International Research Journal of Humanities, Language and Literature
ISSN: (2394-1642)
Impact Factor 6.972Volume 8, Issue 3, March2021
Association of Academic Researchers and Faculties (AARF)
Website-www.aarf.asia, Email : editor@aarf.asia , editoraarf@gmail.com

संस्कृतभाषा कृत्रिमप्रज्ञा च : सम्भावनाः सीमाश्च

डॉ.नीतू अवस्थी
सहायक प्राध्यापक- संस्कृत
राजकीय स्नातकोत्तर महाविद्यालय
नोएडा, गौतमबुद्धनगर

सारांशः
अस्मिन् शोधपत्रे संस्कृतभाषायाः कृत्रिमप्रज्ञातन्त्रज्ञानेन सह सम्बन्धस्य विशदं विवेचनं क्रियते। संस्कृतं विश्वस्य प्राचीनतमासु वैज्ञानिकभाषासु अन्यतमं मन्यते, तस्य व्याकरणं च गणनात्मकदृष्ट्या अत्यन्तं सुव्यवस्थितम् अस्ति। पाणिनिना प्रणीतम् अष्टाध्यायी इति व्याकरणग्रन्थे यानि सूत्राणि प्रतिपादितानि, तानि आधुनिकसङ्गणकविज्ञानस्य औपचारिकभाषासिद्धान्तैः सह साम्यं प्रदर्शयन्ति। तथापि संस्कृतस्य कृत्रिमप्रज्ञायां प्रयोगस्य मार्गे बहवः सीमाः अपि वर्तन्ते, यासां विश्लेषणम् अत्र प्रस्तूयते। प्राकृतिकभाषासंसाधनम्, यान्त्रिकानुवादः, ज्ञानप्रतिनिधित्वम् इत्यादिषु क्षेत्रेषु संस्कृतस्य योगदानं किं भवितुम् अर्हति इति अत्र सप्रमाणं विमृश्यते।

मुख्यशब्दाः : संस्कृतम्, कृत्रिमप्रज्ञा, पाणिनिव्याकरणम्, प्राकृतिकभाषासंसाधनम्, यान्त्रिकानुवादः, गणनात्मकभाषाविज्ञानम्।

१. प्रस्तावना
एकविंशतितमे शताब्दे कृत्रिमप्रज्ञायाः (Artificial Intelligence) क्षेत्रे यत् परिवर्तनं दृश्यते तत् अभूतपूर्वम् इति वक्तुं शक्यते। गहनशिक्षणम् (Deep Learning), बृहद्भाषाप्रतिरूपाणि (Large Language Models), स्वाभाविकभाषासंसाधनम् (Natural Language Processing) इत्यादीनि तन्त्रज्ञानानि मानवसमाजस्य बहुषु क्षेत्रेषु आमूलपरिवर्तनं कुर्वन्ति। एतस्मिन् सन्दर्भे एकस्याः प्राचीनभाषायाः – संस्कृतस्य – भूमिकायाः पुनर्विचारः आवश्यकः प्रतीयते।

संस्कृतभाषा केवलं साहित्यिकभाषा एव नास्ति, अपितु सा एका तार्किकव्यवस्था अपि अस्ति। भर्तृहरिणा वाक्यपदीये प्रतिपादितः स्फोटसिद्धान्तः, पाणिनेः अष्टाध्याय्याः सूत्रपद्धतिः, नव्यन्यायस्य औपचारिकतर्कपद्धतिः – एतानि सर्वाणि आधुनिकगणनात्मकचिन्तनस्य पूर्वरूपाणि इति बहवः विद्वांसः

© Associationof Academic Researchers and Faculties (AARF)
A Monthly Double-Blind Peer Reviewed Refereed Open Access International e-Journal - Included in the International Serial Directories.

मन्यन्ते। Rick Briggs इत्यस्य NASA-संस्थायां १९८५ तमे वर्षे प्रकाशितं शोधपत्रं "Knowledge Representation in Sanskrit and Artificial Intelligence" इति संस्कृतस्य गणनात्मकक्षमतायाः विषये विस्तृतां चर्चां प्रस्तौति।

किन्तु यदा वयं संस्कृतस्य कृत्रिमप्रज्ञायां वास्तविकप्रयोगं विचारयामः, तदा चित्रं किञ्चित् जटिलतरं भवति। सैद्धान्तिकी क्षमता एव पर्याप्ता न भवति — व्यावहारिकाः साधनानि, पर्याप्ताः आँकडासमुच्चयाः (datasets), प्रशिक्षितकर्मशक्तिः च आवश्यकाः भवन्ति। अस्मिन् शोधपत्रे उभयोः पक्षयोः — सम्भावनानां सीमानां च — सन्तुलितं मूल्याङ्कनं प्रस्तूयते।

यद्यदाचरति श्रेष्ठस्तत्तदेवेतरो जनः ।
स यत्प्रमाणं कुरुते लोकस्तदनुवर्तते ॥
— श्रीमद्भगवद्गीता ३.२१

अयं श्लोकः संस्कृतपरम्परायाः प्रामाणिकतां सूचयति। यथा श्रेष्ठानां कार्याणि अनुसरणीयानि भवन्ति, तथैव पाणिन्यादिप्राचीनवैयाकरणानां कार्यं गणनात्मकभाषाविज्ञानस्य क्षेत्रे प्रमाणभूतं भवितुम् अर्हति इति अस्य भावः।

२. संस्कृतभाषायाः गणनात्मकं महत्त्वम्

२.१ पाणिनीयव्याकरणस्य औपचारिकसंरचना

पाणिनिना विरचितं अष्टाध्यायी इति व्याकरणं लगभगं ३,९५९ सूत्रेषु संस्कृतभाषायाः सम्पूर्णं वर्णनं करोति। एतत् व्याकरणं केवलं नियमसङ्ग्रहः नास्ति — अपितु एतत् एकं जनयित्रि व्याकरणम् (Generative Grammar) अस्ति यत् परिमितनियमैः अपरिमितवाक्यानां निर्माणं करोति। Noam Chomsky इत्यस्य परिवर्तनात्मकजनयित्रिव्याकरणम् (Transformational Generative Grammar) पाणिनेः कार्यस्य सहस्रवर्षानन्तरं प्रतिपादितम् इति ध्यातव्यम्।

पाणिनेः सूत्रेषु प्रत्ययविधानम्, सन्धिनियमाः, समासपद्धतिः, कारकव्यवस्था च यथा व्यवस्थितानि सन्ति, तथा तानि सङ्गणकीयसंसाधनस्य कृते अत्यन्तम् अनुकूलानि। गेरार्ड ह्यूट (Gerard Huet) इत्यस्य Sanskrit Heritage Engine इति परियोजना एतस्य व्यावहारिकम् उदाहरणम् अस्ति, यत्र पाणिनीयसूत्राणां गणनात्मकं कार्यान्वयनं कृतम् अस्ति।

२.२ संस्कृतस्य तार्किकसंरचना

संस्कृतभाषायाः वाक्यरचना प्रायः असन्दिग्धा भवति। यत्र अन्यासु भाषासु बहुअर्थकता (ambiguity) एकं गम्भीरं समस्यारूपम् अस्ति, तत्र संस्कृतस्य विभक्तिव्यवस्था कारकसम्बन्धान् स्पष्टतया निर्धारयति।

© Associationof Academic Researchers and Faculties (AARF)
A Monthly Double-Blind Peer Reviewed Refereed Open Access International e-Journal - Included in the International Serial Directories.

उदाहरणार्थम्, "रामः रावणं हन्ति" इति वाक्ये कर्ता कर्म च विभक्तिभिः स्पष्टतया ज्ञायेते – पदक्रमः तु परिवर्तनीयः भवति कवितायां गद्ये च, किन्तु अर्थः स्थिरः तिष्ठति।

एषा विशेषता प्राकृतिकभाषासंसाधनस्य कृते महत्त्वपूर्णा भवति यतः सङ्गणकाय अर्थनिर्णयः सरलतरः भवति। तथापि एतत् अतिसरलीकरणं स्यात् यदि वयं वदेम संस्कृतं पूर्णतया असन्दिग्धम् इति। काव्ये श्लेषालङ्कारस्य विपुलः प्रयोगः, शास्त्रीयग्रन्थेषु परिभाषावैविध्यं च सन्दिग्धतायाः उदाहरणानि सन्ति।

३. कृत्रिमप्रज्ञायाः संस्कृतेन सह सम्बद्धताक्षेत्राणि

३.१ प्राकृतिकभाषासंसाधनम् (NLP)

संस्कृतस्य प्राकृतिकभाषासंसाधनं बहुविधानि आयामानि समाविशति। पदच्छेदः (Tokenization), रूपविज्ञानविश्लेषणम् (Morphological Analysis), वाक्यविन्यासविश्लेषणम् (Syntactic Parsing), अर्थविश्लेषणम् (Semantic Analysis) इत्यादयः – एतेषु सर्वेषु क्षेत्रेषु गतद्वयदशकेषु उल्लेखनीयं कार्यं कृतम् अस्ति। भारतीयप्रौद्योगिकीसंस्थानस्य (IIT) बनारसहिन्दूविश्वविद्यालयस्य च शोधकर्तारः संस्कृतपदच्छेदकयन्त्राणि, सन्धिविच्छेदकानि, समासविग्रहकानि च विकसितवन्तः।

तथापि संस्कृत-NLP-क्षेत्रे एका मौलिकी समस्या अस्ति – प्रशिक्षणसामग्र्याः न्यूनता। आधुनिकवृहद्भाषाप्रतिरूपाणि यथा GPT-श्रेणी, BERT-आदीनि सहस्रकोटिशब्दात्मकेषु आँकडासमुच्चयेषु प्रशिक्षितानि भवन्ति। संस्कृतस्य अङ्कीयकृतं (digitized) साहित्यं तुलनात्मकदृष्ट्या अल्पम् एव अस्ति, यद्यपि GRETIL, SARIT इत्यादिपरियोजनाभिः संस्कृतग्रन्थानाम् अङ्कीयकरणं प्रगतिशीलम् अस्ति।

३.२ यान्त्रिकानुवादः

संस्कृतात् अन्यभाषासु अन्यभाषाभ्यः संस्कृते च यान्त्रिकानुवादः एकं चुनौतीपूर्णं क्षेत्रम् अस्ति। संस्कृतस्य समासबाहुल्यम्, दीर्घसमासानां प्रचुरता, शास्त्रीयपरिभाषाणां बहुलता च अनुवादयन्त्राणां कृते विशेषसमस्यां जनयन्ति। उदाहरणार्थम् "सच्चिदानन्दस्वरूप" इत्यस्य पदस्य अनुवादे सत्+चित्+आनन्द+स्वरूप इति विग्रहः प्रथमं कर्तव्यः, ततः प्रत्येकस्य अवयवस्य अनुवादः कर्तव्यः, अन्ते च सम्पूर्णार्थस्य पुनर्निर्माणं कर्तव्यम्।

Google Translate, Microsoft Translator इत्यादिषु वाणिज्यिकानुवादसाधनेषु संस्कृतस्य समर्थनं सीमितम् अस्ति, गुणवत्ता च अन्यभाषाणां तुलनया न्यूना वर्तते। अस्य कारणद्वयम् अस्ति – प्रथमं समानान्तरसामग्र्याः (parallel corpora) अभावः, द्वितीयं च संस्कृतस्य भाषावैज्ञानिकजटिलता।

अनन्तपारं किल शब्दशास्त्रं स्वल्पं तथायुर्बहवश्च विघ्नाः ।
सारं ततो ग्राह्यमपास्य फल्गु हंसैर्यथा क्षीरमिवाम्बुमध्यात् ॥
– नीतिशतकम्, भर्तृहरिः

भर्तृहरेः अयं श्लोकः शब्दशास्त्रस्य अनन्तविस्तारं सूचयति। कृत्रिमप्रज्ञायाः क्षेत्रे अपि अयम् एव सिद्धान्तः प्रयोजनीयः – विपुलाद् आँकडासमुच्चयात् सारभूतां सूचनां ग्रहीतुं यत्नः कर्तव्यः, यथा हंसः जलमध्यात् क्षीरं पृथक् करोति।

३.३ ज्ञानप्रतिनिधित्वं तर्कश्च

कृत्रिमप्रज्ञायाः एकं प्राचीनतमं क्षेत्रं ज्ञानप्रतिनिधित्वम् (Knowledge Representation) अस्ति। अत्र संस्कृतपरम्परायाः योगदानं विशेषतया नव्यन्यायदर्शने दृश्यते। नव्यन्यायस्य परिभाषापद्धतिः एका औपचारिकभाषा (formal language) इव कार्यं करोति – अवच्छेदकत्वम्, निरूपकत्वम्, प्रतियोगित्वम् इत्यादयः नव्यन्यायपरिभाषाः तार्किकसम्बन्धानां सूक्ष्मवर्णने समर्थाः सन्ति।

Briggs (1985) इत्यस्य मतानुसारं संस्कृतस्य वाक्यसंरचना ज्ञानप्रतिनिधित्वस्य कृते उपयुक्ता भवति। किन्तु एषः दावः पूर्णतया अपरीक्षितः अस्ति इति न वक्तव्यम् – गतचत्वारिंशद्वर्षेषु अस्मिन् विषये बहवः शोधपत्राणि प्रकाशितानि, किन्तु व्यापकव्यावहारिकप्रयोगः अधुनापि दुर्लभः एव।

४. सम्भावनाः

४.१ संस्कृत आधारित ज्ञानकोश निर्माणम्

संस्कृतसाहित्ये विपुलं ज्ञानभण्डारं विद्यते – वैदिकसाहित्यम्, दार्शनिकग्रन्थाः, गणितशास्त्रम्, ज्योतिषशास्त्रम्, आयुर्वेदः, अर्थशास्त्रम् इत्यादयः। एतेषां ग्रन्थानाम् अङ्कीयकरणं तत्पश्चात् कृत्रिमप्रज्ञाआधारितविश्लेषणं च कृतं चेत्, एकस्य अद्वितीयज्ञानकोशस्य निर्माणं सम्भवति। Sanskrit Knowledge Graph इत्यादयः परियोजनाः अस्मिन् दिशि प्रयत्नशीलाः सन्ति।

४.२ शैक्षणिक प्रयोगाः

कृत्रिमप्रज्ञाआधारितशैक्षणिकसाधनानि संस्कृतशिक्षणे क्रान्तिकारिपरिवर्तनं कर्तुं समर्थानि भवन्ति। संवादात्मकशिक्षणप्रणाल्यः (Interactive Tutoring Systems) संस्कृतव्याकरणस्य शिक्षणं सरलीकर्तुं शक्नुवन्ति। पदच्छेदस्य सन्धिविच्छेदस्य च स्वचालितसाधनानि छात्राणां कृते अत्यन्तम् उपयोगीनि भवन्ति। Little Red Dot Foundation इत्यस्य Sanskrit Learning App तथा IIT Bombay इत्यस्य Sanskrit Heritage Site इत्यादीनि एतस्य दिशि प्रशंसनीयानि प्रयासानि सन्ति।

४.३ प्राचीन हस्तलिपीनां पुनरुद्धारः

भारतस्य पुस्तकालयेषु संग्रहालयेषु च सहस्रशः संस्कृतहस्तलिप्यः विद्यन्ते याः कालवशात् क्षीणाः जीर्णाः वा सन्ति। गहनशिक्षणआधारिताः अक्षरपहचानप्रणाल्यः (OCR) एतासां हस्तलिपीनां पठनं पुनरुद्धारं च कर्तुं समर्थाः भवन्ति। National Mission for Manuscripts (NMM) इत्यस्य कार्यम् अस्मिन् सन्दर्भे उल्लेखनीयम्।

५. सीमाः

५.१ आँकडासमुच्चयस्य न्यूनता

कृत्रिमप्रज्ञायाः आधुनिकतन्त्रज्ञानानि आँकडाधारितानि (data-driven) सन्ति। तानि विपुलेषु प्रशिक्षणसामग्रीषु निर्भराणि भवन्ति। संस्कृतस्य अङ्कीयकृतसामग्री अन्यभाषाणां तुलनया अत्यल्पा अस्ति। यत्र अङ्ग्रेजीभाषायाः कृते अब्जशद्वात्मकाः आँकडासमुच्चयाः उपलब्धाः सन्ति, तत्र संस्कृतस्य कृते कोटिशद्वात्मकाः अपि दुर्लभाः। एषा विषमता संस्कृतकृत्रिमप्रज्ञायाः प्रगतौ मूलभूतं बाधकम् अस्ति।

५.२ भाषा वैज्ञानिक जटिलता

संस्कृतस्य यत् वैशिष्ट्यं तस्य कृत्रिमप्रज्ञायां सम्भावनां जनयति, तदेव कदाचित् समस्यारूपम् अपि भवति। सन्धिप्रक्रिया पदसीमानां निर्धारणम् अत्यन्तं कठिनं करोति। एकस्य संस्कृतवाक्यस्य बहुविधाः पदच्छेदविकल्पाः सम्भवन्ति, येषां मध्ये सम्यक्विकल्पस्य चयनं सन्दर्भज्ञानम् अपेक्षते। समासानां विग्रहः अपि सर्वदा एकमेव विधानम् अनुसरति इति न कथयितुं शक्यते – प्रसङ्गः अर्थः च विग्रहे निर्णायकौ भवतः।

५.३ विशेषज्ञकर्मशक्तेः अभावः

संस्कृतस्य गणनात्मकसंसाधने द्वैतज्ञानम् आवश्यकम् – एकतः संस्कृतभाषायाः व्याकरणस्य साहित्यस्य च गहनं ज्ञानम्, अपरतः सङ्गणकविज्ञानस्य कृत्रिमप्रज्ञायाः च कौशलम्। उभयत्र निपुणाः विशेषज्ञाः विरलाः सन्ति। भारतस्य प्रमुखसंस्थासु यथा IIT, IIIT, IISER इत्यादिषु संस्कृतसङ्गणकविज्ञानस्य अन्तर्विषयकशोधकार्यक्रमाः सीमिताः एव सन्ति।

५.४ सांस्कृतिक सन्दर्भस्य गणनात्मक वर्णनस्य कठिनता

संस्कृतसाहित्यस्य बहवः अंशाः सांस्कृतिकदार्शनिकसन्दर्भेषु गहनतया निहिताः सन्ति। ध्वन्यालोकस्य व्यञ्जनावृत्तिः, काव्यशास्त्रस्य रसविधानम्, वेदान्तस्य अध्यारोप-अपवादपद्धतिः – एतानि सर्वाणि सांस्कृतिकसन्दर्भं विना पूर्णतया अवगन्तुं न शक्यन्ते। कृत्रिमप्रज्ञा प्रतिमानानि (patterns) अवगच्छति, किन्तु सांस्कृतिकसूक्ष्मताः अनुभवजन्यं ज्ञानं च तस्य सीमायाः परे वर्तन्ते।

शैले शैले न माणिक्यं मौक्तिकं न गजे गजे।
साधवो न हि सर्वत्र चन्दनं न वने वने ॥
– चाणक्यनीतिः

चाणक्यस्य अयं श्लोकः स्मारयति यत् प्रत्येकं स्थाने प्रत्येकं वस्तु न लभ्यते। तथैव कृत्रिमप्रज्ञायाः प्रत्येकक्षेत्रे समानसफलता न सम्भवति – संस्कृतस्य कानिचित् क्षेत्राणि कृत्रिमप्रज्ञायै अनुकूलानि सन्ति, कानिचित् तु तस्याः सीमायाः बहिः।

© Associationof Academic Researchers and Faculties (AARF)

६. वर्तमानशोधस्य स्थितिः

गतदशके संस्कृतगणनात्मकभाषाविज्ञाने उल्लेखनीया प्रगतिः दृश्यते। JNU-संस्थायाः संस्कृतगणनात्मकभाषाविज्ञानकेन्द्रम्, IIIT Hyderabad इत्यस्य भाषातन्त्रज्ञानशोधकेन्द्रम्, IIT Bombay इत्यस्य CFILT प्रयोगशाला इत्यादयः संस्थाः अस्मिन् क्षेत्रे सक्रियाः सन्ति। International Sanskrit Computational Linguistics Symposium इति सम्मेलनं प्रतिद्वितीयवर्षं सम्पद्यते यत्र विश्वभरतः शोधकर्तारः स्वकीयानि कार्याणि प्रस्तुवन्ति।

Transformer-आधारितानां प्रतिरूपाणां आगमनेन संस्कृत-NLP क्षेत्रे नवाः सम्भावनाः उद्भूताः। SanskritBERT, IndicBERT इत्यादीनि प्रतिरूपाणि संस्कृतपाठानां संसाधने प्रयुक्तानि। तथापि एतेषां प्रतिरूपाणां निष्पादनं (performance) अङ्ग्रेजीभाषायाः प्रतिरूपाणां निष्पादनस्य तुलनया न्यूनम् एव अस्ति, अस्य मुख्यं कारणं च आँकडान्यूनता एव।

७. उपसंहारः

संस्कृतभाषायाः कृत्रिमप्रज्ञायां प्रयोगस्य विषयः आशाजनकः अस्ति, किन्तु यथार्थवादिनो भवेम इति आवश्यकम्। संस्कृतं कृत्रिमप्रज्ञायाः कृते सर्वोत्तमा भाषा अस्ति इति दावः अतिशयोक्तिपूर्णः स्यात् – तथैव संस्कृतस्य गणनात्मकमहत्त्वं नगण्यम् इति अपि न्याय्यं न स्यात्।

वस्तुस्थितिः मध्ये कुत्रचित् अस्ति। पाणिनेः व्याकरणं गणनात्मकभाषाविज्ञानस्य इतिहासे एकं महत्त्वपूर्णं योगदानम् अस्ति। संस्कृतस्य विभक्तिव्यवस्था तार्किकसंरचना च NLP-शोधे मूल्यवती अन्तर्दृष्टिं प्रददाति। किन्तु व्यावहारिकप्रयोगस्य कृते पर्याप्ताः आँकडासमुच्चयाः, प्रशिक्षितकर्मशक्तिः, निरन्तरं वित्तपोषणं च आवश्यकानि।

भविष्ये यदि अन्तर्विषयकसहकार्यं वर्धते, यदि संस्कृतसामग्र्याः अङ्कीयकरणं गतिशीलतया प्रचलति, यदि च भारतसर्वकारस्य तथा अन्ताराष्ट्रियसंस्थानां समुचितं समर्थनं लभ्यते – तर्हि संस्कृतस्य कृत्रिमप्रज्ञाक्षेत्रे योगदानं निश्चयेन विस्तृततरं भविष्यति। अस्मिन् दिशि अग्रसरणं भवेत् इति शोधकर्तृणां दायित्वम् अस्ति।

सन्दर्भसूची (References)

1. ब्रिग्स, आर. (1985). "**संस्कृत और कृत्रिम बुद्धिमत्ता में ज्ञान का निरूपण**". एआई मैगज़ीन, 6(1), 22-38।
2. हुए, जी. (2009). "**संस्कृत संगणकीय भाषाविज्ञान**". कंप्यूटर विज्ञान व्याख्यान नोट्स, खंड 5402। स्प्रिंगर।
3. गोयल, पी., हुए, जी., कुलकर्णी, ए., शार्फ, पी., एवं बंकर, आर. (2012). "**संस्कृत प्रसंस्करण के लिए एक वितरित मंच**". कोलिंग 2012 की कार्यवाही, मुंबई।
4. कुलकर्णी, ए. (2013). "**संस्कृत के लिए डायनेमिक प्रोग्रामिंग आधारित निर्धारक निर्भरता पार्सर**". द्वितीय अंतर्राष्ट्रीय निर्भरता भाषाविज्ञान सम्मेलन (DepLing 2013) की कार्यवाही।

5. हेलविग, ओ. (2016). "**संस्कृत पाठों में वाक्य सीमाओं का पता लगाना**". कोलिंग 2016 की कार्यवाही, ओसाका, जापान, पृष्ठ 288-297।
6. कृष्णा, ए., सतुलुरी, पी., एवं गोयल, पी. (2017). "**संस्कृत शब्द विभाजन के लिए एक डेटासेट**". संयुक्त SIGHUM कार्यशाला, ACL 2017 की कार्यवाही।
7. अरालिकट्टे, आर., दे ल्होन्यू, एम., श्रीवास्तव, ए., एवं सोएगार्ड, ए. (2021). "**संस्कृत और संगणकीय भाषाविज्ञान**". एसोसिएशन फॉर कम्प्यूटेशनल लिंग्विस्टिक्स के लेनदेन, 9, 1312-1324।
8. संधान, जे., बेहरा, ओ., सामन्ता, एस., सोभा, एल., एवं एकबाल, ए. (2022). "**TransLIST: एक ट्रांसफॉर्मर-आधारित भाषावैज्ञानिक रूप से सूचित संस्कृत टोकनाइज़र→**. EMNLP 2022 की कार्यवाही।
9. ककवानी, डी., कुंचुकुट्टन, ए., गोल्ला, एस., आदि. (2020). "**IndicNLPSuite: भारतीय भाषाओं के लिए एकभाषीय कॉर्पस, मूल्यांकन मानक और पूर्व-प्रशिक्षित बहुभाषी भाषा मॉडल→**. EMNLP 2020 (Findings) की कार्यवाही।
10. तेरदलकर, ए., एवं भट्टाचार्य, ए. (2023). "**SanskritShala: वेब-आधारित इंटरफेस के साथ एक न्यूरल संस्कृत NLP टूलकिट**". ACL 2023 (सिस्टम डेमोंस्ट्रेशन) की क ार्यवाही, 240-250।
11. महेश्वरी, एच., जैन, एस., एवं मोरवाल, एस. (2024). "**रूपात्मक प्रसंस्करण के साथ संस्कृत-हिन्दी न्यूरल मशीन अनुवाद का संवर्धन→**. जर्नल ऑफ नैचुरल लैंग्वेज इंजीनियरिंग, 30(2), 345-368।
12. चॉम्स्की, एन. (1957). "**वाक्यविन्यास संरचनाएँ**". माउटन एंड कंपनी, द हेग।
13. राष्ट्रीय पांडुलिपि मिशन (2019). "**वार्षिक प्रतिवेदन 2018-2019**". संस्कृति मंत्रालय, भारत सरकार।