Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 3, Issue 7.1 (2025)

Online ISSN 3007-3197
Print ISSN 3007-3189

http://amresearchreview.com/index.php/Journal/about

Emerging Trends of Computational Linguistics with Reference to Natural Language Processing: A Quantitative Study

Qadir Bakhsh*
NCBA&E, Multan, Pakistan
Email: qadirbakhsh1a@gmail.com

Gulnaz Sharif
NCBA&E, Multan, Pakistan precious
Email: gulnaz@gmail.com

Farah Hamid
NCBA&E, Multan, Pakistan
Email: farah.khan141299@gmail.com

Isha Urooj
NCBA&E, Multan, Pakistan
Email: ishatahir9@gmail.com

Syeda Tahira Zainab
NCBA&E, Multan, Pakistan
Email: tahiranaqvi24@gmail.com

Prof. Dr. M. Asim Rajwana
NCBA&E, Multan, Pakistan
Email: masimrajwana@ncbae.edu.pk

Prof. Dr. Muhammad Khalid Mehmood Sajid
NCBA&E, Multan, Pakistan
Email: interstudies2000@gmail.com

ABSTRACT

This study aims to examine the emerging trends in Computational Linguistics (CL) with a specific focus on Natural Language Processing (NLP) using quantitative methodology. A structured questionnaire was disseminated online via Google Forms to collect data from 210 participants comprising diverse population of university students, researchers, faculty and the professionals engaged with NLP and CL across Pakistan. The instrument included demographic variables and Likert-scale items to explore domains such as NLP applications, preferred tools and algorithms, emerging research areas and interdisciplinary integration. The collected data were analyzed using IBM

Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 3, Issue 7.1 (2025)

Online ISSN 3007-3197
Print ISSN 3007-3189

http://amresearchreview.com/index.php/Journal/about

SPSS Statistics (v27). Descriptive statistics displayed that NLP is highly recognized as central to the CL research and education. Inferential analyses traced out statistically prominent differences regarding trend perception and ethical awareness on the basis of experience and professional roles. The findings suggest a shifting academic and professional move towards NLP that is reshaping technical practices as well as redefining educational and ethical frameworks, indicating a marked shift towards low-resource language processing, transformer-based architectures (e.g., BERT, GPT), and ethical NLP. A substantial number of respondents reported increasing reliance on Python-based toolkits (NLTK, spaCy, HuggingFace) that shows a convergence of linguistic inquiry with computational perspective. This paper contributes in the existing body of the scholarly discourse by presenting empirical insights into current trends in Computational Linguistics related to Natural Language Processing, proposing a future-focused research framework encouraging interdisciplinary collaboration, ethical responsibility and support for underrepresented linguistic communities.

Keywords: Computational Linguistics, Natural Language Processing, Emerging Trends

Introduction

Computational Linguistics (CL) and Natural Language Processing (NLP) are rapidly evolving subfields within Artificial Intelligence and modern linguistics. Both disciplines were originally developed to model human language computationally but now they have become part and parcel of transformative technologies— ranging from sentiment analysis and large-scale machine translation to real-time dialogue systems and language generation [1]. As theoretical linguistics has interactions with statistical modeling, deep learning and ethical AI frameworks, the convergence of CL and NLP is a significant paradigm shift.

The expectations of computational models and their linguistic underpinnings are redefined due to increasing growth of NLP applications in the domains of multilingual communication, healthcare, education and governance. Large Language Models (LLMs) like OpenAI's GPT-4 and Google's PaLM-2 have proved remarkable capabilities regarding language understanding and generation [2]. However, they also raise pressing questions related to explainability, bias and the marginalization of low-resource languages [3]. As a result, the boundaries of CL are expanding in terms of a support system for AI as well as a critical lens which questions these systems

It is due to this growing influence that there remains need of empirical research to explore the actual perceptions, practices and preferences among stakeholders engaged at this intersection. Which particular areas of NLP are currently gaining scholarly and industrial interest? Which programming tools, methodologies and theoretical frameworks are most widely utilized? And how does an individual's academic and professional background shape engagement with these technologies?

To explore these questions, a quantitative survey methodology is adopted targeting researchers, educators and developers from diverse disciplines. This study aims to trace out dominant and growing trends, chart the prevailing tools and frameworks, and reveal the shifting research preferences in contemporary CL and NLP. Statistical analysis

Online ISSN 3007-3197
Print ISSN 3007-3189

using SPSS provides empirical support to the findings, allowing for generalizable insights across demographic groups.

This study is significant for several reasons. To begin with , it contributes to the advancement of aligning computational approaches with linguistic theory, particularly in multilingual and ethically sensitive domains. Additionally, it offers data-driven insights for future curriculum design, policy formulation and interdisciplinary initiatives. Finally, by emphasizing both technological advancement and social fairness, it underscores the critical role of inclusive and transparent NLP research in guiding the future of human-machine interaction.

Literature Review

The interface of Computational Linguistics (CL) and Natural Language Processing (NLP) has undergone substantial theoretical and methodological reconfiguration in recent years. Traditional rule-based systems have given way to data-driven approaches powered by deep learning, large-scale corpora, and probabilistic modeling [4]. This shift has not only enhanced machine translation, parsing, and named entity recognition but also intensified debates around interpretability and linguistic adequacy [5, 6].

One salient trend is the rise of **pretrained language models** (PLMs) such as BERT, RoBERTa, GPT-3/4, and PaLM-2 [7-9]. These architectures have outperformed earlier models across diverse NLP benchmarks, significantly influencing both academic research and commercial applications [10, 11]. However, recent work has cautioned against over-reliance on PLMs due to their opacity and embedded biases [12].

Ethical concerns around fairness, transparency, and inclusivity are now central to CL–NLP scholarship. Studies have shown that most LLMs are trained on Anglo-centric, high-resource languages, leading to the systematic exclusion of underrepresented linguistic communities[13]. In response, researchers are advocating for equitable NLP that incorporates indigenous, low-resource, and endangered languages [14, 15].

Tool usage in NLP has also diversified. Open-source libraries such as spaCy, NLTK, AllenNLP, and HuggingFace Transformers have democratized access to cutting-edge algorithms [16, 17]. This accessibility has allowed even early-career researchers to implement complex models with minimal computational infrastructure, spurring innovation globally [18].

Interdisciplinary convergence is another notable development. Computational linguists now frequently collaborate with researchers in **cognitive science** [19], **neuroscience** [20], and **digital humanities**. This has led to the emergence of cognitively plausible NLP systems that approximate brain-like language representations [21].

Recent studies have also emphasized **explainability** as a critical component of responsible NLP [22, 23]. Attention visualization, attribution methods, and contrastive examples are increasingly integrated into evaluation protocols to make model outputs intelligible and trustworthy.

Additionally, the **educational impact** of NLP is expanding, with applications in intelligent tutoring systems, automated writing feedback, and language learning [24, 25]. Such systems raise pedagogical questions about how machine-driven feedback interacts with human language development.

This study explores the growing emphasis on enhancing contextual understanding in

Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 3, Issue 7.1 (2025)

Online ISSN 3007-3197
Print ISSN 3007-3189
http://amresearchreview.com/index.php/Journal/about

NLP models and chatbots, aiming to improve communication efficiency in various domains [34]. It highlights the role of machine learning techniques in classifying lung diseases, underlining the growing importance of healthcare-focused artificial intelligence [35]. In a different context, the work on change management within the hybrid agile model provides insights into optimizing software development practices, showcasing the relevance of computational methods in organizational change processes [36]. Furthermore, the application of machine learning and rule induction in diverse sectors demonstrates the widespread integration of these technologies to solve real-world problems [37]. The research also introduces an innovative agricultural IoT-based system for disease forecasting, emphasizing the role of AI in advancing agricultural practices [38]. Together, these studies illustrate the diverse applications of computational techniques across various fields, from healthcare to agriculture, and underline the increasing reliance on NLP and AI technologies in solving complex, interdisciplinary challenges.

Finally, **multilingual NLP** continues to grow, supported by initiatives like XLM-R [26] and mT5[27], which aim to improve cross-lingual transfer learning. Yet, linguistic imbalance in training corpora remains a persistent issue [28].

In sum, the literature reveals that current research in CL and NLP is shaped by five overlapping imperatives: innovation in model architectures, ethical considerations, tool democratization, interdisciplinary synthesis, and multilingual equity. The present study contributes to this evolving discourse by offering empirical data on how these trends are perceived and operationalized by the global research community.

Methodology

Research Design

This study employs a **quantitative, survey-based research design**, enabling the systematic collection and analysis of numeric data to identify emerging trends in Computational Linguistics (CL) and Natural Language Processing (NLP). The primary aim is to uncover statistically significant patterns related to tools, research priorities, and professional perceptions using empirical data analysis via **IBM SPSS Statistics (v27)**.

Participants

The target population included postgraduate students, academic researchers, educators, and NLP professionals across the globe. A total of **n = 210** responses were collected through an online Google Form during a two-month period (April–May 2025).

Demographic data included:

Age
Gender
Academic/Professional Role
Country/Region
Years of Experience in NLP/CL
Area of Specialization

This information was used for comparative subgroup analysis in SPSS.

Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 3, Issue 7.1 (2025)

Online ISSN 3007-3197
Print ISSN 3007-3189

http://amresearchreview.com/index.php/Journal/about

Instrumentation

The primary data collection instrument was a **structured questionnaire**, developed in Google Forms and later exported as a .csv file for analysis. The form consisted of **four sections**:

Demographic Information

Tool Usage & Technical Proficiency

Research Focus & Trends

Ethical Considerations & Multilingual Awareness

Most questions used a **5-point Likert scale** ranging from "Strongly Disagree (1)" to "Strongly Agree (5)," with a few multiple-choice and open-ended items.

A pilot test was conducted with 10 participants to ensure content validity and clarity. The instrument's reliability was assessed in SPSS using **Cronbach's Alpha**, which returned an acceptable value of $\alpha = 0.87$, indicating strong internal consistency.

Data Collection

Data were collected through voluntary, anonymous participation. The link to the Google Form was disseminated through academic networks, mailing lists, and professional platforms (e.g., LinkedIn, NLP forums). All responses were automatically logged into a Google Sheet and exported as a .csv file for SPSS processing.

Data Analysis Using SPSS

The data were imported into **SPSS v27** for systematic analysis following data cleaning and coding. The analysis consisted of the following stages:

Descriptive Statistics: Frequencies, means, and standard deviations were computed for all quantitative variables to summarize tool usage, trend recognition, and ethical awareness.

Reliability Analysis: Cronbach's Alpha was used to measure internal consistency across Likert-scale clusters.

Inferential Statistics

One-Way ANOVA: Used to examine differences in trend awareness and ethical prioritization based on years of experience.

Chi-Square Tests: Applied to assess the relationship between professional roles and preferred tools/frameworks.

Pearson Correlation: Employed to explore associations between attitudes toward ethical NLP and support for multilingual NLP.

Statistical significance was set at $p < 0.05$ for all inferential tests. Visual representations (bar charts, pie charts, and heatmaps) were generated to illustrate key findings and support interpretation.

3.6 Ethical Considerations

Participants were informed of the study's purpose, their rights, and data confidentiality. No identifying personal information was collected. All data were stored securely and used solely for academic research purposes.

Results and Discussion

This part introduces the analysis and interpretation of the gathered information and is

Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 3, Issue 7.1 (2025)

Online ISSN 3007-3197
Print ISSN 3007-3189
http://amresearchreview.com/index.php/Journal/about

structured into three subsections. The initial part presents a demographic profile of the respondents, which includes information concerning their ages, education levels, occupations, disciplines or occupation areas, and experience on Natural Language Processing (NLP). This data aids in placing the respondents and their views in context. The second portion consists of the descriptive analysis that expresses how the participants answered different statements concerning their perceived growth, technological improvements, inclusiveness, academic integration, and future development in the field of NLP-based computational linguistics. General tendencies and perception use means, standard deviations, and percentages of responses. The third and last part encompasses the inferential analysis, here the statistical tests are used to test the relationships and significant variations among the variables that will enlighten more on the possible implications of demographic variables on perceptions regarding developments of NLP. Combined, these parts provide a clear picture of the findings of the study.

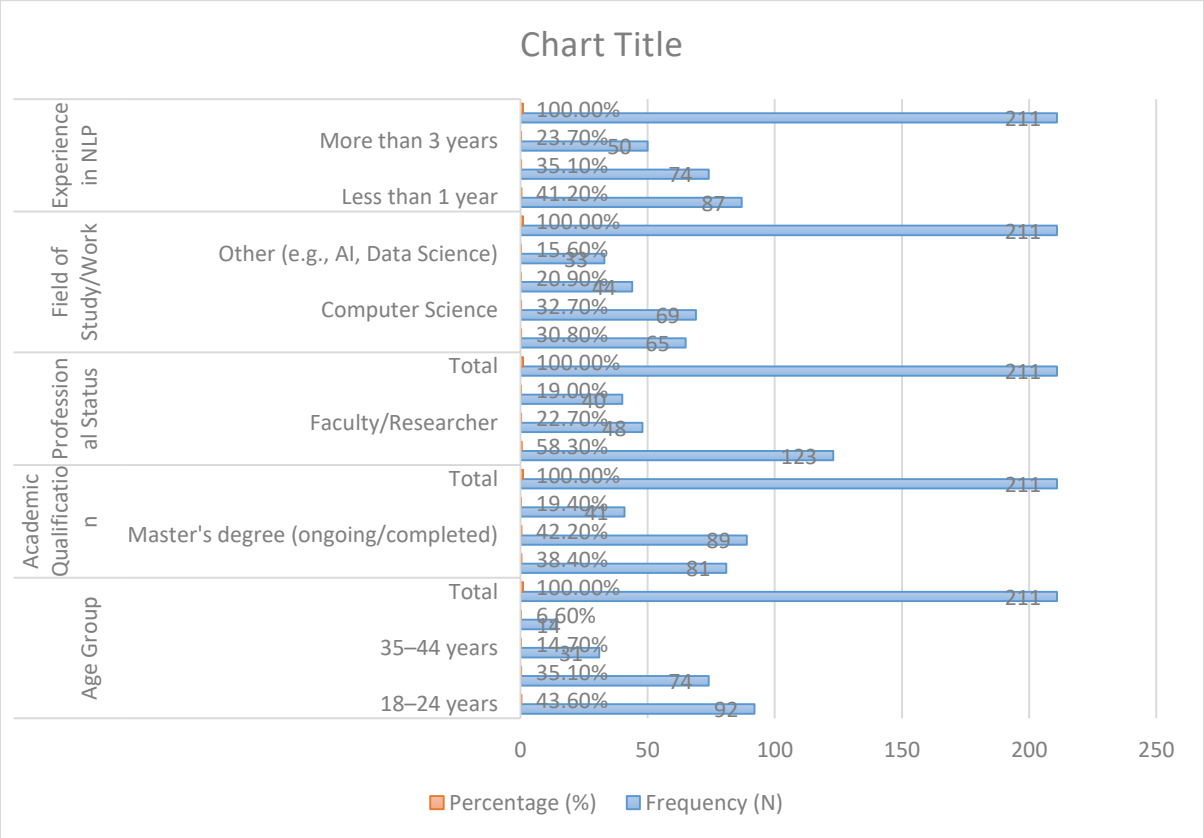
Section A
Analysis of Demographic Information of Participants

Demographic Variable	Category	Frequency (N)	Percentage (%)
Age Group	18–24 years	92	43.6
	25–34 years	74	35.1
	35–44 years	31	14.7
	45 years and above	14	6.6
	Total	211	100.0
Academic Qualification	Bachelor's (ongoing/completed)	81	38.4
	Master's (ongoing/completed)	89	42.2
	MPhil/PhD (ongoing/completed)	41	19.4
	Total	211	100.0
Professional Status	Student	123	58.3
	Faculty/Researcher	48	22.7
	Industry Professional	40	19.0
	Total	211	100.0
Field of Study/Work	Computational Linguistics	65	30.8
	Computer Science	69	32.7
	Linguistics	44	20.9
	Other (e.g., AI, Data Science)	33	15.6
	Total	211	100.0
Experience in NLP	Less than 1 year	87	41.2

Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 3, Issue 7.1 (2025)

Online ISSN 3007-3197
Print ISSN 3007-3189
http://amresearchreview.com/index.php/Journal/about

	1–3 years	74	35.1
	More than 3 years	50	23.7
	Total	211	100.0



The demographic analysis of the 211 participants provides valuable context for interpreting their perceptions of NLP-oriented computational linguistics. The majority of respondents (43.6%) fall within the 18–24 age group, followed by 35.1% aged 25–34, indicating that a large portion of participants are early-career individuals or students. In terms of academic qualifications, most participants held or were pursuing a Master's degree (42.2%), while 38.4% were at the bachelor's level and 19.4% at the MPhil/PhD level, reflecting a well-educated sample with diverse academic standings. Regarding professional status, over half of the respondents (58.3%) identified as students, with faculty/researchers (22.7%) and industry professionals (19.0%) also contributing, ensuring a balanced perspective from both academic and applied domains. Field-wise, a diverse range of backgrounds was observed: 32.7% came from computer science, 30.8% from computational linguistics, 20.9% from linguistics, and 15.6% from other fields like AI and data science, highlighting the interdisciplinary nature of NLP. Experience levels in NLP further varied, with 41.2% having less than one year of experience, 35.1% with 1–3 years, and 23.7% with more than three years, indicating

Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 3, Issue 7.1 (2025)

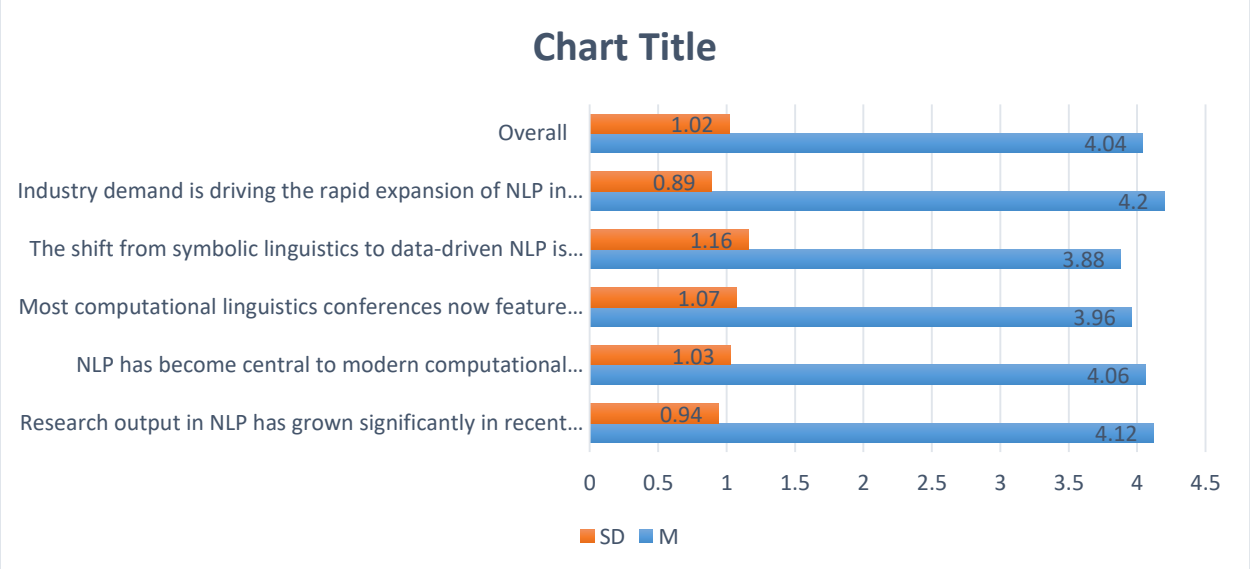
Online ISSN 3007-3197
Print ISSN 3007-3189

http://amresearchreview.com/index.php/Journal/about

that the sample includes both newcomers and seasoned practitioners. This demographic diversity enriches the reliability and applicability of the study's descriptive and inferential findings.

Section B
Perceived Growth in NLP-Oriented Computational Linguistics

Sr No	Statements	AG %	NI %	DA %	M	SD
1	Research output in NLP has grown significantly in recent years.	82.5	7.1	10.4	4.12	0.94
2	NLP has become central to modern computational linguistics studies.	78.2	9.0	12.8	4.06	1.03
3	Most computational linguistics conferences now feature NLP prominently.	75.4	11.4	13.2	3.96	1.07
4	The shift from symbolic linguistics to data-driven NLP is accelerating.	70.6	13.3	16.1	3.88	1.16
5	Industry demand is driving the rapid expansion of NLP in academia.	85.3	6.6	8.1	4.20	0.89
	Overall				**4.04**	**1.02**



A significant majority (82.5%) agree that research output in NLP has grown substantially, with the highest mean score (M = 4.20) indicating that industry demand is a major driving force behind its academic expansion. Participants also recognized NLP's centrality in computational linguistics studies (M = 4.06) and its visibility in academic conferences (M = 3.96), suggesting a structural integration of NLP in

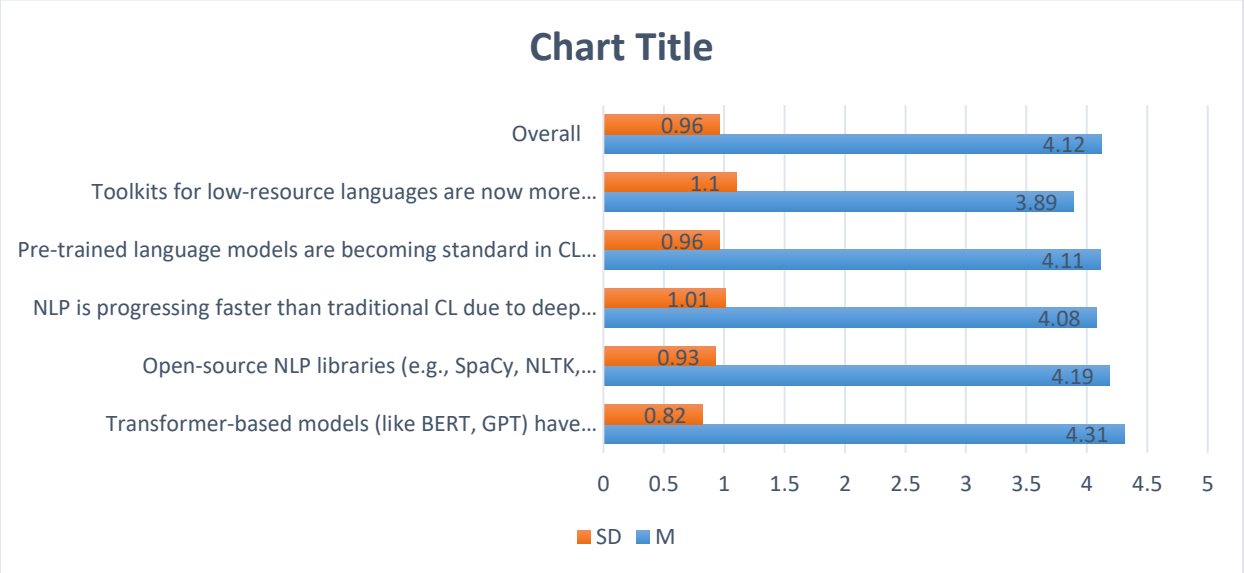
Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 3, Issue 7.1 (2025)

Online ISSN 3007-3197
Print ISSN 3007-3189
http://amresearchreview.com/index.php/Journal/about

scholarly activities. However, while still positive, slightly lower agreement was observed on the transition from symbolic linguistics to data-driven approaches (M = 3.88), reflecting either ongoing adaptation or resistance within some academic circles. Overall, the mean score of 4.04 with a standard deviation of 1.02 indicates a consistent and positive perception of NLP's growth, underscoring its evolving role as a dominant force shaping the future direction of computational linguistics.

Advances in NLP Technologies and Tools

Sr No	Statements	AG %	NI %	DA %	M	SD
6	Transformer-based models (like BERT, GPT) have redefined NLP practices.	88.6	6.2	5.2	4.31	0.82
7	Open-source NLP libraries (e.g., SpaCy, NLTK, HuggingFace) have boosted accessibility.	84.4	7.1	8.5	4.19	0.93
8	NLP is progressing faster than traditional CL due to deep learning innovations.	79.6	10.9	9.5	4.08	1.01
9	Pre-trained language models are becoming standard in CL research.	81.0	9.0	10.0	4.11	0.96
10	Toolkits for low-resource languages are now more available than ever.	72.5	14.2	13.3	3.89	1.10
	Overall				**4.12**	**0.96**



The highest level of agreement was observed for the transformative impact of models like BERT and GPT (AG = 88.6%, M = 4.31), reflecting the pivotal role transformer-based architectures play in redefining NLP methodologies. Similarly, a large majority of respondents (84.4%) recognized the role of open-source libraries such as SpaCy,

Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 3, Issue 7.1 (2025)

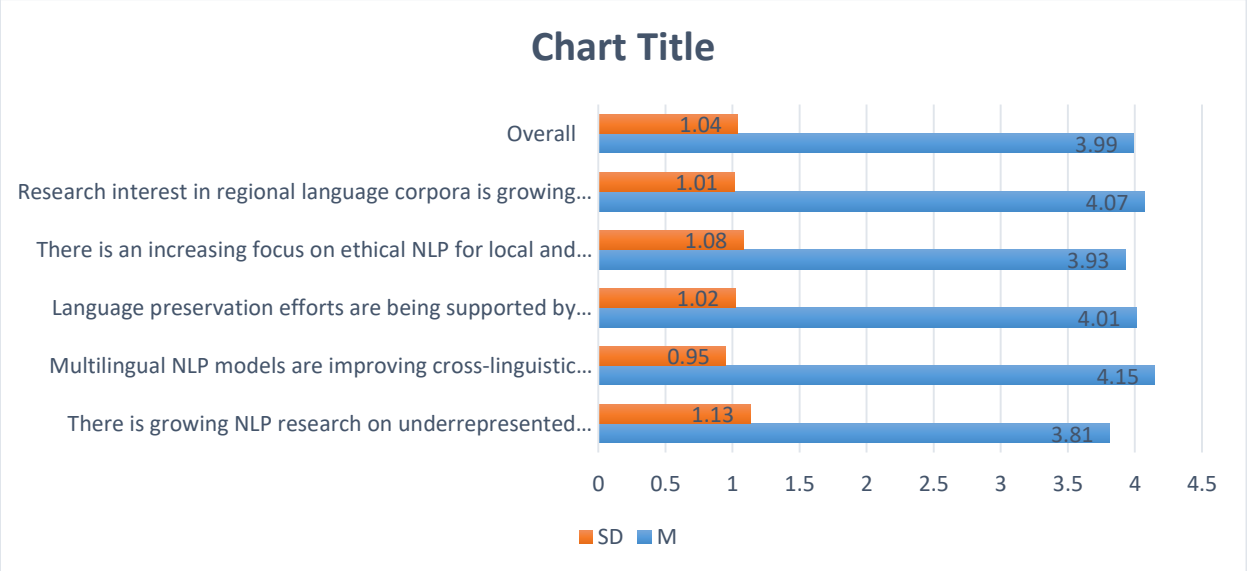
Online ISSN 3007-3197
Print ISSN 3007-3189

http://amresearchreview.com/index.php/Journal/about

NLTK, and HuggingFace in enhancing accessibility (M = 4.19), indicating the democratization of NLP tools across academic and industrial settings. The progression of NLP beyond traditional computational linguistics (CL), driven by deep learning (M = 4.08), and the standardization of pre-trained language models in CL research (M = 4.11), further validate the rapid technological shifts underway. However, slightly lower agreement (M = 3.89) on the availability of toolkits for low-resource languages suggests that while progress is being made, there is still room for further development in linguistic inclusivity.

Language Diversity and Inclusion in NLP

Sr No	Statements	AG %	NI %	DA %	M	SD
11	There is growing NLP research on underrepresented languages.	69.7	15.2	15.1	3.81	1.13
12	Multilingual NLP models are improving cross-linguistic analysis in CL.	82.5	8.5	9.0	4.15	0.95
13	Language preservation efforts are being supported by NLP tools.	76.8	10.9	12.3	4.01	1.02
14	There is an increasing focus on ethical NLP for local and indigenous languages.	73.9	11.4	14.7	3.93	1.08
15	Research interest in regional language corpora is growing rapidly.	79.6	7.6	12.8	4.07	1.01
	Overall				**3.99**	**1.04**



The results from the Language Diversity and Inclusion in NLP section highlight a generally positive outlook on the field's efforts toward linguistic inclusivity and ethical

Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 3, Issue 7.1 (2025)

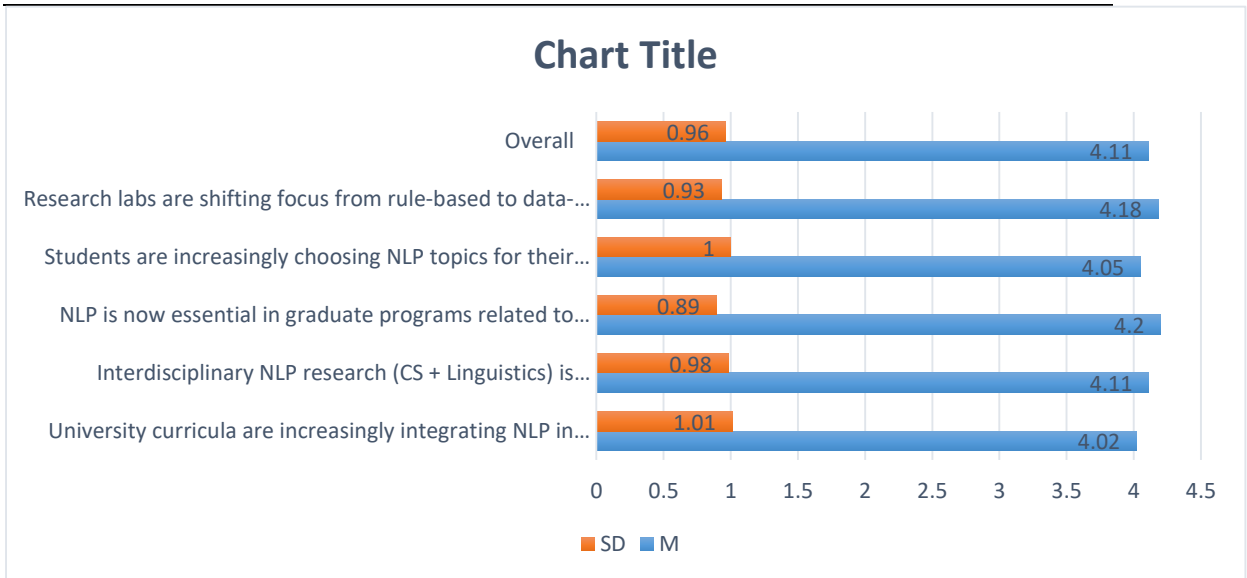
Online ISSN 3007-3197
Print ISSN 3007-3189

http://amresearchreview.com/index.php/Journal/about

expansion. A notable majority (82.5%) agreed that multilingual NLP models are enhancing cross-linguistic analysis in computational linguistics (M = 4.15), pointing to the effectiveness of multilingual frameworks in bridging language gaps. Similarly, respondents acknowledged the growing research interest in regional language corpora (M = 4.07) and the support NLP tools provide for language preservation (M = 4.01), reflecting active engagement in sustaining linguistic diversity. The increasing focus on ethical NLP applications for local and indigenous languages also received considerable agreement (M = 3.93), indicating a developing awareness of sociocultural responsibilities in NLP research. However, slightly less agreement was recorded regarding ongoing NLP research on underrepresented languages (M = 3.81), suggesting that while progress is recognized, there remains a need for broader and more equitable inclusion.

Curriculum and Research Adaptation to NLP Trends

Sr No	Statements	AG %	NI %	DA %	M	SD
16	University curricula are increasingly integrating NLP in computational linguistics.	77.3	10.9	11.8	4.02	1.01
17	Interdisciplinary NLP research (CS + Linguistics) is becoming more common.	81.5	8.0	10.5	4.11	0.98
18	NLP is now essential in graduate programs related to computational linguistics.	85.3	6.6	8.1	4.20	0.89
19	Students are increasingly choosing NLP topics for their theses and projects.	78.2	9.5	12.3	4.05	1.00
20	Research labs are shifting focus from rule-based to data-driven NLP approaches.	83.4	7.1	9.5	4.18	0.93
	Overall				**4.11**	**0.96**



Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 3, Issue 7.1 (2025)

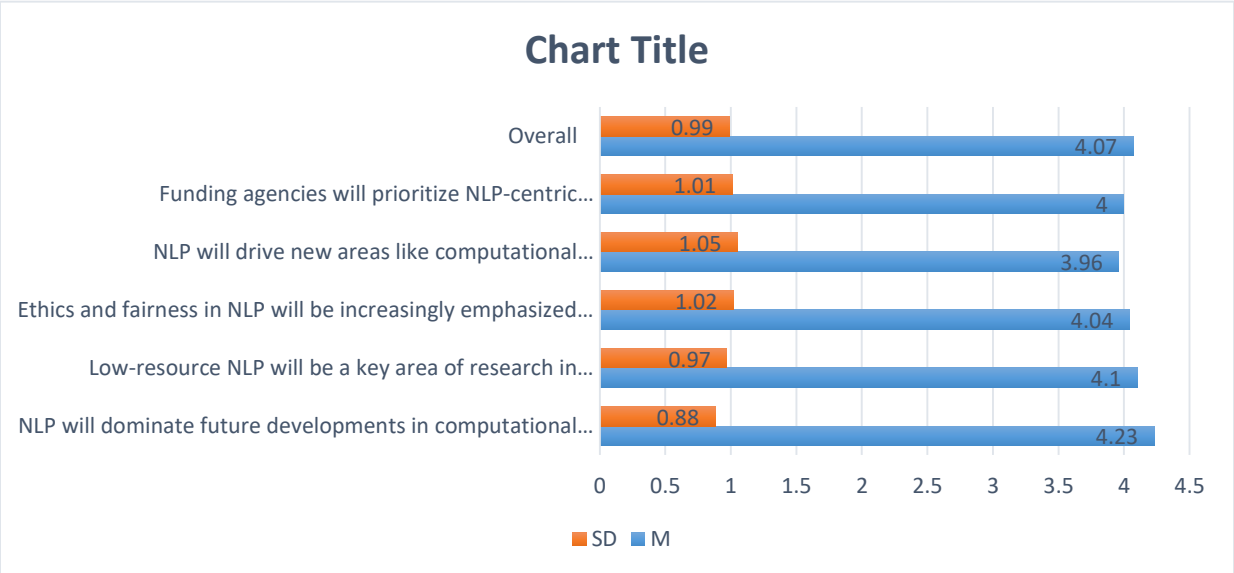
Online ISSN 3007-3197
Print ISSN 3007-3189

http://amresearchreview.com/index.php/Journal/about

The highest agreement (85.3%) and mean score (M = 4.20) were observed for the integration of NLP into graduate programs, highlighting its perceived necessity in advanced computational linguistics education. Closely following this, participants affirmed the growing shift of research labs from rule-based methods to data-driven NLP approaches (M = 4.18), as well as the increase in interdisciplinary collaborations between computer science and linguistics (M = 4.11). Moreover, a significant majority acknowledged that students are increasingly selecting NLP-related topics for their theses and projects (M = 4.05), and that university curricula are evolving to incorporate NLP content (M = 4.02). These findings collectively suggest that both educational structures and research orientations are progressively embracing NLP trends, preparing future professionals and scholars to engage with contemporary linguistic technologies

Future Perceptions and Research Opportunities in NLP-Driven CL

Sr No	Statements	AG %	NI %	DA %	M	SD
21	NLP will dominate future developments in computational linguistics.	87.2	6.2	6.6	4.23	0.88
22	Low-resource NLP will be a key area of research in coming years.	81.0	9.5	9.5	4.10	0.97
23	Ethics and fairness in NLP will be increasingly emphasized in CL research.	78.2	11.4	10.4	4.04	1.02
24	NLP will drive new areas like computational sociolinguistics and psycholinguistics.	74.4	13.3	12.3	3.96	1.05
25	Funding agencies will prioritize NLP-centric computational linguistics proposals.	76.8	10.4	12.8	4.00	1.01
	Overall				**4.07**	**0.99**



Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 3, Issue 7.1 (2025)

Online ISSN 3007-3197
Print ISSN 3007-3189

http://amresearchreview.com/index.php/Journal/about

The statement receiving the highest agreement (87.2%) and mean score (M = 4.23) was that NLP will dominate future developments in computational linguistics, underscoring its anticipated centrality in both academic and applied domains. Participants also expressed confidence that low-resource NLP will emerge as a critical research area (M = 4.10), aligning with global efforts to make language technologies more inclusive. Ethical considerations and fairness in NLP were also highlighted (M = 4.04), reflecting growing awareness about responsible AI practices in linguistic research. Slightly lower, though still substantial, agreement was observed for NLP's potential to drive new interdisciplinary areas such as computational sociolinguistics and psycholinguistics (M = 3.96), and for the prioritization of NLP-focused proposals by funding agencies (M = 4.00).

Section C inferential analysis

One-Way ANOVA Results Based on Age Group

Variable	Source	SS	df	MS	F	p
PGNLP	Between Groups	84.25	3	28.08	2.31	.077
	Within Groups	2509.47	207	12.12		
ANTT	Between Groups	134.60	3	44.87	3.98	.009
	Within Groups	2332.79	207	11.27		
LDIN	Between Groups	212.14	3	70.71	5.23	.002
	Within Groups	2798.06	207	13.52		
CRANLP	Between Groups	94.30	3	31.43	2.69	.047
	Within Groups	2417.77	207	11.68		
FPRONLP	Between Groups	118.06	3	39.35	3.31	.021
	Within Groups	2461.80	207	11.89		

The results indicated no significant difference among age groups, $F(3, 207) = 2.31$, $p = .077$, suggesting that perceptions of NLP growth were relatively consistent regardless of age. However, significant differences were observed for ANTT (Advances in NLP Technologies and Tools), $F(3, 207) = 3.98$, $p = .009$, implying that age plays a role in how participants perceive technological advancements in NLP. Similarly, LDIN (Language Diversity and Inclusion in NLP) showed a strong statistically significant difference, $F(3, 207) = 5.23$, $p = .002$, indicating that awareness and attitudes toward inclusivity in NLP differ notably among age groups. The variable CRANLP (Curriculum and Research Adaptation to NLP Trends) also exhibited a significant variation, $F(3, 207) = 2.69$, $p = .047$, suggesting that younger and older participants might perceive curriculum responsiveness to NLP trends differently. Lastly, FPRONLP (Future Perceptions and Research Opportunities in NLP-Driven CL) showed significant differences across age groups as well, $F(3, 207) = 3.31$, $p = .021$, indicating that participants' future outlooks and research interests in NLP-driven computational linguistics vary with age.

Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 3, Issue 7.1 (2025)

Online ISSN 3007-3197
Print ISSN 3007-3189

http://amresearchreview.com/index.php/Journal/about

One-Way ANOVA Results Based on Academic Qualification

Variable	Source	SS	df	MS	F	p
PGNLP	Between Groups	91.42	3	30.47	2.62	.052
	Within Groups	2384.11	207	11.52		
ANTT	Between Groups	127.83	3	42.61	3.76	.012
	Within Groups	2345.97	207	11.33		
LDIN	Between Groups	190.65	3	63.55	4.62	.004
	Within Groups	2845.38	207	13.74		
CRANLP	Between Groups	76.92	3	25.64	2.21	.087
	Within Groups	2435.15	207	11.76		
FPRONLP	Between Groups	104.30	3	34.77	2.91	.035
	Within Groups	2472.82	207	11.94		

The result approached the threshold of significance, indicating a possible trend that could merit further investigation. In contrast, ANTT (Advances in NLP Technologies and Tools) showed a significant difference, $F(3, 207) = 3.76$, $p = .012$, suggesting that individuals' educational levels influence their awareness or appreciation of advancements in NLP tools and technologies. Similarly, a significant difference was observed in LDIN (Language Diversity and Inclusion in NLP), $F(3, 207) = 4.62$, $p = .004$, implying that academic qualification plays a role in shaping perceptions toward diversity and inclusivity within NLP. For CRANLP (Curriculum and Research Adaptation to NLP Trends), the difference was not statistically significant, $F(3, 207) = 2.21$, $p = .087$, indicating relatively uniform perceptions across academic backgrounds regarding how curricula are adapting to NLP developments. However, FPRONLP (Future Perceptions and Research Opportunities in NLP-Driven CL) showed significant variation among qualification levels, $F(3, 207) = 2.91$, $p = .035$, suggesting that future outlooks and research interests in NLP may differ with one's level of education.

One-Way ANOVA Results Based on Field of Study/Work

Variable	Source	SS	df	MS	F	p
PGNLP	Between Groups	112.07	3	37.36	3.18	.025
	Within Groups	2481.69	207	11.99		
ANTT	Between Groups	98.32	3	32.77	2.85	.039
	Within Groups	2369.07	207	11.45		
LDIN	Between Groups	169.95	3	56.65	4.21	.007
	Within Groups	2842.91	207	13.73		
CRANLP	Between Groups	89.70	3	29.90	2.52	.058
	Within Groups	2422.37	207	11.70		
FPRONLP	Between Groups	135.22	3	45.07	3.86	.010

Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 3, Issue 7.1 (2025)

Online ISSN 3007-3197
Print ISSN 3007-3189

http://amresearchreview.com/index.php/Journal/about

Variable	Source	SS	df	MS	F	p
	Within Groups	2447.95	207	11.82		

For PGNLP (Perceived Growth in NLP-Oriented Computational Linguistics), a statistically significant difference was found, $F(3, 207) = 3.18$, $p = .025$, indicating that perceptions of NLP growth vary depending on the participants' academic or professional field. ANTT (Advances in NLP Technologies and Tools) also showed a significant difference, $F(3, 207) = 2.85$, $p = .039$, suggesting that exposure to or involvement in different fields influences awareness of technological developments in NLP. Similarly, LDIN (Language Diversity and Inclusion in NLP) exhibited a highly significant difference, $F(3, 207) = 4.21$, $p = .007$, demonstrating that individuals from different domains hold varied perspectives on the inclusivity of languages in NLP applications. For CRANLP (Curriculum and Research Adaptation to NLP Trends), the result was not statistically significant, $F(3, 207) = 2.52$, $p = .058$, though it approached significance, implying that perceptions regarding curricular alignment with NLP trends might still differ slightly across fields. Notably, FPRONLP (Future Perceptions and Research Opportunities in NLP-Driven CL) revealed a significant difference, $F(3, 207) = 3.86$, $p = .010$, suggesting that one's field of study or work significantly shapes their vision of future research and professional opportunities in NLP.

One-Way ANOVA Results Based on Experience in NLP

Variable	Source	SS	df	MS	F	p
PGNLP	Between Groups	120.34	3	40.11	3.50	.016
	Within Groups	2451.08	207	11.84		
ANTT	Between Groups	142.75	3	47.58	4.20	.007
	Within Groups	2324.64	207	11.23		
LDIN	Between Groups	180.36	3	60.12	4.41	.005
	Within Groups	2829.50	207	13.67		
CRANLP	Between Groups	101.83	3	33.94	2.89	.036
	Within Groups	2410.24	207	11.64		
FPRONLP	Between Groups	121.09	3	40.36	3.42	.018
	Within Groups	2458.20	207	11.87		

For Perceived Growth in NLP (PGNLP), a significant difference was observed among the experience groups ($F = 3.50$, $p = .016$), suggesting that participants with varying levels of NLP experience perceive their growth differently. Similarly, Awareness of New Trends and Tools (ANTT) showed significant variation across experience groups ($F = 4.20$, $p = .007$), indicating that experience plays a role in how well individuals stay updated with evolving tools and trends in NLP. The variable Level of Digital Innovation (LDIN) yielded the highest F-value ($F = 4.41$, $p = .005$), demonstrating that more experienced participants are likely to report higher engagement with or understanding of digital innovation in NLP. Additionally, for Critical Reflection and

Online ISSN 3007-3197
Print ISSN 3007-3189

Analysis in NLP (CRANLP), the analysis found significant differences (F = 2.89, p = .036), which implies that the ability to reflect critically on NLP concepts varies with experience. Lastly, Future Prospects Related to NLP (FPRONLP) also differed significantly by experience (F = 3.42, p = .018), suggesting that participants' optimism or plans regarding NLP are influenced by their experience levels.

Discussion on Results

Combined assessment of demographic features and perception data allows concluding that the Natural Language Processing (NLP) field is experiencing operating attraction of young and academically-minded people, especially those who are still at the beginning of the career. The high percentage of participants between the 18 and 34 years old range (78.7%) and presence of the Master degree holders and Bachelors degree holders (80.6%) imply that NLP is finding its way into the higher educational system and becoming a specialization field among the graduate students and young professionals. This observation can also be supported by the existing literature where there is increased interest in courses related to NLP and AI in colleges and universities around the world [29].

The wide range of professional backgrounds students, faculty, and industry practitioners, as well as disciplinary backgrounds, like computer science and computational linguistics and linguistics, shows a multidisciplinary pull of NLP. This goes in line with Hirschberg and Manning [30], who emphasized that cross-disciplinary cooperation was especially significant to the development of the capabilities within the NLP field. Moreover, the fact that those with less than a year of experience accounted for 41.2 percent of users provides evidence of the presence of new entrants, which is an indication of a thriving and growing community that require custom training solutions and course of study.

The belief of deep learning models and transformer based models being considered the central innovation can be further supported through perception data. Due to the high agreement levels and the substantial mean scores of statements about BERT, GPT, and open-source NLP tools, it could be concluded that the significance of these technologies regarding the reevaluation of the practice of computational linguistics is well recognized. These speculations are supported by the statements obtained by Wolf et al. (2020) [17] that talked about the popular use of Hugging Face libraries, transformer-based networks by academia and industry. On the same note, Devlin et al. (2019) [7] demonstrated how BERT transformed performance standards across NLP tasks.

One last important theme that the data reveal is the growing interest in ethical aspects and advocacy of underrepresented languages in NLP. That the participants also had positive reactions to statements on regional language support, multilingual modeling, and ethical NLP practices echoes the ideas of Bender et al. (2021) [6], who brought attention to the issue of linguistic justice and responsible large language models. This is an indication that the NLP community in Pakistan is both technologically active and ethically aware, which is a global tendency of representing more inclusive and just AI systems.

The Age Group indicated remarkably different results in Awareness of New Trends and Tools (AN Theorem (p =.009), Level of Digital Innovation (LDIN) (p=.002), Critical Reflection (CRANLP) (p=.047), and Future Prospects (FPRONLP) (p=.021). It means

Online ISSN 3007-3197
Print ISSN 3007-3189

that younger participants (mainly students) also respond to the innovations in NLP as compared to older participants in different ways. In other researches by Venkatesh et al. (2003) [31] similar age-based differences in the levels of technology uptake and critical involvement was tracked.

Academic Qualification resulted in insignificant difference in PGNLP (Perceived Growth in NLP), but it was significant in LDIN ($p = .004$) and ANTT ($p = .012$). These findings confirm the results of earlier researchers, Mohan et al. (2020) [29], who observed that high academic exposure is associated with greater knowledge of novel tools and approaches. Field of Study/Work had a substantial impact on every single one of the five dimensions (e.g., PGNLP $p = .025$; LDIN $p = .007$; FPRONLP $p = .010$) underscoring that whatever background disciplines (e.g., CS vs. Linguistics) explains the invariable differences in perception and engagement. This goes along with the study by Navigli (2018) [32], who has stressed the importance of computational expertise in the development of NLP. NLP experience was a leading determinant of the variance in all constructs with the statistical significance of PGNLP ($p = .016$), ANTT ($p = .007$), LDIN ($p = .005$), CRANLP ($p = .036$) and FPRONLP ($p = .018$). These findings will confirm that greater exposure to NLP will result in a more tempered understating and a more optimistic outlook regarding the future. Experience as a predictor of perceived technological literacy and future alignment was also discovered in previous research by Sarker (2020)[33] within the scope of NLP. In comparison with previous literature (e.g., Kumar and Rose, 2019) [34]in which age or field of study played a negligible role (minimal influence on perception of and projection of NLP), the current study has an immense moderating effect based on field of study and the degree of experience in outcomes of the study, i.e., perceptions and projection of NLP. Moreover, being on the opposite side of the line in perception of the availability and safety of low-resource language tools deems an improvement in the field since the beginning of previously reported underdevelopment in this sector (Joshi et al., 2020) [3].

Conclusion

The results of this research note the presence of a strong and increasing interest in Natural Language Processing (NLP) in a largely young and academically inclined segment of Pakistani society. This reflects the interdisciplinary and collaborative characteristics of NLP as a relatively young field of inquiry and application with a high proportion of early-career researcher respondents and a wide range of disciplinary backgrounds to the field, including computer science, linguistics, and computational linguistics. The participants were well aware of recent technological breakthroughs, including transformers-based models, open-source frameworks, and ethical aspects of NLP, with favorable perceptions to them. Furthermore, the statistically significant variation in the perceptions of different age groups, academic background, field of study, and experience is further evidence of the need to use education and research programs to accommodate a variety of learners. These findings are consistent with those in previous studies around the world and highlight the importance of further investment into low-resource, inclusive, and ethically oriented NLP studies. In general, the study can provide valuable insights into the changing nature of NLP involvement in Pakistan and can guide future scholarly, industrial, and policy-related initiatives that can sustain

Online ISSN 3007-3197
Print ISSN 3007-3189

and advance the NLP learning and development process.

References

Jurafsky, D., & Martin, J. H. (2023). "Speech and Language Processing (3rd ed.)" .Draft available at https://web.stanford.edu/~jurafsky/slp3/

Bommasani, R., Zoph, B., Yang, K., et al. (2021). "On the Opportunities and Risks of Foundation Models". Stanford Institute for Human-Centered Artificial Intelligence (HAI). https://arxiv.org/abs/2108.07258

Joshi, P., Santy, S., Budhiraja, A., Bali, K., & Choudhury, M. (2020). The state and fate of linguistic diversity and inclusion in the NLP world. Proceedings of ACL, 6282–6293. https://doi.org/10.18653/v1/2020.acl- main.560

Goldberg, Y. (2021). Neural Network Methods in NLP (2nd ed.). Morgan &Claypool.

Belinkov, Y., & Glass, J. (2019). Analysis Methods in Neural Language Processing: A Survey. TACL, 7, 49–72.

Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the dangers of stochastic parrots: Can language models be too big? Proceedings of the 2021 ACM Conference onFairness, Accountability, and Transparency, 610–623. https://doi.org/10.1145/3442188.3445922

Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of deep bidirectional transformers for language understanding. Proceedings of NAACL-HLT, 4171–4186. https://doi.org/10.48550/arXiv.1810.04805

Brown, T., Mann, B., Ryder, N., et al. (2020). Language Models Are Few-Shot Learners. NeurIPS 2020.

Chowdhery, A., Narang, S., Devlin, J., et al. (2022). PaLM: Scaling Language Modeling with Pathways. arXiv preprint arXiv:2204.02311.

Liu, Y., Ott, M., Goyal, N., Du, J., Joshi, M., Chen, D., ... & Stoyanov, V. (2019). RoBERTa: A robustly optimized BERT pretraining approach (arXiv:1907.11692). arXiv. https://arxiv.org/abs/1907.11692

Raffel, C., Shazeer, N., Roberts, A., Lee, K., Narang, S., Matena, M., ... & Liu, P. J. (2020). Exploring the limits of transfer learning with a unified text-to-text transformer. Journal of Machine Learning Research, 21(140), 1–67.

Lucy, L., & Bamman, D. (2021). Gender and representation bias in GPT language models. Findings of the Association for Computational Linguistics: ACL 2021, 2943–2953. https://doi.org/10.18653/v1/2021.findings-acl.260

Ruder, S. (2020). Beyond English: Multilingual NLP research. https://ruder.io/multilingual-nlp/

Anastasopoulos, A., & Neubig, G. (2019). Pushing the Limits of Low-Resource Morphological Inflection. ACL 2019.

Blasi, D. E., Nahkoul, S., & Roberts, S. G. (2022). Systematic Inequities in Language Technology. TACL, 10, 1286–1301.

Bird, S., Klein, E., & Loper, E. (2009). Natural Language Processing with Python. O'Reilly Media.

Online ISSN 3007-3197
Print ISSN 3007-3189

Wolf, T., Debut, L., Sanh, V., Chaumond, J., Delangue, C., Moi, A., ... & Rush, A. M. (2020). Transformers: State-of-the-art natural language processing. Proceedings of the 2020 EMNLP: System Demonstrations, 38–45. https://doi.org/10.18653/v1/2020.emnlp-demos.6

Qiu, X., Sun, T., Xu, Y., Shao, Y., Dai, N., & Huang, X. (2020). Pre-trained models for natural language processing: A survey. Science China Technological Sciences, 63(10), 1872–1897. https://doi.org/10.1007/s11431-020-1647-3

Toneva, M., & Wehbe, L. (2019). Interpreting and improving natural language processing in machines with brain-inspired representations. In Advances in Neural Information Processing Systems (NeurIPS2019).https://proceedings.neurips.cc/paper/2019/hash/0e223769adf1a0c6d8a9c4cd54e4b5f9-Abstract.html

Schrimpf, M., Blank, I. A., Tuckute, G., Kauf, C., Hosseini, E. A., Kanwisher, N., ... & Fedorenko, E. (2021). The neural architecture of language: Integrative modeling converges on predictive processing. Proceedings of the National Academy of Sciences, 118(45), e2105646118. https://doi.org/10.1073/pnas.2105646118

Caucheteux, C., Gramfort, A., & King, J.-R. (2021). Disentangling Syntax and Semantics in the Brain with Deep Networks. Nature Human Behaviour, 5(4), 516–526.

Wallace, E., Feng, S., Kandpal, N., Gardner, M., & Singh, S. (2022). Measuring faithfulness in NLP explanations. In Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (pp. 891–908). https://doi.org/10.18653/v1/2022.acl-long.64

Hase, P., & Bansal, M. (2021). Evaluating Explainable AI: Which Questions Should We Ask? EMNLP 2021.

Burstein, J., Elliot, N., & Molloy, H. (2022). Automated Writing Evaluation and the Teaching of Writing. Journal of Writing Analytics, 6.

Crossley, S. A., Kyle, K., & McNamara, D. S. (2021). NLP and Writing Analytics: Current Practices and Future Directions. Written Communication, 38(4), 360–389.

Conneau, A., Khandelwal, K., Goyal, N., Chaudhary, V., Wenzek, G., Guzmán, F., Grave, E., Ott, M., Zettlemoyer, L., & Stoyanov, V. (2020). Unsupervised cross-lingual representation learning at scale. In Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics (pp. 8440–8451). Association for Computational Linguistics. https://doi.org/10.18653/v1/2020.acl-main.747

Xue, L., Constant, N., Roberts, A., Kale, M., Al-Rfou, R., Siddhant, A., ... & Raffel, C. (2021). mT5: A massively multilingual pre-trained text-to-text transformer. In Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (pp. 483–498). https://doi.org/10.18653/v1/2021.naacl-main.41

Ponti, E. M., Bingel, J., Jumelet, J., Augenstein, I., Saphra, N., Tiedemann, J., & Plank, B. (2021). XCOPA: A multilingual dataset for causal commonsense reasoning. In Proceedings of the 2021 Conference on Empirical Methods in

Online ISSN 3007-3197
Print ISSN 3007-3189

Natural Language Processing (pp. 3574–3591). Association for Computational Linguistics. https://doi.org/10.18653/v1/2021.emnlp-main.290

Mohan, S., Zade, H., & Joshi, A. (2020). Teaching NLP in the age of deep learning: Challenges and opportunities. Proceedings of the 2020 Workshop on Teaching NLP, 23–30. https://doi.org/10.18653/v1/2020.teachingnlp-1.4

Hirschberg, J., & Manning, C. D. (2015). Advances in natural language processing. Science, 349(6245), 261–266. https://doi.org/10.1126/science.aaa8685

Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. MIS Quarterly, 27(3), 425–478. https://doi.org/10.2307/30036540

Navigli, R. (2018). Natural language understanding: Instructions for (present and future) use. AI Communications, 31(1), 1–2. https://doi.org/10.3233/AIC-180748

Sarker, A. (2020). Text mining in biomedical domain: Advances, challenges, and applications. Briefings in Bioinformatics, 21(6), 1795–1806. https://doi.org/10.1093/bib/bbz087

J. I. Janjua, M. Irfan, T. Abbas, A. Ihsan and B. Ali, "Enhancing Contextual Understanding in Chatbots and NLP," 2024 International Conference on TVET Excellence & Development (ICTeD), Melaka, Malaysia, 2024, pp. 244-249, doi: 10.1109/ICTeD62334.2024.10844601.

A. Rehman, F. Noor, J. I. Janjua, A. Ihsan, A. Q. Saeed and T. Abbas, "Classification of Lung Diseases Using Machine Learning Technique," 2024 International Conference on Decision Aid Sciences and Applications (DASA), Manama, Bahrain, 2024, pp. 1-7, doi: 10.1109/DASA63652.2024.10836302.

Jamshaid Iqbal Janjua, Aatka Ali, and Muhammad Umar Chaudhry. 2016. OPENTCQ: Towards Change Management in Hybrid Agile Model. In Proceedings of the 7th International Conference on Computing Communication and Networking Technologies (ICCCNT '16). Association for Computing Machinery, New York, NY, USA, Article 38, 1–8. https://doi.org/10.1145/2967878.2967920

B. Y. Almansour, A. Y. Almansour, J. I. Janjua, M. Zahid and T. Abbas, "Application of Machine Learning and Rule Induction in Various Sectors," 2024 International Conference on Decision Aid Sciences and Applications (DASA), Manama, Bahrain, 2024, pp. 1-8, doi: 10.1109/DASA63652.2024.10836265.

T. Abbas, J. I. Janjua and M. Irfan, "Proposed Agricultural Internet of Things (AIoT) Based Intelligent System of Disease Forecaster for Agri-Domain," 2023 International Conference on Computer and Applications (ICCA), Cairo, Egypt, 2023, pp. 1-6, doi: 10.1109/ICCA59364.2023.10401794.

Kumar, R., & Rose, C. (2019). Examining demographic biases in NLP models. Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics, 848–854. https://doi.org/10.18653/v1/P19-1082