18738 | 18738 |
| ar2 | 0.8552 | 0.7613 | 0.8502 | 0.7611 | 0.8093 | 0.7608 |

Columns (1) and (2) examine the innovation investment mechanism. The coefficient of lnNLP on RD is 0.6377 and significant at the 1% level. After RD is included in the NQP regression, its coefficient is 0.0001 and significant at the 1% level, while lnNLP remains significantly positive. These results are consistent with the mechanism proposed in H2. Columns (3) and (4) examine operational efficiency. The coefficient of lnNLP on OER is −0.0106 and significant at the 10% level, indicating that NLP technology is associated with a lower operating expense ratio. In the NQP regression, OER has a coefficient of −0.0030 and is significant at the 5% level, while lnNLP remains significantly positive. These findings are consistent with H3. Columns (5) and (6) examine data factor utilization. The coefficient of lnNLP on lndata is 0.0230 and significant at the 5% level. After lndata is included in the NQP regression, its coefficient is 0.0004 and significant at the 10% level, while lnNLP remains significantly positive. These results are consistent with H4, suggesting that improved development

and utilization of data constitute another potential channel through which NLP technology enhances enterprise new-quality productive forces.

### 4.5. Heterogeneity Test

(1) Market Heterogeneity

Table 5 examines heterogeneity in market competition and regional marketization. The coefficient of lnNLP is insignificant in the high-competition group but positive and significant at the 1% level in the low-competition group, suggesting that the effect of NLP technology is more evident under lower market competition. Regarding regional marketization, lnNLP is positive and significant at the 1% level in the high-marketization group but insignificant in the low-marketization group. These results suggest that market conditions may shape the productivity effects of NLP technology.

**Table 5.** Results on Market Heterogeneity

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| | Market Competition | | Degree of Marketization | |
| | High | Low | High | Low |
| | NQP | NQP | NQP | NQP |
| lnNLP | 0.001 | 0.001*** | 0.001*** | 0.001 |
| | (1.378) | (3.042) | (3.502) | (1.306) |
| Controls | Yes | Yes | Yes | Yes |
| Firm/Year | Yes | Yes | Yes | Yes |
| _cons | 0.003 | 0.020*** | 0.018*** | 0.006 |
| | (0.392) | (3.914) | (3.929) | (0.541) |
| N | 5858 | 12543 | 16316 | 2320 |
| r2_a | 0.771 | 0.761 | 0.764 | 0.750 |

(2) Heterogeneity in Industrial Characteristics

Table 6 examines heterogeneity across industry sectors and factor intensity, following the classification of Yin et al. (2018) [14]. The coefficient of lnNLP is insignificant in the primary sector but positive and significant in the secondary and tertiary sectors. Similarly, the coefficient is insignificant in labor-intensive industries but positive and significant in technology and capital-intensive industries. These findings suggest that the effect of NLP technology is more evident in industries with greater technological and capital intensity.

**Table 6.** Results on Industry Heterogeneity

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| | Primary Sector | Secondary Sector | Tertiary Sector | Labor-intensive | Technology-intensive | Capital-intensive |
| | NQP | NQP | NQP | NQP | NQP | NQP |
| lnNLP | 0.020 | 0.001*** | 0.001* | 0.001 | 0.001*** | 0.001** |
| | (1.464) | (2.725) | (1.909) | (0.794) | (3.441) | (2.562) |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes |
| Firm/Year | Yes | Yes | Yes | Yes | Yes | Yes |
| _cons | -0.028 | 0.024*** | 0.003 | 0.005 | 0.015*** | 0.035*** |
| | (-0.894) | (4.722) | (0.421) | (0.614) | (2.653) | (3.430) |
| N | 200 | 14014 | 4473 | 4221 | 9763 | 4728 |
| ar2 | 0.737 | 0.767 | 0.756 | 0.786 | 0.769 | 0.770 |

(3) Firm Heterogeneity

Columns (1) and (2) of Table 7 examine differences effects in firm size. First, the coefficient of lnNLP is positive and significant for large firms but insignificant for small and medium-sized enterprises (SMEs), suggesting that larger firms may be better positioned to absorb the costs of technology deployment and organizational adjustment. Second, following Wu et al. (2021) [15], we construct a digitalization index from annual-report keywords and divide firms into high- and low-digitalization groups. The coefficient of lnNLP is positive and significant in the high-digitalization group but insignificant in the low-digitalization group, suggesting that stronger digital infrastructure facilitates the integration of NLP technology into business processes. Third, human capital is measured by the proportion of R&D personnel in the total workforce. The coefficient of lnNLP is positive and significant in the high-human-capital group but insignificant in the low-human-capital group, suggesting that greater R&D human capital may facilitate the effective application of NLP technology.

**Table 7.** Results on Firm Heterogeneity

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| | Large Companies | Small and Medium-Sized Enterprises | High Digitalization | Low Digitalization | High Human Capital | Low Human Capital |
| | NQP | NQP | NQP | NQP | NQP | NQP |
| lnNLP | 0.001*** | 0.000 | 0.001*** | 0.001 | 0.001** | 0.001 |
| | (3.566) | (0.213) | (3.533) | (0.737) | (2.357) | (1.493) |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes |
| Firm/Year | Yes | Yes | Yes | Yes | Yes | Yes |
| _cons | 0.019*** | -0.022 | 0.020*** | 0.013* | 0.019*** | 0.013** |
| | (4.240) | (-0.875) | (3.471) | (1.882) | (3.124) | (2.060) |
| N | 17267 | 1289 | 8676 | 9438 | 5948 | 10944 |
| ar2 | 0.764 | 0.767 | 0.778 | 0.769 | 0.833 | 0.795 |

# 5. RESEARCH FINDINGS AND POLICY RECOMMENDATIONS

## 5.1. Research Findings

Using Chinese A-share listed companies from 2015 to 2023, this study constructs an NLP technology adoption index based on the Management's Discussion and Analysis (MD&A) sections of annual reports and examines its impact on enterprise new-quality productive forces. The main findings are as follows.

First, NLP technology significantly enhances enterprise new-quality productive forces. This finding remains robust after accounting for endogeneity, employing alternative measures of the core variables, and adjusting the sample. Second, NLP technology enhances enterprise new-quality productive forces through three main channels: increasing innovation investment, improving operational efficiency by reducing operating expense ratios, and strengthening the development and utilization of data as a factor of production. These findings highlight the importance of integrating NLP technology with corporate R&D, operations, and data utilization. Third, the effect of NLP technology varies across market, industry, and firm characteristics. The positive effect is more pronounced in environments with lower market competition and higher levels of marketization, in the secondary and tertiary sectors, in technology- and capital-intensive industries, and among larger firms with stronger digital infrastructure and higher levels of human capital.

### 5.2. Policy Recommendations

Based on these findings, efforts to promote the effective application of NLP technology should focus on the government, industry, and enterprise levels. At the government level, policymakers should improve the market environment and market-based factor allocation, strengthen digital infrastructure and public technology services, and reduce barriers to enterprise technology adoption. At the industry level, NLP applications should be promoted according to industry-specific characteristics and needs, particularly in R&D, production, services, and management, while encouraging the sharing of common technologies and best practices. At the enterprise level, firms should integrate NLP technology into R&D, operations, and data management, increase innovation investment, improve operational efficiency and data utilization, and strengthen digital infrastructure and human capital. These measures can facilitate the effective transformation of NLP technology into improvements in enterprise new-quality productive forces.

## ACKNOWLEDGEMENTS

This research was supported by the 2025 Soft Science Project of Henan Provincial Department of Science and Technology titled "Research on the Impact of Natural Language Intelligent Processing Technology on the New Quality Productivity of Enterprises in Henan Province" (Grant No. 252400410058).

## REFERENCES

[1] Che, W. X., Dou, Z. C., Feng, Y. S., et al. (2023). Natural Language Processing in the Era of Large Models: Challenges, Opportunities, and Development. Science in China: Information Sciences, 53(9), 1645–1687. (In Chinese)

[2] Zhou, W., & Xu, L. Y. (2023). On New-Quality Productive Forces: Connotations, Characteristics, and Key Focus Areas. Reform, (10), 1–13. (In Chinese)

[3] He, X. J., Wang, Q. P., & Yang, M. J. (2025). A Study on the Impact of Artificial Intelligence on New-Quality Productive Forces: Empirical Evidence from Chinese A-Share Listed Companies. Technology and Economy, 44(2), 97–114. (In Chinese)

[4] Ji, Y. Y., Zhou, X., & Zhang, Q. (2023). Digital Transformation and Corporate Innovation: An Analysis from the Perspective of R&D Investment and R&D Efficiency. Financial Research, (4), 111–129. (In Chinese)

[5] Ang, J. S., Cole, R. A., & Lin, J. W. (2000). Agency Costs and Ownership Structure. The Journal of Finance, 55(1), 81–106.

[6] Jones, C. I., & Tonetti, C. (2020). Nonrivalry and the Economics of Data. American Economic Review, 110(9), 2819–2858.

[7] Song, J., Zhang, J. C., & Pan, Y. (2024). A Study on the Impact of ESG Development on Enterprises' New-Quality Productivity: Empirical Evidence from Chinese A-Share Listed Companies. Contemporary Economic Management, 46(6), 1–11. (In Chinese)

[8] Zhang, X. E., Wang, W., & Yu, Y. B. (2025). A Study on the Impact of Digital and Intelligent Transformation on Enterprises' New-Quality Productivity. Studies in Science of Science, 43(5), 943–954. (In Chinese)

[9] Qiu, X., Sun, T., Xu, Y., et al. (2020). Pre-trained Models for Natural Language Processing: A Survey. Science China: Technological Sciences, 63, 1872–1897.

[10] Xie, W. M., & Fang, H. X. (2011). Financial Development, Financing Constraints, and Corporate R&D Investment. Financial Research, (5), 171–183. (In Chinese)

[11] Li, W. G., & Yu, M. G. (2015). Equity Structure and Innovation in Privatized Enterprises. Management World, (4), 112–125. (In Chinese)

[12] Xu, X., Li, K. A. B., & Tian, X. X. (2021). Research Progress on Data as a Factor of Production. Dynamics of Economics, (4), 142–158. (In Chinese)

[13] Goldsmith-Pinkham, P., Sorkin, I., & Swift, H. (2020). Bartik Instruments: What, When, Why, and How. American Economic Review, 110(8), 2586–2624.

[14] Yin, M. Q., Sheng, L., & Li, W. B. (2018). Executive Incentives, Innovation Investment, and Firm Performance: An Empirical Study by Industry from an Endogeneity Perspective. Nankai Management Review, 21(1), 109–117. (In Chinese)

[15] Wu, F., Hu, H. Z., Lin, H. Y., et al. (2021). Corporate Digital Transformation and Capital Market Performance: Empirical Evidence from Stock Liquidity. Management World, 37(7), 130–144. (In Chinese)