தமிழ் செயற்கை நுண்ணறிவு ஆய்விதழ்
மதிப்பாய்வு செய்யப்படும் இதழ்

Tamil Journal of Artificial Intelligence
Peer-Reviewed Journal

Volume 1, Issue 2 – July 2026

பல்மாதிரி செயற்கை நுண்ணறிவு (Multimodal AI) எழுத்தைத் தாண்டி காட்சி, ஒலி, காணொளியைப் புரிந்துகொள்ளும் AI

முனைவர் கி. நாகேந்திரன், உதவிப்பேராசிரியர், தமிழ்த்துறை, ஸ்ரீ எஸ்.இராமசாமி நாயுடு ஞாபகார்த்தக் கல்லூரி, சாத்தூர், விருதுநகர் மாவட்டம், தமிழ்நாடு, இந்தியா.
ORCiD: https://orcid.org/0000-0003-2865-7832

அறிமுகம்

மனித அறிவு ஒரே ஒரு தகவல் வடிவத்தின் அடிப்படையில் இயங்குவதில்லை. நாம் ஒரு நிகழ்வைப் புரிந்துகொள்ளும்போது மொழி, காட்சி, ஒலி, உடல் இயக்கம், இடவியல் மற்றும் சூழல் ஆகியவற்றை ஒருங்கிணைக்கிறோம். செயற்கை நுண்ணறிவும் இதேபோன்ற ஒருங்கிணைந்த புரிதலை நோக்கி நகர்கிறது. இதன் முக்கிய வடிவமே பல்மாதிரி செயற்கை நுண்ணறிவு (Multimodal AI). உரை, படம், ஒலி, காணொளி மற்றும் பிற தரவு வடிவங்களை ஒருங்கிணைத்து புரிந்துகொள்ளவும் பதிலளிக்கவும் கூடிய அமைப்புகள் தற்போது AI ஆராய்ச்சியின் முக்கியப் பகுதியாக உள்ளன.

பல்மாதிரி AI காட்சி, ஒலி, காணொளியைப் புரிந்துகொள்ளும் AI

முதற்கட்ட செயற்கை நுண்ணறிவு அமைப்புகள் பெரும்பாலும் குறிப்பிட்ட தரவு வடிவங்களை மையமாகக் கொண்டிருந்தன. உரையைப் புரிந்துகொள்ளும் மொழி மாதிரி, படங்களை அடையாளம் காணும் பார்வை மாதிரி, ஒலியை எழுத்தாக மாற்றும் பேச்சு மாதிரி எனத் தனித்தனி அமைப்புகள் உருவாக்கப்பட்டன. பல்மாதிரி AI (Multimodal AI) இந்தப் பிரிவுகளை ஒன்றிணைக்கிறது. ஒரு மாணவர் கணிதப் பிரச்சினையின் புகைப்படத்தை வழங்கி, அதனைப் படிப்படியாக விளக்குமாறு கேட்கலாம். ஒரு மருத்துவர் மருத்துவப் படத்துடன் நோயாளியின் அறிகுறிகளை இணைத்து பகுப்பாய்வு செய்யும் அமைப்பைப் பயன்படுத்தலாம். ஒரு ஆராய்ச்சியாளர் காணொளி, உரை மற்றும் தரவுத்தொகுப்பை ஒரே பணிச்சூழலில் ஆராயலாம். பல்மாதிரி AI-யின் (Multimodal AI) முக்கிய முன்னேற்றம் தகவல்களை ஒன்றாகப் பார்ப்பதில் மட்டுமல்ல; அவற்றுக்கிடையேயான தொடர்பைப் புரிந்துகொள்வதிலும் உள்ளது.

ஒரு காணொளியில் ஒருவர் பேசும் வார்த்தைகள், அவரது செயல், சுற்றுப்புறச் சூழல் மற்றும் காட்சியில் தோன்றும் பொருட்கள் ஆகியவற்றை இணைத்து பொருள் உருவாக்குவது இதன் மேம்பட்ட வடிவமாகும். இத்தகைய திறன் கல்வியில் பெரும் மாற்றத்தை ஏற்படுத்தலாம். பார்வைத் திறன், கேள்வித் திறன் அல்லது வாசிப்புத் திறன் ஆகியவற்றில் வேறுபாடு கொண்ட மாணவர்களுக்குப் பல்வேறு வடிவங்களில் ஒரே கருத்தை வழங்க முடியும். ஒரு இலக்கியப் பாடத்தில் கவிதையின் உரையையும் அதன் ஒலிப்பதிவையும் காட்சிப்படுத்தலையும் இணைக்கலாம். வரலாற்றுப் பாடத்தில் பழைய ஆவணம், வரைபடம், புகைப்படம் மற்றும் வாய்மொழி விளக்கம் ஆகியவற்றை ஒருங்கிணைக்கலாம். இதனால் கற்றல் ஒரு உரைநூல் மைய அனுபவத்திலிருந்து பல உணர்வுகளை உள்ளடக்கிய அனுபவமாக மாறக்கூடும்.

ஆய்வுத் துறையிலும் இதன் பயன்பாடு விரிவடைந்து வருகிறது. புவி அறிவியலில் செயற்கைக்கோள் படங்கள், வானிலைத் தரவுகள், நிலவியல் அளவீடுகள் மற்றும் அறிவியல் உரைகள் ஆகியவற்றை இணைத்து பகுப்பாய்வு செய்யும் அடித்தள மாதிரிகள் உருவாகின்றன. 2026ஆம் ஆண்டில் வெளியான ஆய்வுகள், பூமி அறிவியலுக்கான foundation models உணர்தல் நிலையிலிருந்து பல்மாதிரி பகுத்தறிதல் மற்றும் முகவர் சார்ந்த அறிவியல் பணிச்சுற்றுகளுக்குச் செல்லும் போக்கைக் காட்டுகின்றன.

Volume 1, Issue 2 – July 2026

மருத்துவத் துறையில் பல்மாதிரி AI (Multimodal AI) மருத்துவப் பதிவுகள், ஆய்வக முடிவுகள், மருத்துவப் படங்கள் மற்றும் மருத்துவர் குறிப்புகள் போன்ற பல்வேறு தகவல்களை இணைக்க உதவலாம். ஆனால் இத்தகைய பயன்பாட்டில் துல்லியம் மிக முக்கியமானது. ஒரு படத்தின் அடிப்படையில் மட்டும் உருவாக்கப்பட்ட தவறான விளக்கம் மருத்துவ முடிவை பாதிக்கக்கூடும். எனவே AI வழங்கும் முடிவு மனித நிபுணரால் சரிபார்க்கப்பட வேண்டியது அவசியம்.

பல்மாதிரி AI-யின் (Multimodal AI) மற்றொரு முக்கிய முன்னேற்றம் காட்சி சார்ந்த பகுத்தறிதல் ஆகும். 2026ஆம் ஆண்டின் ஆய்வுகள், சில உடல் மற்றும் இடவியல் சார்ந்த பணிகளில் வெறும் மொழி அடிப்படையிலான பகுத்தறிதலைவிட, காட்சி மற்றும் மொழி இணைந்த செயல்முறைகள் பயனுள்ளதாக இருக்கக்கூடும் என்பதை ஆராய்கின்றன. இதன் அடுத்த கட்டம் மனித-கணினி தொடர்பை மாற்றக்கூடும். எதிர்கால AI (Future AI) அமைப்புடன் நாம் தட்டச்சு செய்வது மட்டுமல்லாமல், பேசலாம்; ஒரு பொருளைக் காட்டலாம்; ஒரு ஆவணத்தைப் பகிரலாம்; காணொளியை இயக்கலாம்; உடனடியாக அதன் அடிப்படையில் விவாதிக்கலாம். இதனால் கணினி ஒரு “கட்டளைப் பெறும் கருவி” என்ற நிலையிலிருந்து “சூழலைப் புரிந்துகொள்ளும் துணை” என்ற நிலைக்குச் செல்லும். இருப்பினும், பல்மாதிரி அமைப்புகள் புதிய சவால்களையும் உருவாக்குகின்றன. வெவ்வேறு தரவு வடிவங்களில் உள்ள பிழைகள் ஒன்றோடொன்று இணைந்து தவறான முடிவை உருவாக்கலாம். தனியுரிமை மற்றொரு முக்கியப் பிரச்சினையாகும். முகம், குரல், இருப்பிடம், உடல் இயக்கம் போன்ற தகவல்கள் ஒரே அமைப்பில் இணைக்கப்படும்போது தனிநபர் பற்றிய மிக நுணுக்கமான சுயவிவரம் உருவாகும் வாய்ப்பு உள்ளது. எனவே தரவு குறைப்பு, அனுமதி, பாதுகாப்பான சேமிப்பு மற்றும் வெளிப்படையான பயன்பாட்டுக் கொள்கைகள் அவசியம்.

முடிவுரை

பல்மாதிரி AI (Multimodal AI) என்பது வெறும் தொழில்நுட்ப மேம்பாடு அல்ல; மனிதர்கள் தகவலைப் புரிந்துகொள்ளும் முறைக்கு நெருக்கமாக AI அமைப்புகளை நகர்த்தும் முயற்சியாகும். உரை, காட்சி, ஒலி, காணொளி மற்றும் சூழல் ஆகியவற்றை ஒருங்கிணைக்கும் திறன் கல்வி, மருத்துவம், அறிவியல், தொழில், ஊடகம் மற்றும் படைப்புத் துறைகளில் புதிய வாய்ப்புகளை உருவாக்குகிறது. அதே நேரத்தில் தனியுரிமை, நம்பகத்தன்மை மற்றும் மனித மேற்பார்வை ஆகியவை அதன் வளர்ச்சியின் அடிப்படை நிபந்தனைகளாக இருக்க வேண்டும்.

குறிப்புகள்

[1] கூகுள் கிளவுட். “மல்டிமோடல் ஏஐ.” 2026.
[2] ஐபிஎம். “வாட் இஸ் மல்டிமோடல் ஏஐ?” 2024.
[3] ஜாவோ, சியாங்யு, மற்றும் பலர். “எர்த் சயின்ஸ் ஃபவுண்டேஷன் மாடல்ஸ்: ஃப்ரம் பெர்செப்ஷன் டு ரீஸனிங் அண்ட் டிஸ்கவரி.” *ஆர்க்சிவ்*, 2026.
[4] வூ, ஜியாலாங், மற்றும் பலர். “விஷுவல் ஜெனரேஷன் அன்லாக்ஸ் ஹியூமன்-லைக் ரீஸனிங் த்ரூ மல்டிமோடல் வேர்ல்ட் மாடல்ஸ்.” *ஆர்க்சிவ்*, 2026.
[5] ஓபன்ஏஐ. “மாடல்ஸ்.” *ஓபன்ஏஐ டெவலப்பர் டாக்குமென்டேஷன்*, 2026.
[6] ஆப்பிள். “வாட்ஸ் நியூ இன் தி ஃபவுண்டேஷன் மாடல்ஸ் ஃப்ரேம்வொர்க்.” *டபிள்யூடபிள்யூடிசி* 2026, 2026.

நிதிசார் கட்டுரையாளர் உறுதிமொழி: இல்லை
கட்டுரையாளர் நன்றியுரை: இல்லை

Volume 1, Issue 2 – July 2026

கட்டுரையாளர் உறுதிமொழி: இக்கட்டுரையில் எவ்வித முரண்பாடும் இல்லை என்று உறுதிமொழி அளிக்கிறேன்.

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