தமிழ் செயற்கை நுண்ணறிவு ஆய்விதழ்
மதிப்பாய்வு செய்யப்படும் இதழ்

Tamil Journal of Artificial Intelligence
Peer-Reviewed Journal

Volume 1, SPL 1 – July 2026

தமிழ் இலக்கியப் படைப்புகளை ஆவணப்படுத்துவதில் செயற்கை நுண்ணறிவுக் கருவிகளின் பங்கும் பயன்பாடுகளும்

நா. அபிநயா, உதவிப்பேராசிரியர், தமிழ்த்துறை (சுயநிதி), ஸ்ரீ எஸ்.இராமசாமி நாயுடு ஞாபகார்த்தக் கல்லூரி, சாத்தூர், தமிழ்நாடு, இந்தியா.
ORCiD: https://orcid.org/0009-0004-2920-1815

அறிமுகம்

தமிழ் இலக்கியம், உலகின் தொன்மையான மற்றும் வளமான இலக்கிய மரபுகளுள் ஒன்றாகும். சங்க இலக்கியம், காப்பியங்கள், பக்தி இலக்கியம், சிற்றிலக்கியங்கள், நவீன கவிதை, உரைநடை, நாடகம், மற்றும் சமகால இலக்கியம் எனப் பல ஆயிரம் ஆண்டுகளில் படைக்கப்பட்ட பல இலட்சம் படைப்புகள் தமிழில் உள்ளன. இந்த இலக்கியச் செல்வங்கள், தமிழ் மக்களின் பண்பாட்டு அடையாளம், வரலாறு, மற்றும் அறிவுத்திறனின் வெளிப்பாடுகளாகும். இருப்பினும், இந்தப் படைப்புகளில் பெரும்பாலானவை ஓலைச் சுவடிகள், கையெழுத்துப் பிரதிகள், மற்றும் அரிய அச்சு நூல்களாக மட்டுமே கிடைக்கின்றன. அவற்றைப் பாதுகாப்பதும், ஆவணப்படுத்துவதும், எதிர்காலத் தலைமுறையினருக்கு எடுத்துச் செல்வதும் தமிழ் இலக்கிய ஆய்வின் முக்கியமான பணிகளாகும். பாரம்பரியமாக, இலக்கியப் படைப்புகளை ஆவணப்படுத்துவது என்பது கையேடு உழைப்பு, நேரம் எடுக்கும் செயல்முறைகள், மற்றும் மனித ஆய்வாளர்களின் திறன் மற்றும் அனுபவத்தை மட்டும் சார்ந்திருந்தது. ஒரு ஓலைச் சுவடியைப் படியெடுத்தல், அதன் உரையைத் திருத்துதல், பல பிரதிகளை ஒப்பிடுதல், மற்றும் ஒரு தரவுத்தளத்தை உருவாக்குதல் ஆகிய பணிகள் ஆண்டுக்கணக்காக நீடிக்கும். இந்தச் சூழலில், செயற்கை நுண்ணறிவு (AI) தொழில்நுட்பங்கள், இந்த ஆவணப்படுத்தும் பணியை மிகவும் திறம்படவும், விரைவாகவும், துல்லியமாகவும் செய்யும் ஒரு புரட்சிகரமான கருவியாக அமைந்துள்ளன. AI-யின் ஒளியியல் எழுத்து அங்கீகாரம், கையெழுத்து அங்கீகாரம், இயற்கை மொழி செயலாக்கம், இயந்திரக் கற்றல், மற்றும் பெரும் தரவு மேலாண்மை ஆகிய திறன்கள், பழைய கையெழுத்துப் பிரதிகளை டிஜிட்டல் மயமாக்குவதிலும், அவற்றின் உரைகளைப் பகுப்பாய்வு செய்வதிலும், மற்றும் இலக்கியத் தரவுத்தளங்களை உருவாக்குவதிலும் மிகவும் பயனுள்ளதாக அமைகின்றன. இக்கட்டுரை, தமிழ் இலக்கியப் படைப்புகளை ஆவணப்படுத்துவதில் AI கருவிகளின் பங்கு, அவற்றின் பயன்பாடுகள், சவால்கள், மற்றும் எதிர்காலச் சாத்தியங்களை ஒரு முழுமையான கல்வியியல் ஆய்விற்கு உட்படுத்துகிறது.

தமிழ் இலக்கிய ஆவணப்படுத்தலின் தற்போதைய நிலையும் சவால்களும்

தமிழ் இலக்கியப் படைப்புகளை ஆவணப்படுத்துவதில் பல சவால்கள் உள்ளன. முதலாவதாக, ஓலைச் சுவடிகள் மற்றும் கையெழுத்துப் பிரதிகள் எண்ணற்றவை. அவை பல நூலகங்கள், அருங்காட்சியகங்கள், தனியார் சேகரிப்புகள், மற்றும் கோயில்களில் சிதறிக்கிடக்கின்றன. இவற்றை ஒருங்கிணைந்த தரவுத்தளமாக்குவது ஒரு பெரும் பணியாகும். இரண்டாவதாக, பல ஓலைச் சுவடிகள் பழுதடைந்த நிலையில் உள்ளன. அவற்றின் எழுத்துகள் மங்கியோ, கிழிந்தோ, அல்லது பூச்சிகளால் அழிந்தோ உள்ளன. இவற்றை மீட்டெடுப்பதும், படியெடுப்பதும் மிகவும் சிரமமானது. மூன்றாவதாக, தமிழ் எழுத்து முறை காலப்போக்கில் பல மாற்றங்களைச் சந்தித்துள்ளது. வட்டெழுத்து, கிரந்தம், பழைய தமிழ் எழுத்துருக்கள் என்று பல்வேறு வடிவங்களில் உள்ள எழுத்துகளை இனங்காண்பது ஒரு சிறப்புத் திறனைக் கோருகிறது. நான்காவதாக, ஒரு இலக்கியப் படைப்புக்கு ஒன்றுக்கு மேற்பட்ட பிரதிகள்

Volume 1, SPL 1 – July 2026

இருக்கும். ஒவ்வொரு பிரதியிலும் சில சொற்கள், வரிகள், அல்லது பகுதிகள் மாறுபடலாம். இந்தப் பிரதிகளை ஒப்பிட்டு, மூல உரையைக் கண்டறிவது (Textual Criticism) மிகவும் நேரம் எடுக்கும் ஒரு பணியாகும். ஐந்தாவதாக, இந்த ஆவணங்களைத் தேடவும், பகுப்பாய்வு செய்யவும், மற்றும் பயன்படுத்தவும் ஏதுவான ஒரு முழுமையான டிஜிட்டல் தரவுத்தளம் இல்லை. இந்தச் சவால்களைச் சமாளிக்க AI கருவிகள் மிகவும் பயனுள்ளவையாக அமைகின்றன.

ஒளியியல் எழுத்து அங்கீகாரம் (OCR) - ஓலைச் சுவடிகளின் டிஜிட்டல் மயமாக்கல்

தமிழ் இலக்கிய ஆவணப்படுத்தலில் AI-யின் மிக முக்கியமான பயன்பாடு, ஒளியியல் எழுத்து அங்கீகாரம் (Optical Character Recognition - OCR) தொழில்நுட்பமாகும். OCR தொழில்நுட்பம், ஒரு காகிதம் அல்லது ஓலைச் சுவடியில் உள்ள எழுத்துகளை ஒரு படமாகப் படம் பிடித்து, அவற்றைக் கணினி உரையாக (Computer-readable Text) மாற்றுகிறது. தமிழில், குறிப்பாக வட்டெழுத்து, கிரந்தம், மற்றும் பழைய தமிழ் எழுத்துருக்களை இனங்காணும் OCR கருவிகள் உருவாக்கப்பட்டுள்ளன. இந்த AI கருவிகள், ஒரு ஓலைச் சுவடியின் படத்தை உள்ளீடாகப் பெற்று, அதன் ஒவ்வொரு எழுத்தையும் இனங்கண்டு, அவற்றைத் தமிழ் யுனிகோட் எழுத்துகளாக மாற்றுகின்றன. இந்த OCR தொழில்நுட்பத்தின் சிறப்பு, அது பழுதடைந்த, மங்கிய, அல்லது கிழிந்த எழுத்துகளையும் இனங்காணும் திறன் கொண்டது. AI-யின் இயந்திரக் கற்றல் வழிமுறைகள், பல ஆயிரம் ஓலைச் சுவடிகளின் படங்களில் பயிற்சியளிக்கப்பட்டு, பல்வேறு வகையான எழுத்து முறைகள் மற்றும் கையெழுத்து நடைகளைக் கற்றுக்கொள்கின்றன. இதனால், ஒரு பழைய ஓலைச் சுவடியில் உள்ள ஒரு மங்கிய எழுத்தை, அதன் வடிவம், சூழல், மற்றும் பிற எழுத்துகளுடனான தொடர்பை வைத்து AI துல்லியமாக இனங்காண முடிகிறது. இது, கையேடு முறையில் படியெடுக்க எடுக்கும் பல மாதங்கள் அல்லது வருடங்கள் நேரத்தை சில நிமிடங்களுக்கு அல்லது சில மணிநேரங்களுக்குக் குறைக்கிறது.

கையெழுத்து அங்கீகாரம் (HWR) வாயிலான பிரதி படியெடுத்தல்

ஓலைச் சுவடிகளுக்கு அடுத்தபடியாக, பல தமிழ் இலக்கியப் படைப்புகள் கையெழுத்துப் பிரதிகளாகவே கிடைக்கின்றன. இவை பத்தொன்பதாம் மற்றும் இருபதாம் நூற்றாண்டுகளில் எழுதப்பட்டவை. இந்தக் கையெழுத்துப் பிரதிகளின் எழுத்து நடை, ஒவ்வொரு எழுத்தாளருக்கும் வேறுபடும். AI-யின் கையெழுத்து அங்கீகாரம் (Handwriting Recognition - HWR) தொழில்நுட்பம், இந்தக் கையெழுத்துப் பிரதிகளைப் படித்து, அவற்றை டிஜிட்டல் உரையாக மாற்றுகிறது. HWR கருவிகள், கையெழுத்தின் வேகம், சாய்வு, மற்றும் எழுத்துகளுக்கிடையேயான இடைவெளி ஆகியவற்றைப் பகுப்பாய்வு செய்து, ஒவ்வொரு எழுத்தாளரின் தனித்துவமான எழுத்து நடையைக் கற்றுக்கொள்கின்றன. AI HWR கருவிகள், இந்தப் பிரதிகளை விரைவாகப் படியெடுத்து, அவற்றை ஆய்வுக்கு உட்படுத்த உதவுகின்றன.

இயற்கை மொழி செயலாக்கம் (NLP) வாயிலான உரைத் திருத்தம் மற்றும் பதிப்பு ஒப்பீடு

AI-யின் இயற்கை மொழி செயலாக்கம் (NLP) திறன், ஒரு டிஜிட்டல் உரையில் உள்ள இலக்கணப் பிழைகள், எழுத்துப் பிழைகள், மற்றும் சொற்பொருள் முரண்பாடுகளைத் தானாகவே கண்டறிந்து திருத்தும் திறன் கொண்டது. இது, ஓலைச் சுவடிகள் அல்லது கையெழுத்துப் பிரதிகளில் இருந்து OCR அல்லது HWR மூலம் படியெடுக்கப்பட்ட உரைகளில் உள்ள பிழைகளைத் திருத்துவதற்கு மிகவும் பயனுள்ளதாகும். NLP கருவிகள், ஒரு வாக்கியத்தின் இலக்கணக் கட்டமைப்பைப் பகுப்பாய்வு செய்து, அதில் உள்ள தவறான சொற்கள் அல்லது இலக்கணப் பிழைகளை இனங்கண்டு, அவற்றிற்கான திருத்தங்களைப் பரிந்துரைக்கின்றன. மேலும், NLP கருவிகள், ஒரு இலக்கியப் படைப்பின் பல்வேறு பிரதிகளை

தமிழ் செயற்கை நுண்ணறிவு ஆய்விதழ்
மதிப்பாய்வு செய்யப்படும் இதழ்

Tamil Journal of Artificial Intelligence
Peer-Reviewed Journal

Volume 1, SPL 1 – July 2026

ஒப்பிட்டு ஆய்வு செய்யும் பணியிலும் (Collation) பயன்படுகின்றன. ஒரு காப்பியத்திற்கு அல்லது ஒரு கவிதைத் தொகுப்பிற்கு பத்து அல்லது இருபது வெவ்வேறு பிரதிகள் இருக்கலாம். ஒவ்வொரு பிரதியிலும் சில சொற்கள், வரிகள், அல்லது பகுதிகள் மாறுபடலாம். AI NLP கருவிகள், இந்தப் பிரதிகளை ஒரே நேரத்தில் ஒப்பிட்டு, ஒவ்வொரு வரியிலும் ஏற்படும் மாறுபாடுகளைத் துல்லியமாகப் பட்டியலிடும். இது, ஆய்வாளர்கள் மூல உரை (Original Text) எது என்பதைக் கண்டறிவதற்கு மிகவும் உதவியாக அமைகிறது. இந்தப் பணி, பாரம்பரியமாக ஆய்வாளர்கள் கையால் செய்யும் போது பல மாதங்கள் எடுக்கும்; AI கருவிகள் அதை சில நிமிடங்களில் செய்துவிடுகின்றன.

இயந்திரக் கற்றல் வாயிலான பகுப்பாய்வு மற்றும் காலம் நிர்ணயம்

AI-யின் இயந்திரக் கற்றல் வழிமுறைகள், ஒரு இலக்கியப் படைப்பின் ஆசிரியர் யார், அது எந்தக் காலத்தில் எழுதப்பட்டது, மற்றும் அது ஒரே ஆசிரியரால் எழுதப்பட்டதா அல்லது பலரால் எழுதப்பட்டதா என்பதைப் பகுப்பாய்வு செய்யும் திறன் கொண்டவை. இது, ஆசிரியர் பாணி பகுப்பாய்வு (Stylometric Analysis) என்று அழைக்கப்படுகிறது. AI கருவிகள், ஒரு படைப்பின் சொல்லாடல், வாக்கிய அமைப்பு, சொற்களின் சராசரி நீளம், மற்றும் சில குறிப்பிட்ட சொற்கள் அல்லது சொற்றொடர்களின் பயன்பாட்டு அதிர்வெண் ஆகியவற்றைப் பகுப்பாய்வு செய்கின்றன. ஒவ்வொரு ஆசிரியருக்கும் இந்தப் பண்புகள் தனித்துவமானவை. AI கருவிகள், இந்தத் தரவுகளை ஒரு அறியப்பட்ட ஆசிரியரின் படைப்புகளுடன் ஒப்பிட்டு, அந்தப் படைப்பின் ஆசிரியர் யார் என்பதைத் தீர்மானிக்கின்றன. தமிழ் இலக்கியத்தில், பல பண்டைய நூல்களின் ஆசிரியர் யார் என்பது குறித்து அறிஞர்களிடையே கருத்து வேறுபாடுகள் உள்ளன. எடுத்துக்காட்டாக, திருக்குறளின் ஆசிரியர் ஒருவரா அல்லது பலரா? சில சங்க இலக்கியப் பாடல்களின் ஆசிரியர் யார்? இந்தக் கேள்விகளுக்கு AI-யின் ஆசிரியர் பாணி பகுப்பாய்வு ஒரு புதிய மற்றும் துல்லியமான பதிலை வழங்கும். மேலும், AI கருவிகள், ஒரு படைப்பின் மொழியியல் பண்புகளைப் பகுப்பாய்வு செய்து, அது எந்தக் காலத்தைச் சேர்ந்தது என்பதையும் மதிப்பிடும். இது, தமிழ் இலக்கியத்தின் காலவரிசையைத் துல்லியமாக வரையறுக்க உதவுகிறது.

பெரும் தரவு மேலாண்மை வாயிலான இலக்கியத் தரவுத்தளங்கள்

AI-யின் பெரும் தரவு மேலாண்மை (Big Data Management) திறன், ஆயிரக்கணக்கான தமிழ் இலக்கியப் படைப்புகளை ஒரு ஒருங்கிணைந்த டிஜிட்டல் தரவுத்தளத்தில் சேமிக்கவும், அவற்றை எளிதில் தேடவும், பகுப்பாய்வு செய்யவும் உதவுகிறது. ஒரு முழுமையான தமிழ் இலக்கியத் தரவுத்தளம், பின்வரும் அம்சங்களை உள்ளடக்கியதாக இருக்கும்: ஒவ்வொரு படைப்பின் முழு உரை, அதன் ஆசிரியர், காலம், வகை, முக்கிய சொற்கள், கருப்பொருள்கள், மற்றும் பிரதிகளுக்கிடையேயான மாறுபாடுகள். AI கருவிகள், இந்தத் தரவுத்தளத்தில் உள்ள தரவுகளைத் தானாகவே வகைப்படுத்தி, அவற்றுக்கிடையேயான தொடர்புகளைக் கண்டறியும். எடுத்துக்காட்டாக, ஒரு ஆய்வாளர் "காதல்" என்ற கருப்பொருளில் எழுதப்பட்ட அனைத்து சங்க இலக்கியப் பாடல்களையும் தேட விரும்பினால், AI கருவிகள் சில நொடிகளில் அந்தப் பாடல்களைத் தேர்ந்தெடுத்து வழங்கும். மேலும், AI கருவிகள், இந்தத் தரவுத்தளத்தைப் பயன்படுத்தி, தமிழ் இலக்கியத்தின் காலப்போக்கில் ஏற்பட்ட கருப்பொருள் மாற்றங்கள், மொழியியல் மாற்றங்கள், மற்றும் பாணி மாற்றங்களைப் பகுப்பாய்வு செய்யும். இது, தமிழ் இலக்கிய ஆய்வுக்கு ஒரு புதிய பரிமாணத்தைத் திறக்கிறது.

AI கருவிகளின் பயன்பாட்டில் உள்ள நெறிமுறைச் சிக்கல்கள்

AI கருவிகள் தமிழ் இலக்கிய ஆவணப்படுத்தலில் மிகவும் பயனுள்ளவையாக இருந்தாலும், அவற்றின் பயன்பாட்டில் பல நெறிமுறைச் சிக்கல்களைக் கவனத்தில் கொள்ள

Volume 1, SPL 1 – July 2026

வேண்டும். முதலாவதாக, தரவுத் தனியுரிமை (Data Privacy) மற்றும் படைப்புரிமை (Copyright). பல தமிழ் இலக்கியப் படைப்புகள், குறிப்பாக சமகால இலக்கியம், படைப்புரிமைச் சட்டங்களால் பாதுகாக்கப்படுகின்றன. AI கருவிகள், இந்தப் படைப்புகளை டிஜிட்டல் மயமாக்கி, அவற்றைத் தரவுத்தளத்தில் சேமிக்கும் போது, படைப்புரிமை மீறல் ஏற்படாமல் இருக்க, உரிமையாளர்களின் அனுமதி பெறுவது அவசியம். மேலும், பழைய ஓலைச் சுவடிகள் மற்றும் கையெழுத்துப் பிரதிகள் பல நூலகங்கள் மற்றும் தனியார் சேகரிப்புகளில் உள்ளன. அவற்றை டிஜிட்டல் மயமாக்கும் போது, அந்த நிறுவனங்கள் மற்றும் தனிநபர்களின் ஒத்துழைப்பு மற்றும் அனுமதி பெறப்பட வேண்டும். இரண்டாவதாக, AI கருவிகளின் தரவுச் சார்பு (Data Bias). AI கருவிகள், அவை பயிற்சியளிக்கப்பட்ட தரவுகளில் உள்ள சார்புகளைப் பெற்றிருக்கும். தமிழ் இலக்கிய ஆவணப்படுத்தலில், AI கருவிகள் கிடைக்கும் தரவுகளில் பயிற்சியளிக்கப்படுகின்றன. கிடைக்கும் தரவுகள் பெரும்பாலும் புகழ்பெற்ற, பெரிய இலக்கியப் படைப்புகளை மையமாகக் கொண்டவை. இதனால், சிறிய, கிராமப்புற, அல்லது புறக்கணிக்கப்பட்ட இலக்கியப் படைப்புகள் AI கருவிகளால் கவனிக்கப்படாமல் போகலாம். இந்தச் சார்பைத் தவிர்க்க, AI கருவிகள் பல்வேறு வகையான இலக்கியப் படைப்புகள், பிராந்தியங்கள், மற்றும் காலங்களில் இருந்து சேகரிக்கப்பட்ட பன்முகத் தரவுகளில் பயிற்சியளிக்கப்பட வேண்டும். மூன்றாவதாக, AI கருவிகளின் வரம்புகள் மற்றும் மனிதத் தலையீட்டின் அவசியம். AI கருவிகள் ஒரு ஓலைச் சுவடியின் எழுத்துகளை இனங்காணவும், ஒரு படைப்பின் பிரதிகளை ஒப்பிடவும் முடியும்; ஆனால், அவை ஒரு இலக்கியப் படைப்பின் அழகியல், உணர்வுநயம், மற்றும் பண்பாட்டு முக்கியத்துவத்தை முழுமையாகப் புரிந்துகொள்ள முடியாது. இலக்கிய ஆவணப்படுத்தல் என்பது வெறும் உரையைப் படியெடுப்பது மட்டுமல்ல; அது அந்த உரையின் விளக்கம், வர்ணனை, மற்றும் கல்வியியல் மதிப்பீடு ஆகியவற்றையும் உள்ளடக்கியது. இந்தப் பணிகளுக்கு மனித ஆய்வாளர்களின் திறன், அனுபவம், மற்றும் உள்ளுணர்வு மிகவும் அவசியமானவை. எனவே, AI கருவிகள் ஒரு துணைக் கருவியாகவே பயன்படுத்தப்பட வேண்டும்; மனித ஆய்வாளர்களுக்கு மாற்றாக அல்ல. நான்காவதாக, தொழில்நுட்பச் செலவு மற்றும் அணுகல். AI கருவிகளைச் செயல்படுத்துவதற்கு அதிக விலையுயர்ந்த கணினி உள்கட்டமைப்பு, மென்பொருட்கள், மற்றும் பயிற்சி பெற்ற ஊழியர்கள் தேவைப்படுகின்றன. பல சிறிய நூலகங்கள், அருங்காட்சியகங்கள், மற்றும் ஆய்வு நிறுவனங்கள் இந்தச் செலவுகளைச் சமாளிப்பதில் சிரமப்படுகின்றன. எனவே, AI கருவிகள் மலிவு விலையில் கிடைக்கும்படி செய்வதும், அவற்றைப் பயன்படுத்துவதற்கான பயிற்சிகளை வழங்குவதும் அவசியமான ஒரு கல்வியியல் தேவையாகும்.

AI கருவிகளின் பயன்பாட்டிற்கான ஒரு முன்மொழியப்பட்ட கட்டமைப்பு

தமிழ் இலக்கியப் படைப்புகளை ஆவணப்படுத்துவதற்கு AI கருவிகளைத் திறம்படப் பயன்படுத்த, பின்வரும் படிநிலைக் கட்டமைப்பைப் பின்பற்றலாம். முதலாவதாக, ஒரு தேசிய அளவிலான தமிழ் இலக்கிய டிஜிட்டல் மயமாக்கல் திட்டத்தை உருவாக்குதல். இந்தத் திட்டத்தின் மூலம், நாடு முழுவதும் உள்ள நூலகங்கள், அருங்காட்சியகங்கள், மற்றும் தனியார் சேகரிப்புகளில் உள்ள அனைத்து ஓலைச் சுவடிகள், கையெழுத்துப் பிரதிகள், மற்றும் அரிய அச்சு நூல்களை டிஜிட்டல் மயமாக்குவதற்கு ஒரு விரிவான திட்டத்தை உருவாக்க வேண்டும். இதற்கு AI OCR, HWR, மற்றும் படமாக்கும் தொழில்நுட்பங்கள் பயன்படுத்தப்பட வேண்டும். இரண்டாவதாக, ஒரு பன்முகத் தரவுத்தளத்தை உருவாக்குதல். டிஜிட்டல் மயமாக்கப்பட்ட அனைத்து இலக்கியப் படைப்புகளையும் ஒரு ஒருங்கிணைந்த, தேடக்கூடிய, மற்றும் பகுப்பாய்வு செய்யக்கூடிய தரவுத்தளத்தில் சேமிக்க வேண்டும். இந்தத் தரவுத்தளம், உரை, படம், ஒலி, மற்றும் காணொலி ஆகிய அனைத்து வடிவங்களையும் உள்ளடக்கியதாக இருக்க

தமிழ் செயற்கை நுண்ணறிவு ஆய்விதழ்
மதிப்பாய்வு செய்யப்படும் இதழ்
5
Tamil Journal of Artificial Intelligence
Peer-Reviewed Journal

Volume 1, SPL 1 – July 2026

வேண்டும். மேலும், இந்தத் தரவுத்தளத்தை ஆய்வாளர்கள், மாணவர்கள், மற்றும் பொதுமக்கள் இலவசமாக அணுகும் வகையில் உருவாக்க வேண்டும். மூன்றாவதாக, AI கருவிகளின் துல்லியத்தை மேம்படுத்துவதற்கான தொடர் பயிற்சித் திட்டங்கள். AI கருவிகள், புதிய தரவுகளின் அடிப்படையில் தொடர்ந்து பயிற்சியளிக்கப்பட வேண்டும். குறிப்பாக, பல்வேறு வட்டாரங்கள், காலங்கள், மற்றும் கையெழுத்து நடைகளில் இருந்து சேகரிக்கப்பட்ட பன்முகத் தரவுகள் AI-க்கு வழங்கப்பட வேண்டும். மேலும், AI கருவிகளின் முடிவுகளை மனித ஆய்வாளர்கள் தொடர்ந்து மதிப்பீடு செய்து, அவற்றை மேம்படுத்துவதற்கான பின்னூட்டங்களை வழங்க வேண்டும். நான்காவதாக, AI கருவிகளின் பயன்பாட்டிற்கான நெறிமுறைக் கட்டமைப்பு மற்றும் வழிகாட்டுதல்களை உருவாக்குதல். இது படைப்புரிமை, தரவுத் தனியுரிமை, பண்பாட்டு உணர்வு, மற்றும் AI-யின் சார்புத் தன்மை ஆகியவற்றைக் கட்டுப்படுத்தும் வழிகாட்டுதல்களை உள்ளடக்கியதாக இருக்க வேண்டும். இந்த வழிகாட்டுதல்களைப் பின்பற்றி, AI கருவிகள் நெறிமுறையான முறையில் பயன்படுத்தப்படுவதை உறுதி செய்ய வேண்டும். ஐந்தாவதாக, AI கருவிகளைப் பயன்படுத்துவதற்கான பயிற்சித் திட்டங்களை உருவாக்குதல். இலக்கிய ஆய்வாளர்கள், நூலகர்கள், மற்றும் மாணவர்களுக்கு AI கருவிகளைப் பயன்படுத்துவதற்கான தொழில்நுட்பப் பயிற்சிகளை வழங்க வேண்டும். இது, AI கருவிகளைத் திறம்படப் பயன்படுத்துவதற்கு உதவுவதோடு, அவற்றின் முடிவுகளைச் சரியாக விளக்குவதற்கும் உதவும்.

AI கருவிகளின் எதிர்காலம்

தமிழ் இலக்கிய ஆவணப்படுத்தலில் AI-யின் எதிர்காலம் மிகவும் நம்பிக்கைக்குரியது. மேம்பட்ட AI மாதிரிகள், பழைய ஓலைச் சுவடிகளில் உள்ள மங்கிய, கிழிந்த, அல்லது அழிந்த எழுத்துகளை மீட்டெடுக்கும் (Restoration) திறனைப் பெறும். எடுத்துக்காட்டாக, AI கருவிகள், ஒரு சுவடியின் ஒரு பகுதி முழுவதும் அழிந்துவிட்டால், அந்தச் சுவடியின் மற்ற பகுதிகளிலும், அதே காலத்தின் பிற இலக்கியப் படைப்புகளிலும் உள்ள சொற்கள் மற்றும் வாக்கிய அமைப்புகளைப் பகுப்பாய்வு செய்து, அழிந்த பகுதியில் என்ன எழுதப்பட்டிருக்கலாம் என்பதை மிகவும் துல்லியமாக முன்கணிக்கும். மேலும், AI கருவிகள், தமிழ் இலக்கியப் படைப்புகளை மொழிபெயர்ப்பதிலும் பயன்படும். பழைய தமிழ் இலக்கியங்களை நவீன தமிழிலும், ஆங்கிலத்திலும் மொழிபெயர்க்க AI உதவும். இது, தமிழ் இலக்கியத்தை உலகளாவிய அளவில் அறியச் செய்யும் ஒரு சக்திவாய்ந்த கருவியாக அமையும். எதிர்காலத்தில், AI கருவிகள் ஒரு இலக்கியப் படைப்பின் உணர்வுநயம், அழகியல், மற்றும் பண்பாட்டு அடுக்குகளைப் புரிந்துகொண்டு, அவற்றை மொழிபெயர்ப்பில் பாதுகாக்கும் திறனைப் பெறும். கூடுதலாக, AI கருவிகள், தமிழ் இலக்கியத் தரவுத்தளங்களைப் பயன்படுத்தி, இலக்கியப் படைப்புகளுக்கிடையேயான மறைந்த தொடர்புகளைக் கண்டறியும். உதாரணமாக, ஒரு கவிஞரின் கவிதையில் உள்ள ஒரு உவமை, மற்றொரு கவிஞரின் கவிதையில் உள்ள உவமையுடன் எவ்வாறு தொடர்புடையது என்பதை AI கண்டறியும். இது, தமிழ் இலக்கிய ஆய்வுக்கு ஒரு புதிய பரிமாணத்தைத் திறக்கும்.

முடிவுரை

தமிழ் இலக்கியப் படைப்புகளை ஆவணப்படுத்துவது, தமிழ் மொழி மற்றும் பண்பாட்டின் வரலாற்றைப் பாதுகாப்பதற்கான ஒரு முக்கியமான கல்வியியல் பணியாகும். பாரம்பரிய முறைகள், கையேடு உழைப்பு, நேரம், மற்றும் மனிதத் திறனை மட்டும் சார்ந்திருந்தன. இருப்பினும், செயற்கை நுண்ணறிவு தொழில்நுட்பங்கள், இந்தப் பணியை மிகவும் திறம்படவும், விரைவாகவும், துல்லியமாகவும் செய்யும் ஒரு புரட்சிகரமான மாற்றத்தை ஏற்படுத்தியுள்ளன. AI-யின் ஒளியியல் எழுத்து அங்கீகாரம், கையெழுத்து அங்கீகாரம்,

Volume 1, SPL 1 – July 2026

இயற்கை மொழி செயலாக்கம், இயந்திரக் கற்றல், மற்றும் பெரும் தரவு மேலாண்மை ஆகிய திறன்கள், பழைய ஓலைச் சுவடிகள் மற்றும் கையெழுத்துப் பிரதிகளை டிஜிட்டல் மயமாக்குவதிலும், அவற்றின் உரைகளைத் திருத்துவதிலும், பதிப்புகளை ஒப்பிட்டு ஆய்வு செய்வதிலும், ஆசிரியர் பாணி பகுப்பாய்வு செய்வதிலும், மற்றும் இலக்கியத் தரவுத்தளங்களை உருவாக்குவதிலும் மிகவும் பயனுள்ளவையாக அமைகின்றன.

இருப்பினும், AI-யின் பயன்பாட்டில் தரவுத் தனியுரிமை, படைப்புரிமை, தரவுச் சார்பு, மற்றும் AI-யின் வரம்புகள் போன்ற நெறிமுறைச் சிக்கல்களைக் கவனத்தில் கொள்ள வேண்டும். AI என்பது ஒரு துணைக் கருவியாகவே அமைய வேண்டும்; மனித ஆய்வாளர்களின் அறிவு, அனுபவம், மற்றும் உள்ளுணர்வுக்கு மாற்றாக அல்ல. AI கருவிகள், மனித ஆய்வாளர்களின் பணிச்சுமையைக் குறைத்து, அவர்கள் மிகவும் ஆழமான, படைப்பாற்றல் நிறைந்த இலக்கிய ஆய்வுகளில் கவனம் செலுத்த உதவ வேண்டும். சரியான நெறிமுறைக் கட்டமைப்பு, பன்முகத் தரவுகள், மனிதர்-இன்-தி-லூப் அமைப்பு, மற்றும் தொடர் மேம்பாட்டு உத்திகளுடன் AI-யை ஒருங்கிணைப்பதன் மூலம், தமிழ் இலக்கிய ஆவணப்படுத்தல் ஒரு புதிய உயரத்திற்கு எடுத்துச் செல்லப்படும். AI-துணையுடனான தமிழ் இலக்கிய ஆவணப்படுத்தல், தமிழ் இலக்கியத்தின் பாரம்பரியத்தைப் பாதுகாப்பதோடு, அதனை உலகளாவிய அளவில் அறியச் செய்யும் ஒரு சக்திவாய்ந்த கல்வியியல் கருவியாக அமையும்.

குறிப்புகள்

[1] அண்ணாமலை, சு. (2022). தமிழ் இலக்கிய டிஜிட்டல் மயமாக்கல்: ஒரு அறிமுகம். சென்னை: தமிழ்ப் பல்கலைக்கழக வெளியீட்டகம்.

[2] இராமசாமி, நா. (2021). இலக்கிய ஆவணப்படுத்தலில் AI-யின் பங்கு. கோவை: நியூ செஞ்சுரி புத்தக நிலையம்.

[3] தங்கராஜ், மு. (2024). இலக்கிய ஆசிரியர் பாணி பகுப்பாய்வில் இயந்திரக் கற்றல். இந்திய கணினி மொழியியல் ஆய்வுகள், 17(1), 45-62.

[4] பாலகிருஷ்ணன், ரா. (2021). தமிழ் இலக்கியத் தரவுத்தளங்களின் வடிவமைப்பும் மேலாண்மையும். கோவை: கணினி மொழி வெளியீடு.

நிதிசார் கட்டுரையாளர் உறுதிமொழி: இல்லை

கட்டுரையாளர் நன்றியுரை: இல்லை

கட்டுரையாளர் உறுதிமொழி: இக்கட்டுரையில் எவ்வித முரண்பாடும் இல்லை என்று உறுதிமொழி அளிக்கிறேன்.

இக்கட்டுரை கிரியேட்டிவ் காமன்சு ஆட்ரிபியூசன் 4.0வின் **Creative Commons Attribution4.0** கீழ் பன்னாட்டு உரிமம் பெற்றுள்ளது.