தமிழ் செயற்கை நுண்ணறிவு
TAMIL JOURNAL OF ARTIFICIAL INTELLIGENCE
PEER-REVIEWED JOURNAL

# தமிழ் செயற்கை நுண்ணறிவு ஆய்விதழ்

மதிப்பாய்வு செய்யப்படும் இதழ்

# Tamil Journal of Artificial Intelligence

*Peer-Reviewed Journal*

தொகுதி 1, இதழ் 1 – ஜனவரி 2026

Volume 1, Issue 1 – January 2026

முதன்மை ஆசிரியர் / Editor-in-Chief

திருமதி ஜெ. புவனாசுந்தரி / Mrs. J. Bhuvanaasundhari

முனைவர் ஞா. அந்தோனி சுரேஷ் / Dr. G. Antony Suresh

SUSTAINABLE DEVELOPMENT GOALS

*To save environment and end climate change critical situations avoid printing.*

# தமிழ் செயற்கை நுண்ணறிவு ஆய்விதழ்

*மதிப்பாய்வு செய்யப்படும் இதழ்*

# Tamil Journal of Artificial Intelligence

*Peer-Reviewed Journal*

**தொகுதி 1, இதழ் 1 – ஜனவரி 2026**

**Volume 1, Issue 1 – January 2026**

**முதன்மை ஆசிரியர் / Editor-in-Chief**

**திருமதி ஜெ. புவனா சுந்தரி / Mrs. J. Bhuvanaasundhari**

**முனைவர் ஞா.அந்தோனி சுரேஷ் / Dr. G.Antony Suresh**

# Title Verso


| | |
|---|---|
| Journal title | : Tamil Journal of Artificial Intelligence |
| Tamil Title | : தமிழ் செயற்கை நுண்ணறிவு ஆய்விதழ் |
| Abbreviated key-title | : |
| E-ISSN Number | : |
| Editor-in-Chief | : Mrs. J. Bhuvanaasundhari// Dr. G. Antony Suresh |
| Publication frequency | : Bi-Annual (Calendar Cycle - January and July) |
| Volume/Issue Type | : Volume 1 Issue 1 January 2026 |
| Place of Publication | : Virudhunagar |
| Start of Publication | : January 2026 |
| Paper Size (Format) | : Digital A4 Size |
| Medium of Publication | : E-version |
| Subject | : Multidisciplinary Studies |
| Language | : Tamil |
| Access Type (URL or DOI): | Gold OA, Online, Indexed long time in Internet Archive |
| Subscription Type | : APC |
| Publisher | : Maheswari Publishers, (The publishing unit of PADIAN EDUCATIONAL TRUST-TN-32-0003213) |
| Publisher Website | : https://pandianeducationaltrust.com/ |
| Journal Site | : https://pandianeducationaltrust.com/tamil-journal-of-artificial-intelligence.html |


**Pandian Educational Trust supports SDG Goals and is a member of UN SDG Publisher compact.**

## PUBLISHER'S MESSAGE

### Aim & Objectives

***Tamil Journal of Artificial Intelligence* - TJAI** (E-ISSN:) published Bi-yearly (January -July) which is published by Maheswari Publishers, patronized by Pandian Educational Trust, Virudhunagar, Tamil Nadu, India. Tamil Journal of Artificial Intelligence aims to bring down AI in Tamil Language to promote learning support for the Tamil academicians and scholars in the field of teaching. Research through this academic medium motivates in all aspects of main and inter-disciplines of the core area of study with authentic e-publication. Making known of Artificial Intelligence and future studies in the Tamil world aids the scholarly community to gather knowledge on Al Research and its application in the all fields of learning. The audacity and revelation of academic knowledge on AI could foster higher knowledge amidst the Tamil community. All of these motivate best distribution of research that produces positive outcomes for the betterment of research and education and the unification of the people in our world. In achieving the aim, our journal Tamil Journal of Artificial Intelligence has been created.

### Disclaimer

***Tamil Journal of Artificial Intelligence - TJAI*** is committed to research Ethics and consider plagiarism as a crime done to the original author. So, the authors are advised to follow academic ethics with respect to acknowledgment of quotations from other scholarly works. The Publisher & Editors will not be held responsible for any lapse of the provider regarding plagiarism in their manuscripts. The submissions must be original, must accompany the declaration form stating the research paper as an original work, and has not been published elsewhere for any academic or research purpose. The contributor will be liable for such lapses on any legal issues and publication ethics. Contact: editortjai@pandianeducationaltrust.com for submission and see guidelines information at journal site. For institutional subscription contact the publisher.

### Message from the Editor-in-Chief

***Tamil Journal of Artificial Intelligence TJAI*** is a Peer Reviewed Journal which is committed to academic research, welcomes academicians, scholars and students all over the world who to advance their status of academic career and society by their scholarly ideas in all the field of academic studies. Research is to establish, substantiate facts and restate previous works and to resolve issues. An active venture to endow cogent approach to these types for educational reformations through academic research has become the central intent of the inaugural journal issue to bring down scholarly articles on Artificial Intelligence in Tamil language to the Tamil academic world.


**Publisher Contact:**
**Maheswari Dharmalingam**
Pandian Educational Trust (TN-32-0003213),
Maheswari Publishers, (The publishing unit of PANDIAN EDUCATIONAL TRUST-TN-32-0003213)
3/350, Veterinary Hospital Back Side,
Virudhunagar-626001, Tamil Nadu, India.
Mobile: +91 8526769556,
**Email: editortjai@pandianeducationaltrust.com**



**Copyright © 2026-Authors'**

***Tamil Journal of Artificial Intelligence TJAI*** **(E-ISSN:)** is an Open Access journal and Pdf copy can be CC reused within the terms of the BY license https://creativecommons.org/licenses/by/4.0/. Think before you print so that you can save trees and environment.

# ஆசிரியர் குழு / Editorial Board


**ஜெ. புவனா சுந்தரி,**
முதன்மை ஆசிரியர் TJAI,
டீன் ஆஃப் கம்யூடேஷனல் ரிசெர்ச்,
பாண்டியன் கல்வி அறக்கட்டளை,
தமிழ்நாடு, இந்தியா.
**மின்னஞ்சல்: editortjai@pandianeducationaltrust.com**
**பணியிட முகப்புப் பக்கம்:**
**https://pandianeducationaltrust.com/trustees.html**

**Mrs. J. Bhuvanaasundhari, MCA., B.Ed.,**
Editor-in-Chief TJAI,
Dean of Computational Research,
Pandian Educational Trust, Tamil Nadu, India.
**Email: editortjai@pandianeducationaltrust.com**
**Institutional Profile:**
**https://pandianeducationaltrust.com/trustees.html**

**முனைவர் ஞா அந்தோனி சுரேஷ்,**
முதன்மை ஆசிரியர் TJAI,
டீன் ஆஃப் தமிழ் ரிசெர்ச்,
பாண்டியன் கல்வி அறக்கட்டளை,
தமிழ்நாடு, இந்தியா.
உதவிப் பேராசிரியர், சதக்கத்துல்லா அப்பா கல்லூரி.
திருநெல்வேலி-627011, தமிழ்நாடு, இந்தியா.
**மின்னஞ்சல் : deants/@pandianeducationaltrust.com**
**surtamil6@gmail.com**
**பணியிட முகப்புப் பக்கம்:**
**https://pandianeducationaltrust.com/trustees.html**

**Dr. G. Antony Suresh,**
Editor-in-Chief T.JAI
Dean of Tamil Studies,
Pandian Educational Trust, Tamil Nadu, India.
College Assistant Professor, Sadakatullah Appa
(Autonomous), Tirunelveli - 627011, Tamil Nadu, India.
**Email: deants@pandianeducationaltrust.com**
**surtamil6@gmail.com**
**Institutional Profile:**
**https://pandianeducationaltrust.com/trustees.html**

**முதுமுனைவர் அ.இராஜலட்சுமி,**
உதவிப் பேராசிரியர்,
கொங்குநாடு கலை மற்றும் அறிவியல் கல்லூரி
(தன்னாட்சி),
கோயம்புத்தூர்-641029. தமிழ்நாடு, இந்தியா.
**மின்னஞ்சல் : arajalakshmita@kongumaducollege.ac.in**
**பணியிட முகப்புப் பக்கம்:**
**https://kongunaducollege.ac.in/28/department-tamil-member-faculty**

**Dr. A. Rajalakshmi,**
Assistant Professor of Tamil,
Kongunadu Arts and Science College (Autonomous),
Coimbatore-641029,
Tamil Nadu, India.
**Email: arajalakshmita@kongunaducollege.ac.in**
**Institutional Profile:**
**https://kongunaducollege.ac.in/28/department-tamil-member-faculty**

**முனைவர் ந. இரஞ்சன்,**
உதவிப் பேராசிரியர், தமிழ்த்துறை,
எஸ்.ஆர்.எம். பல்கலைக்கழகம்
காட்டாங்குளத்தூர், சென்னை- 603203
தமிழ்நாடு, இந்தியா.
**மின்னஞ்சல்: ranjann@srmist.edu.in**
**பணியிட முகப்புப் பக்கம்:**
**https://www.srmist.edu.in/faculty/dr-n-ranjan/**

**Dr. N. Ranjan,**
Assistant Professor of Tamil, SRM University,
Kattankulathur. Chennai,
Tamil Nadu, India.
**Email: ranjann@smist.edu.in**
**Institutional Profile:**
**https://www.srmist.edu.in/faculty/dr-n-ranjan/**

**அஸ்வாக் அஹமது அகு**
உதவிப் பேராசிரியர் (சீனியர் கிரேட்)
கணினி பயன்பாடுத் துறை,
பி.எஸ்.அப்துர் ரஹ்மான் கிரசண்ட் அறிவியல் மற்றும்
தொழில்நுட்ப நிறுவனம்,
சென்னை, தமிழ்நாடு, இந்தியா.
**மின்னஞ்சல்: ashfauk@crescent.education**
**பணியிட முகப்புப் பக்கம்:**
**https://crescent,education/dr-a-k-ashfauk-ahamed-computer-applications/**

**Dr.A.K.Ashfauk Ahamed**
Assistant Professor (Sr.Gr)
Department of Computer Applications
B.S.Abdur Rahman Crescent Institute of Science and
Technology
Chennai, Tamil Nadu, India.
**Email: ashfauk@crescent.education**
**Institutional Profile: https://crescent.education/dr-a-k-ashfauk-ahamed-computer-applications/**

**முனைவர் த. மகேஸ்வரி,**
தலைவர், பாண்டியன் கல்வி அறக்கட்டளை
விருதுநகர், தமிழ்நாடு. இந்தியா.
**மின்னஞ்சல்: president@pandianeducationaltrust.com**
**பணியிட முகப்புப் பக்கம்:**
**https://pandianeducationaltrust.com/trustees.html**

**Dr. D. Maheswari,**
President, Pandian Educational Trust
Virudhunagar, Tamil Nadu, India.
**Email: president@pandianeducationaltrust.com**
**Institutional Profile:**
**https://pandianeducationaltrust.com/trustees.html**

## பொருளடக்கம் / Contents


| வ. எண் | தலைப்பு / பெயர் | பக்க எண் |
|---|---|---|
| 1. | மெட்டா ஏ.ஐ (Meta Al) டிஜிட்டல் தொடர்புகளின் எதிர்காலமும் செயற்கை நுண்ணறிவின் பரிணாமமும்<br>திருமதி.அ.கோமதி | 1 |
| 2. | தமிழ் மொழி செயலாக்கத்திற்கு ஊக்கமளிக்கும் AI கருவிகள்<br>முனைவர் கி. நாகேந்திரன் | 5 |
| 3. | தமிழ் ஓலைச்சுவடிகளைப் பாதுகாக்கச் செயற்கை நுண்ணறிவின் (AI) பயன்பாடு<br>முனைவர் த. மகேஸ்வரி | 9 |
| 4. | தமிழ் மொழி மற்றும் இலக்கியத்திற்கான நவீன செயற்கை நுண்ணறிவுக் (AI) கருவிகள்<br>முனைவர் கு.ராஜேஷ் குமார் | 12 |
| 5. | க்ளாட் ஏ.ஐ (Claude Al) எனும் செயற்கை நுண்ணறிவு<br>முனைவர் ரா. விநாயகமூர்த்தி | 15 |

தமிழ் செயற்கை நுண்ணறிவு ஆய்விதழ்
*மதிப்பாய்வு செய்யப்படும் இதழ்*

**Tamil Journal of Artificial Intelligence**
*Peer-Reviewed Journal*



**Volume 1, Issue 1 – January 2026**

# மெட்டா ஏ.ஐ (Meta AI) டிஜிட்டல் தொடர்புகளின் எதிர்காலமும் செயற்கை நுண்ணறிவின் பரிணாமமும்

**திருமதி. அ. கோமதி,** உதவிப் பேராசிரியர், வணிகவியல் துறை (S/F), காமராஜ் கல்லூரி, தூத்துக்குடி, தமிழ்நாடு, இந்தியா.
**ORCiD: https://orcid.org/0000-0003-2943-8519**

## அறிமுகம்

தற்கால டிஜிட்டல் யுகத்தில், செயற்கை நுண்ணறிவு (AI) நமது அன்றாட வாழ்வின் ஒரு பகுதியாக மாறிவிட்டது எனலாம். மெய்நிகர் உதவியாளர்கள் முதல் மொழிபெயர்ப்புக் கருவிகள் வரை, செயற்கை நுண்ணறிவுத் தொழில்நுட்பங்கள் நாம் ஒருவருக்கொருவர் தொடர்பு கொள்ளும் விதத்தையும், தொழில்நுட்பத்தைக் கையாளும் முறையையும் முற்றிலுமாக மாற்றியமைத்து வருகின்றன என்பது உண்மை. செயற்கை நுண்ணறிவுத் துறையில் நிகழ்ந்துள்ள முன்னேற்றங்களில் ஒன்று 'மெட்டா ஏ.ஐ' (Meta AI) ஆகும். ஃபேஸ்புக், இன்ஸ்டாகிராம் மற்றும் வாட்ஸ்அப் ஆகிய செயலிகளின் தாய் நிறுவனமான மெட்டாவால் உருவாக்கப்பட்ட இந்த அதிநவீனத் தொழில்நுட்பம், டிஜிட்டல் தொடர்புகளை மிகவும் எளிமையானதாகவும், தனிப்பயனாக்கப்பட்டதாகவும், திறமையானதாகவும் மாற்றும் வகையில் வடிவமைக்கப்பட்டுள்ளது. மேம்பட்ட இயந்திரக் கற்றல் (Machine Learning) வழிமுறைகள் மற்றும் இயற்கை மொழிச் செயலாக்க (NLP) நுட்பங்களின் துணையுடன் இயங்கும் மெட்டா ஏ.ஐ. உரை, உரையாடல், படங்கள் மற்றும் காணொளிகளை உருவாக்கும் ஆற்றல் கொண்டது. நாம் இக்கட்டுரையில், மெட்டா ஏ.ஐ-யின் சிறப்பம்சங்கள், நன்மைகள், பயன்பாட்டுச் செயல்பாடுகள் குறித்து விரிவாகக் காண்போம்.

## மெட்டா ஏ.ஐ-யின் சிறப்பம்சங்கள்

மெட்டா ஏ.ஐ. டிஜிட்டல் தொடர்புகளுக்கான ஒரு நல்ல கருவியாகத் தன்னை நிலைநிறுத்திக் கொள்ளும் வகையில் பல புதுமையான அம்சங்களைக் கொண்டுள்ளது. முதன்மையாக, இது பயனர்களுடன் இயல்பான முறையில் உரையாடல் உதவியாளராகச் செயல்படுகிறது. அதில் கேள்விகளுக்குப் பதிலளிப்பது. பரிந்துரைகளை வழங்குவது. கட்டுரைகளை எழுதுவது போன்ற பணிகளைச் செய்யும் வசதி மேலும் மெட்டாவின் தளங்களான ஃபேஸ்புக், வாட்ஸ்அப், இன்ஸ்டாகிராம் மற்றும் மெசஞ்சர் ஆகியவற்றில் இந்த சேவை தடையின்றி கிடைக்கிறது. பயனர்கள் உள்ளிடும் குறுஞ்செய்திகளின் அடிப்படையில் தத்ரூபமான படங்களை உருவாக்கும் திறனையும் இது பெற்றுள்ளது. "imagine" (கற்பனை) என்ற கட்டளையைப் பயன்படுத்தி பயனர்கள் கலைப்படைப்புகள், அலங்கார வடிவமைப்புகள் அல்லது தனிப்பயன் அசைவூட்டப் படங்களை (GIFs) எளிதாக உருவாக்கிக் கொள்ள முடியும். இதனுடன், பயனர்கள் கட்டளைகளைத் தட்டச்சு செய்வதற்குப் பதிலாக நேரடியாகப் பேசி, உரையாடல் வடிவில் பதில்களைப் பெறும் குரல்வழிப் பயன்பாட்டையும் (Voice Mode) மெட்டா ஏ.ஐ வழங்குகிறது. இது கைகளின் உதவியின்றிச் செயல்படுவதற்கும், மொழிபெயர்ப்புக்கும் பெரிதும் உதவுகிறது. பயனர்களின் விருப்பங்களைப் பகுப்பாய்வு செய்து, அவர்களுக்கு ஏற்றவாறு தனிப்பயனாக்கப்பட்ட உள்ளடக்கப் பரிந்துரைகளை வழங்குவது இதன் மற்றொரு சிறப்பாகும். இது சமூக வலைத்தளங்களில் பயனர்களின் ஈடுபாட்டை அதிகரிக்கிறது எனலாம். இவை மட்டுமின்றி, புகைப்படங்களைச் சமூக ஊடகங்களுக்கு ஏற்ற வகையிலான குறுகிய காணொளிகளாகவோ அல்லது அசைவூட்டப் படங்களாகவோ மாற்றும் உள்ளடக்கியுள்ளது.

Volume 1, Issue 1 – January 2026

**செயல்படும் விதம் மற்றும் நன்மைகள்**

மெட்டா ஏ.ஐ-யின் செயல்பாடுகள் அனைத்தும் அதிநவீன இயந்திரக் கற்றல் (Machine Learning) வழிமுறைகள் மற்றும் இயற்கை மொழிச் செயலாக்க நுட்பங்களின் (NLP) அடிப்படையிலேயே நிகழ்கின்றன. மெட்டா நிறுவனத்தின் திறந்த மூலமான (Open Source) 'லாமா 3.2' (Llama 3.2) என்னும் மேம்பட்ட செயற்கை நுண்ணறிவு மாதிரியின் (AI Model) அடிப்படையில் கட்டமைக்கப்பட்டுள்ள இந்த மென்பொருள் மெட்டா நிறுவனத்தின் மென்பொருள்களில் எல்லாம் ஒருங்கிணைக்கப்பட்டுச் செயல்படுகிறது. இந்தத் தொழில்நுட்பக் கட்டமைப்பு பயனர்களின் கேள்விகளை மிகத் துல்லியமாகவும் அவர்களின் தனிப்பட்ட தேவைகளுக்கு ஏற்பவும் புரிந்துகொண்டு பதிலளிக்க மெட்டா ஏ.ஐ-க்கு உதவுகிறது. இதன் மூலம் கிடைக்கும் நன்மைகள் அளப்பரியவை ஆகும். இது செயற்கை நுண்ணறிவுக் கருவிகளுக்கான அணுகலை ஜனநாயகப்படுத்துகிறது எனலாம். ஏனென்றால், எவ்விதத் தொழில்நுட்ப நிபுணத்துவமும் இன்றிச் சாதாரண மக்களும் ஏ.ஐ தொழில்நுட்பத்தை எளிதாகப் பயன்படுத்தும் வாய்ப்பை மெட்டா உருவாக்கி உள்ளது. மேலும், சிறு கட்டுரைகள் எழுதுதல், படங்கள் மற்றும் காணொளிகளை உருவாக்குதல் போன்ற பல்வேறு அன்றாடப் பணிகளை இது எளிதாக்குவதால், தனிநபர்கள் மற்றும் வணிக நிறுவனங்களின் உற்பத்தித்திறன் பன்மடங்கு அதிகரிக்க பயன்படுகிறது.

**மெட்டா ஏ.ஐ-யின் எதிர்காலப் பயன்பாடுகள்**

எதிர்காலத்தில், மெட்டா ஏ.ஐ-யின் பயன்பாட்டு மருத்துவம், கல்வி, வணிகம் எனப் பல துறைகளிலும் விரிவடைய உள்ளன. மருத்துவத் துறையில், நோயாளிகளுக்கான மருத்துவர் சந்திப்புகளைத் திட்டமிடுதல், மருத்துவப் பதிவுகளைப் பாதுகாப்பாக அணுகுதல் மற்றும் தனிப்பயனாக்கப்பட்ட சுகாதார ஆலோசனைகளை வழங்குதல் போன்றவற்றின் மூலம் மருத்துவ நிர்வாகச் செயல்முறைகளை இது எளிதாக்கும். கல்வித்துறையில், ஒவ்வொரு மாணவரின் தனிப்பட்ட கற்றல் திறனுக்கு ஏற்பப் பாடங்களை வடிவமைத்து வழங்குவதன் மூலம், தனிப்பயனாக்கப்பட்ட கற்றல் அனுபவத்தை இது உறுதி செய்யும்; மேலும், ஆசிரியர்களின் மதிப்பீட்டுப் பணிகளுக்கும் இது பேருதவியாக இருக்கும் எனலாம். வணிக நிறுவனங்களைப் பொறுத்தவரை, வாடிக்கையாளர்களுடனான தொடர்பை மேம்படுத்தவும், தரவுகளைப் பகுப்பாய்வு செய்யவும், வாடிக்கையாளர் ஆதரவை வழங்கவும் மெட்டா ஏ.ஐ மிகச்சிறந்த கருவியாகச் செயல்பட்டு வணிக வளர்ச்சியைத் துரிதப்படுத்தும் வாய்ப்பை நாம் காணமுடிகிறது. ஏனைய துறைகளிலும் மெட்டா ஏ.ஐ-யின் பயன்பாடு செரிவாக்கம் பெற வாய்ப்பு அதிகம் உள்ளது. அதைப் பற்றிய ஆய்வுகளும் நடந்து கொண்டு உள்ளன.

**மெட்டா ஏ.ஐ-யின் சவால்களும் வரம்புகளும்**

மெட்டா ஏ.ஐ பல நன்மைகளையும் வாய்ப்புகளையும் வழங்கினாலும், இதில் கவனிக்கப்பட வேண்டிய சில சவால்களும் வரம்புகளும் உள்ளன. முதன்மையானது தரவுப் பாதுகாப்பு (Data Privacy) மற்றும் தனியுரிமை சார்ந்த விஷயங்களாகும். செயற்கை நுண்ணறிவு மாதிரியை மேம்படுத்துவதற்காகப் பயனர்களின் அதிக அளவிலான தரவுகள் சேகரிக்கப்பட்டுப் பயன்படுத்தப்படுவதால், தரவு மீறல்களும் தனியுரிமைப் பாதிப்புகளும் ஏற்பட வாய்ப்புகள் உள்ளன. இரண்டாவதாக, நெறிமுறைச் சிக்கல்கள் மற்றும் சார்புத்தன்மை (Bias) ஆகியவையாகும். மெட்டா ஏ.ஐ பயிற்றுவிக்கப்படும் தரவுகளில் ஏதேனும் சார்புத்தன்மை அல்லது முன்தீர்மானங்கள் சார்ந்த பிரச்சனைகள் இருந்தால், அது உருவாக்கும் முடிவுகளிலும் பாரபட்சம் வெளிப்படலாம். இது சில சமூகச் சூழல்களில் தேவையற்ற பிரச்சனைகளை ஏற்படுத்தக்கூடும் என்பதால், தரவு மேலாண்மையிலும் சோதனையிலும் அதீத கவனம் தேவைப்படுகிறது.

Volume 1, Issue 1 – January 2026

**முடிவுரை**

மெட்டா ஏ.ஐ ஏ-ஐ தொழில்நுட்பத்தில் புதுமையான அம்சங்களையும் நன்மைகளையும் அடக்கிய ஒரு தொழில்நுட்பக் கருவியாக விளங்குகிறது. இது டிஜிட்டல் தொழில்நுட்பத்தை மக்கள் எளிதாக அணுகுவதற்கும், அவர்களின் உற்பத்தித்திறனை மேம்படுத்துவதற்கும் பெருமளவில் உதவுகிறது எனலாம். இருப்பினும், இதில் பாதுகாப்பு சார்ந்த அபாயங்கள் மற்றும் நெறிமுறைச் சிக்கல்களைத் தவிர்ப்பதில் கூடுதல் கவனம் செலுத்தப்பட வேண்டும். செயற்கை நுண்ணறிவுத் தொழில்நுட்பம் தொடர்ந்து பரிணாம வளர்ச்சி அடைந்து வரும் இக்காலகட்டத்தில், மெட்டா ஏ.ஐ போன்ற கருவிகள் பொறுப்புணர்வு சார்ந்த நெறிமுறைகளுடன் உருவாக்கப்பட்டுப் பயன்படுத்தப்படுவதை உறுதி செய்வது அவசியமாகும். எனவே, மனித வாழ்வை மேம்படுத்தும் வகையில் இத்தொழில்நுட்பத்தின் ஆற்றலைச் சரியாகப் பயன்படுத்தினால், டிஜிட்டல் தொடர்புகளின் எதிர்காலம் மிகச் சிறப்பான முறையில் வடிவமைக்கப்படும் என்பதில் எவ்வித ஐயமுமில்லை.

**குறிப்புகள்**

[1] ஏ.ஐ அட் மெட்டா (AI at Meta). (2024). மெட்டா லாமா 3 அறிமுகம்: இதுவரை வெளிவந்த மிகவும் திறன்வாய்ந்த திறந்த மூல பெரிய மொழி மாதிரி (Introducing Meta Llama 3: The most capable openly available LLM to date). மெட்டா (Meta). https://ai.meta.com/blog/meta-llama-3/

[2] பெண்டர், ஈ. எம்., கெப்ரு, டி., மெக்மில்லன்-மேஜர், ஏ., & ஷ்மிட்செல், எஸ். (Bender, E. M. et al.). (2021). *சீரற்ற கிளிகளின் ஆபத்துகள்: மொழி மாதிரிகள் மிகவும் பெரியதாக இருக்க முடியுமா?* (On the dangers of stochastic parrots: Can language models be too big?). 2021 ஏ.சி.எம் (ACM) நேர்மை, பொறுப்புக்கூறல் மற்றும் வெளிப்படைத்தன்மை மாநாட்டுக் கோவை, 610–623. https://doi.org/10.1145/3442188.3445922

[3] புளோரிடி, எல்., & கோல்ஸ், ஜே. (Floridi, L., & Cowls, J.). (2019). *சமூகத்தில் செயற்கை நுண்ணறிவுக்கான ஐந்து கோள்கைகளின் ஒருங்கிணைந்த கட்டமைப்பு* (A unified framework of five principles for AI in society). ஹார்வர்ட் தரவு அறிவியல் ஆய்விதழ் (Harvard Data Science Review), 1(1). https://doi.org/10.1162/99608f92.8add20c0

[4] மெட்டா (Meta). (2023). எமு: புகைப்பட ஊசிகளைப் பயன்படுத்திப் பட உருவாக்க மாதிரிகளை மேம்படுத்துதல் (Emu: Enhancing image generation models using photogenic needles in a haystack). arXiv. https://doi.org/10.48550/arXiv.2309.15807

[5] டோபோல், ஈ. ஜே. (Topol, E. J.). (2019). உயர் செயல்திறன் மருத்துவம்: மனித மற்றும் செயற்கை நுண்ணறிவின் ஒருங்கிணைப்பு (High-performance medicine: the convergence of human and artificial intelligence). நேச்சர் மெடிசின் (Nature Medicine), 25(1), 44-56. https://doi.org/10.1038/s41591-018-0300-7

[6] டோவ்ரான், எச்., மார்ட்டின், எல்., ஸ்டோன், கே., ஆல்பர்ட், பி., அல்லமஹரி, ஏ., பாபாயி, ஒய்., ... & சியாலோம், டி. (Touvron, H. et al.). (2023). லாமா 2: திறந்த அடிப்படை மற்றும் நேர்த்தியாகச் சரிசெய்யப்பட்ட அரட்டை மாதிரிகள் (Llama 2: Open foundation and finetuned chat models). arXiv. https://doi.org/10.48550/arXiv.2307.09288

**Volume 1, Issue 1 – January 2026**

[7] வாஸ்வானி, ஏ., ஷசீர், என்., பர்மர், என்., உஸ்கோரிட், ஜே., ஜோன்ஸ், எல்., கோம்ஸ், ஏ. என்., ... & போலோசுகின், ஐ. (Vaswani, A. et al.). (2017). கவனம் மட்டுமே உங்களுக்குத் தேவையானது (Attention is all you need). நரம்பியல் தகவல் செயலாக்க அமைப்புகளின் முன்னேற்றங்கள் (Advances in Neural Information Processing Systems), 30.

[8] ஹோம்ஸ், டபுள்யு., பயாலிக், எம்., & பேடல், சி. (Holmes, W., Bialik, M., & Fadel, C.). (2019). கல்வியில் செயற்ை க நுண்ணறிவு: கற்பித்தல் மற்றும் கற்றலுக்கான வாக்குறுதிகள் மற்றும் தாக்கங்கள் (Artificial intelligence in education: Promises and implications for teaching and learning). பாடத்திட்ட மறுவடிவமைப்பு மையம் (Center for Curriculum Redesign).

**நிதிசார் கட்டுரையாளர் உறுதிமொழி**: இல்லை
**கட்டுரையாளர் நன்றியுரை:** இல்லை
**கட்டுரையாளர் உறுதிமொழி:** இக்கட்டுரையில் எவ்வித முரண்பாடும் இல்லை என்று உறுதிமொழி அளிக்கிறேன்.

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

தமிழ் செயற்கை நுண்ணறிவு ஆய்விதழ்
*மதிப்பாய்வு செய்யப்படும் இதழ்*

**Tamil Journal of Artificial Intelligence**
***Peer-Reviewed Journal***

**Volume 1, Issue 1 – January 2026**

# தமிழ் மொழி செயலாக்கத்திற்கு ஊக்கமளிக்கும் AI கருவிகள்

**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

மொழி என்பது ஒரு நிலத்தின் பாரம்பரியத்தையும், வாழ்வின் அடையாளத்தையும், பேச்சாற்றல் மற்றும் எழுத்தாற்றலின் தொடக்கத்தையும், ஒரு சமூகத்தின் சுவாசத்தையும் தன்னகத்தே தாங்கி நிற்கும் உன்னதப் படைப்பாகும். தொல்காப்பியம் தோன்றிய காலம் முதலே. தமிழ் மொழியானது தன் இலக்கண நெறிகளோடு இயல்பான வாழ்வுமுறையையும் இணைத்து செழித்து வளர்ந்து வந்துள்ளது எனக் காணலாம். மனிதக் குரலுக்கு இணையான குரலில் கணினிகள் பேசத் தொடங்கியுள்ள இன்றைய நவீன உலகிலும், தமிழ் மொழி தொழில்நுட்பத்தோடு இணைந்து பயணிக்கத் தொடங்கியுள்ளது என்பதை வியப்புடன் பார்க்க நேரிடுகிறது. இப்பயணத்தில், செயற்கை நுண்ணறிவு (**Artificial Intelligence)** எனப் அழைக்கப்படும் நவீனத் தொழில்நுட்பம் தமிழுக்குப் பல புதுமையான பரிமாணங்களை வழங்கி வருகிறது எண்டது மிகையில்லை. குறிப்பாக, **21-**ஆம் நூற்றாண்டில் செயற்கை நுண்ணறிவுத் தொழில்நுட்பம் உலகளாவிய அளவில் மாபெரும் மாற்றங்களை ஏற்படுத்தி வரும் நிலையில், மனித மொழியைக் கணினி புரிந்துகொள்ளச் செய்யும் 'இயற்கை மொழிச் சீ' **(Natural Language Processing NLP)** முக்கியப் பங்கு வகிக்கிறது என்பது உண்மையே. இத்துறையில் தமிழ் மொழியின் தற்போதைய நிலை, அதனை ஆவணக்காப்பு செய்யும் செயற்கை நுண்ணறிவுக் கருவிகள் குறித்து இக்கட்டுரையில் விரிவாகக் காண்போம்.

## இயற்கை மொழிச் செயலாக்கமும் தமிழ் மொழியும்

இயற்கை மொழிச் செயலாக்கம் (**NLP)** என்பது மனித மொழியைக் கணினி புரிந்துகொள்ளவும், பகுப்பாய்வு செய்யவும், தகுந்த பதிலளிக்கவும் உதவும் ஒரு அதிநவீனத் தொழில்நுட்பம் சார்ந்த துறையாகும். இது இயந்திர மொழிபெயர்ப்பு (**Machine Translation),** உரை அடிப்படையிலான தேடல் (**Text-based Search),** உரையாடல் அமைப்புகள் (**Conversational Systems)** மற்றும் ஒலி அடிப்படையிலான மொழிப் புரிதல் (**Speech-based Language Understanding)** ஆகியவற்றை உள்ளடக்கியதாகும். மிகவும் தொன்மை வாய்ந்த தமிழ் மொழி, அதன் செறிவான இலக்கண அமைப்புகள், வேர்ச்சொற்களின் ஆழம், பன்முகப் பயன்பாடுகள் மற்றும் பண்டைய இலக்கிய வடிவங்கள் ஆகியவற்றின் காரணமாக இயற்கை மொழிச் செயலாக்கத்தில் தனித்துவமான சிறப்புகளையும் அதேவேளையில் சில சவால்களையும் கொண்டுள்ளது. தமிழ் மொழியின் பல்வேறு வட்டார வழக்குகள், சங்க இலக்கிய மரபுகள் மற்றும் வேர்ச்சொல் அடர்த்தி ஆகியவை கணினிப் பரிமாணத்தில் தமிழ் மொழிச் செயலாக்கத்தை மேலும் நுண்ணியதாகவும் சவாலானதாகவும் மாற்றுகின்றன. எனவே, அதை பற்றிய அறிவு நமக்கு அவசியமாகிறது

## செயற்கை நுண்ணறிவில் தமிழ் எதிர்கொள்ளும் சவால்கள்

தமிழில் இயற்கை மொழிச் செயலாக்கத்தை மேம்படுத்துவதில் பல நடைமுறைச் சவால்கள் உள்ளன. ஆங்கிலம் போன்ற உலகளாவிய மொழிகளுக்குப் பல்வேறு பயன்பாடுகளுக்கான தரவுகள் பரவலாகக் கிடைக்கும் ஆனால் தமிழுக்குப் போதுமான அளவிலான தரவுகள் இன்னும் முழுமையாக உருவாக்கப்படவில்லை எனலாம். மேலும், தமிழ் மொழியில் ஒரு சொல்லுக்குப் பல பொருள்கள் இருப்பதாலும், வாக்கிய அமைப்புகளில்

தமிழ் செயற்கை நுண்ணறிவு ஆய்விதழ்
மதிப்பாய்வு செய்யப்படும் இதழ்

Tamil Journal of Artificial Intelligence
*Peer-Reviewed Journal*

Volume 1, Issue 1 – January 2026

பல்வேறு அமைப்புக்கள் காணப்படுவதாலும், கணினிகள் சரியான பொருளைத் துல்லியமாகக் கண்டறிவதில் சிக்கல்கள் எழுகின்றன. இதுதவிர, ஆங்கிலத்திலிருந்து தமிழுக்கோ அல்லது தமிழிலிருந்து பிற மொழிகளுக்கோ மொழிபெயர்க்கும் இயந்திர மொழிபெயர்ப்புக் கருவிகள் முழுமையான துல்லியத்தை அடைய இன்னும் பெருமளவில் மேம்படுத்தப்பட வேண்டிய கட்டாயத்தில் உள்ளன என்பட்டர்ஹெ உண்மை. இத்தகைய சவால்களைக் கடந்து, கணினி இயந்திரம் புரிந்துகொள்ளும் வண்ணம் தமிழை வடிவமைக்கும் இந்தப் பயணமே மொழியின் மறுமலர்ச்சி தொடங்கும் தருணமாக அமைகிறது எனலாம்.

**தமிழ் மொழிச் செயலாக்கத்திற்கான ஏ.ஐ கருவிகள்**

செயற்கை நுண்ணறிவுத் தொழில்நுட்பத்தின் (**AI Technology**) அபரிமித வளர்ச்சியால் தற்போது தமிழ் மொழிச் செயலாக்கத்திற்குப் பல்வேறு கருவிகள் உருவாக்கப்பட்டு வருகின்றன. மாணவர்களிடையே கணினிக் கல்வியை ஊக்குவிப்பதற்காக உருவாக்கப்பட்ட 'எழில்' (**Ezhil**) என்னும் திறந்த மூலத் (**Open Access**) தமிழ் நிரலாக்க மொழி, தமிழில் மென்பொருள் (**Tamil Software**) உருவாக்கப் பயிற்சியை வழங்கி வருகிறது. சென்னை ஐ.ஐ.டி (**IIT Madras**) மற்றும் பிற கல்வி நிறுவனங்களால் உருவாக்கப்பட்ட 'தமிழி என்.எல்.பி' (**ThamizhiNLP**) திட்டம், உரை சுத்திகரிப்பு (**Text Preprocessing**), பாகுபடுத்தல் (**Parsing**), លក់កប់ **4 (Part-of-Speech (POS) Tagging (or Syntactic Analysis)** एल श्री உருவாக்கம் (**Text Generation (or Language Modeling)**) ஆகிய பணிகளைத் தமிழில் செவ்வனே மேற்கொள்கிறது. மேலும், ஏ.ஐ-ஃபார்-பாரத் (**AI4Bharat**) திட்டத்தின் கீழ் சென்னை ஐ.ஐ.டி வழங்கியுள்ள கருவிகள் மொழிபெயர்ப்பு மற்றும் உரை உருவாக்கத்திற்குப் பெரிதும் உதவுகின்றன எனலாம். குறிப்பாக, இண்டிக்ட்ரான்ஸ் (**IndicTrans**), இண்டிக்பெர்ட் (**IndicBERT**) போன்ற செயற்கை நுண்ணறிவு மாதிரிகள் தமிழில் சிறப்பாகச் செயல்படுகின்றன. மிகமுக்கிமாக கூகுள் ட்ரான்ஸ்லேட் (**Google Translate**) மற்றும் மெட்டா நிறுவனத்தின் 'என்.எல்.எல்.பி' (**Meta NLLB - No Language Left Behind**) ஆகிய திட்டங்கள் குறைந்த தரவுகளைக் கொண்டு மேம்பட்ட இயந்திர மொழிபெயர்ப்பை வழங்கி வருகின்றன. இதுமட்டுமன்றி, கூகுள் ஸ்பீச் ஏ.பி.ஐ (**Google Speech API**) மற்றும் மொஸில்லாவின் காமன் வாய்ஸ் (**Mozilla Common Voice**) போன்ற குரல்வழிச் செயற்கை நுண்ணறிவுக் கருவிகள் மூலம் தமிழுக்கான பேச்சு-உரை (**Speech-to-Text**) மற்றும் உரை-பேச்சு (**Text-to-Speech**) தொழில்நுட்பங்களும் குறிப்பிடத்தக்க வளர்ச்சியை அடைந்துள்ளன என்பதை காணமுடிகிறது.

**பயன்பாடுகளும் எதிர்கால வாய்ப்புகளும்**

தற்பொழுது, செயற்கை நுண்ணறிவுக் கருவிகள் (**AI Softwares**) கல்வி மற்றும் ஆராய்ச்சித் துறைகளிலும் பெரும் தாக்கத்தை ஏற்படுத்தி வருகின்றன. உரை சுருக்கம் (**Text Summarization or Automatic Text Summarization), - 2 (Question-Answer Generation or Automated Question Generation, QAG)** மற்றும் மாணவர் செயல்திறன் **4 (Student Performance Evaluation or Automated Assessment/Learning Analytics)** போன்றவை இத்தொழில்நுட்பத்தின் மூலம் எளிதாக மேற்கொள்ளப்பட்டு வருகின்றன. சங்க இலக்கியங்கள் உள்ளிட்ட பண்டைய தமிழ் நூல்களை எண்ணிம (**Digital**) வடிவில் ஆவணப்படுத்தவும், அவற்றை ஆழமாக ஆய்வு (**Deep Research**) செய்யவும் இக்கருவிகள் உதவுகின்றன. எதிர்காலத்தில், பல்வேறு தமிழ்ப் பேச்சு வழக்குகளை (**Dialects**) அடிப்படையாகக் கொண்ட பன்முக உதவிகள், செயற்கை நுண்ணறிவு வழியிலான தமிழ் மருத்துவ ஆலோசனைகள், விவசாயம் சார்ந்த வழிகாட்டுதல்கள் மற்றும் இலக்கிய விமர்சனங்களுக்கான கணினி மதிப்பீடுகள் எனப் பல துறைகளிலும் செயற்கை நுண்ணறிவுத் தொழில்நுட்பம் தமிழ் மொழியோடு இணைந்து

**Volume 1, Issue 1 – January 2026**

செயல்படும் வாய்ப்புகள் உள்ளன எனலாம். அதற்கான நிலையான வழிகாட்டு முறைகளை **(SOP)** உருவாக்குவது மிக முக்கியமாகும்.

**சமூகவியல் மற்றும் நெறிமுறைச் சிந்தனைகள்**

இவை அனைத்திற்கும் மேலாக, தமிழுக்கான செயற்கை நுண்ணறிவுக் கருவிகள் உருவாக்கப்படும்போது சமூகவியல் மற்றும் நெறிமுறைச் சிந்தனைகளைக் கருத்தில் கொள்வது மிகவும் முக்கியம். தரவுகளில் நம்பகத்தன்மையை உறுதி செய்தல், மொழி மரபைக் காத்தல், சமூகப் பாகுபாடுகளைத் தவிர்த்தல் மற்றும் பண்பாட்டு அடையாளத்தை நிலைநிறுத்துதல் ஆகிய பண்புகளோடு இத்தொழில்நுட்பம் வளர வேண்டும். ஒருதலைப்பட்சமான அல்லது தவறான தரவுகள் சமூகத்தில் எதிர்மறையான தாக்கங்களை ஏற்படுத்தக்கூடும். அதனால், இத்தகைய விழிப்புணர்வு இல்லையெனில், தொழில்நுட்ப உலகில் தமிழ் வெறும் 'பயனர் மொழியாக' மட்டுமே விளங்கும் அபாயம் உள்ளது. அதனை தவிர்க்க இம்மாதிரியான முறைகளை அறிந்து கொள்வது முக்கியமாகும். வருங்காலங்களில், பள்ளி மற்றும் கல்லூரியிலும் **AI** தொழில்நுட்ப குறியீடுகளை பாடமாக இணைப்பது அவசியமாகின்றது.

**முடிவுரை**

செயற்கை நுண்ணறிவு **(AI)** மற்றும் இயற்கை மொழிச் செயலாக்கம் **(NLP)** ஆகியவை தமிழ் மொழிக்காகப் பல புதிய வாய்ப்புகளை உருவாக்கி வருகின்றன. சரியான தரவுகளையும் சமூகச் சிந்தனையையும் உள்ளடக்கி உருவாக்கப்படும் செயற்கை நுண்ணறிவுக் கருவிகள், தமிழ் மொழியை உலகளாவிய மையத்தில் நிலைநிறுத்துவதோடு மட்டுமல்லாமல், அடுத்த தலைமுறைக்கு நமது பண்பாட்டு அடையாளத்தைக் கடத்தும் பாலமாகவும் அமையும். இன்றைய நிலையில் தமிழில் செயற்கை நுண்ணறிவு வெறும் தொழில்நுட்ப முன்னேற்றமாக மட்டுமல்லாமல் தமிழ் மொழி கணினி பயன்பாட்டில் சிறக்க ஓர் அரியதோர் வாய்ப்பாக அமையும். மேலும், தமிழ் மொழியை கணினி அறிவியல் சார்ந்து ஆய்வு செய்து படிக்க இம்மாதிரியான பகுப்பு முறைகள் உதவும்.

**குறிப்புகள்**

[1] ஆர்டிலா, ஆர்., பிரான்சன், எம்., டேவிஸ், கே., ஹென்ரெட்டி, எம்., கோலர், எம்., மேயர், ஜே., மோரேஸ், ஆர்., சாண்டர்ஸ், எல்., பயர்ஸ், எஃப். எம்., & வெபர், ஜி. (Ardila, R. et al.). (2020). காமன் வாய்ஸ்: ஒரு மாழிபரும் பன்மொழிப் பேச்சுத் தாவத்தொகுப்பு (Common voice: A massively-multilingual speech corpus). 12-வது மொழி வளங்கள் மற்றும் மதிப்பீட்டு மாநாட்டுக் கோவை (Proceedings of the 12th Language Resources and Evaluation Conference), 4218-4222. https://aclanthology.org/2020.lrec-1.520

[2] கக்வானி, டி., குஞ்சுகுட்டன், ஏ., கோல்லா, எஸ்., கோகுல், என். சி., பட்டாச்சார்யா, பி., காப்பரா, எம். எம்., & குமார், பி. (Kakwani, D. et al.). (2020). இண்டிக் என்.எல்.பி குட் (IndicNLPSuite): இந்திய மொழிகளுக்கான ஒருமொழித் தரவுத்தொகுப்புகள், மதிப்பீட்டு அளவுகோல்கள் மற்றும் முன்-பயிற்சி பெற்ற பன்மொழி மாதிரிகள். கணக்கீட்டு மொழியியல் சங்கத்தின் கண்டுபிடிப்புகள் (Findings of the Association for Computational Linguistics: EMNLP 2020), 4948-4961. https://doi.org/10.18653/v1/2020.findings-emnlp.445

[3] கூகுள் (Google). (2024). கூகுள் கிளவுட் ஸ்பீச்-டு-டெக்ஸ்ட் ஏ.பி.ஐ (Google Cloud Speech-to-Text API) [கணினி மென்பொருள்]. கூகுள் கிளவுட். https://cloud.google.com/speech-to-text

[4] கூகுள் (Google). (2024). கூகுள் கிளவுட் ட்ரான்ஸ்லேஷன் ஏ.பி.ஐ (Google Cloud Translation API) [கணினி மென்பொருள்]. கூகுள் கிளவுட். https://cloud.google.com/translate

**Volume 1, Issue 1 – January 2026**

[5] சர்வவேஸ்வரன், க., & டயஸ், ஜி. (Sarveswaran, K., & Dias, G.). (2020). தமிழி என்.எல்.பி (ThamizhiNLP): தமிழ் இயற்கை மொழிச் செயலாக்கத்திற்கான ஒரு கட்டமைப்பு. 12-வது மொழி வளங்கள் மற்றும் மதிப்பீட்டு மாநாட்டுக் கோவை (Proceedings of the 12th Language Resources and Evaluation Conference), 6745–6753. https://aclanthology.org/2020.lrec-1.833

[6] என்.எல்.எல்.பி குழு (NLLB Team), வாஸ்டா-ஜுஸ்ஸா, எம். ஆர்., கிராஸ், வ., மற்றும் பலர். (2022). எந்த மொழியும் பின்தங்கவில்லை: மனிதர்களை மையமாகக் கொண்ட இயந்திர மொழிபெயர்ப்பை அளவிடுதல் (No language left behind: Scaling human-centered machine translation). arXiv. https://doi.org/10.48550/arXiv.2207.04672

[7] முத்தையா, மு. (Muthiah, M.). (2023). எழில்: ஒரு தமிழ் நிரலாக்க மொழி (Ezhil: A Tamil programming language) (பதிப்பு 2.0) [கணினி மென்பொருள்]. எழில் மொழித் திட்டம் (Ezhil Language Project). https://ezhillang.org/

**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.

தமிழ் செயற்கை நுண்ணறிவு ஆய்விதழ்
*மதிப்பாய்வு செய்யப்படும் இதழ்*

Tamil Journal of Artificial Intelligence
*Peer-Reviewed Journal*



**Volume 1, Issue 1 – January 2026**

# தமிழ் ஓலைச்சுவடிகளைப் பாதுகாக்கச் செயற்கை நுண்ணறிவின் (AI) பயன்பாடு

**முனைவர் த. மகேஸ்வரி**, முதன்மை ஆசிரியர், தமிழ்மொழி மற்றும் இலக்கிய பன்னாட்டு ஆய்விதழ் (IJTLLS), விருதுநகர், தமிழ்நாடு, இந்தியா.
**ORCiD: https://orcid.org/0000-0002-4187-0120**

## அறிமுகம்

தமிழ் நாகரிகத்தின் தொன்மையான அடையாளமாக விளங்கும் ஓலைச்சுவடிகள், நம் முன்னோர்களின் அறிவுச் செல்வத்தை இன்றளவும் வெளிப்படுத்தும் ஆதாரமாகத் திகழ்கின்றன. சங்க இலக்கியங்கள் முதல் சித்தர் பாடல்கள் வரை, உளவியல், மருத்துவம், வானியல் போன்ற பல்வேறு அறிவியல் துறைகளின் முதன்மை ஆதாரங்களாக இவை விளங்குகின்றன. எனினும், காலத்தின் தொடர்ச்சியான சீற்றத்துக்கு இலக்காகி, இந்த அரிய அறிவுப் பொக்கிஷங்கள் மெல்ல மெல்ல சிதைந்து மறையும் அபாயத்தில் உள்ளன. இந்த நெருக்கடியான சூழலில், செயற்கை நுண்ணறிவு (**AI**) தொழில்நுட்பமானது ஓலைச்சுவடிகளைப் பாதுகாக்கும் பணியில் ஒரு புதிய நம்பிக்கை தருவதாக அமைந்து உள்ளது. பழங்கால ஓலைச்சுவடிகளை எண்ணிம (**Digital**) முறையில் பாதுகாப்பதில் செயற்கை நுண்ணறிவு மாற்றங்களை ஏற்படுத்துகிறது என்பதைப் பற்றி இக்கட்டுரை விளக்குகிறது.

## ஓலைச்சுவடிகளைப் பாதுகாப்பதில் உள்ள சவால்கள்

ஓலைச்சுவடிகளைப் பாதுகாக்கும் பணியில் நாம் பலவிதமான சவால்களை எதிர்கொள்ள வேண்டியுள்ளது. முதலாவதாக, ஓலைச்சுவடிகள் பனை இழைகளால் ஆனவை என்பதால், ஈரப்பதம், தட்பவெப்ப நிலை மாற்றங்கள் மற்றும் தொடர்ச்சியான பயன்பாட்டினால் ஏற்படும் தேய்மானம் ஆகியவற்றால் அவை எளிதில் சிதைந்துவிடுகின்றன. குறிப்பாக, ஆயிரம் முதல் பல நூறு ஆண்டுகளுக்கு முற்பட்ட பழமையான ஓலைகள் சிதைந்த நிலையில் காணப்படுகின்றன. இரண்டாவதாக, கரப்பான் பூச்சி (**Cockroach**), சில்வர் ஃபிஷ் (**Silverfish**) போன்ற பூச்சிகளின் தாக்குதலால் பல அரிய சுவடிகள் நிரந்தரமாக அழிந்துபோன நிலையில் உள்ளன. இத்தகைய சிதைவுகளுக்கு அப்பால், சுவடிகளில் உள்ள எழுத்து வடிவங்களும் படிப்பதற்கு பெரும் சவாலாக அமைகின்றன. பண்டைய தமிழ் எழுத்து முறைகள், கிரந்த எழுத்துகள், வட்டெழுத்துகள் மற்றும் சுவடிகளுக்கே உரிய தனித்துவமான சுருக்கெழுத்துகள், குறியீடுகள் போன்றவை தற்கால செயற்கை நுண்ணறிவு மூலம் 'படிம எழுத்துணரி' (**OCR**) தொழில்நுட்பம் செயலாக்கம் செய்யப்படலாம். மேலும், சங்ககால மற்றும் இடைக்காலத் தமிழ் மொழியின் சொல்லாட்சிகளும் இலக்கண அமைப்புகளும் நவீனத் தமிழிலிருந்து பெரிதும் வேறுபடுவதால், செயற்கை நுண்ணறிவு அமைப்புகளுக்கு இவற்றைப் புரிந்துகொள்ளக் கூடுதல் தரவுகளும் பயிற்சியும் தேவைப்படுகின்றன.

## எண்ணிம மயமாக்கலில் செயற்கை நுண்ணறிவின் பங்கு

ஓலைச்சுவடிகளைப் பாதுகாப்பாக எண்ணிம (**Digital**) வடிவத்திற்கு மாற்றுவதில் பல்வேறு செயற்கை நுண்ணறிவுத் தொழில்நுட்பங்கள் ஒருங்கிணைந்து செயல்படுத்தலாம். இதன் முதல் படியாக இருப்பது உயர்தர ஸ்கேனிங் மற்றும் படமாக்கல் முறையாகும். மல்டி-ஸ்பெக்ட்ரல் (**Multi-spectral**) இமேஜிங் தொழில்நுட்பத்தின் மூலம் பல்வேறு ஒளி அலைநீளங்களைப் பயன்படுத்தி, கண்ணுக்குத் தெரியாமல் மறைந்துபோன ஓலைச் சுவடிஎழுத்துகளைக் கண்டறிய முடிகிறது. மேலும், முப்பரிமாண (**3D**) ஸ்கேனிங் மற்றும் **1200 DPI**-க்கும் அதிகமான துல்லியத்தன்மை கொண்ட படமாக்கல் முறை மூலம் ஓலையின்

**Volume 1, Issue 1 – January 2026**

அமைப்பு மற்றும் சிதைவுகள் மிகத் தெளிவாகப் பதிவு செய்ய முடியும். இதனைத் தொடர்ந்து, படமாக்கப்பட்ட சுவடிகளிலிருந்து எழுத்துகளைப் பிரித்தெடுக்க கணினியின் துணைகொண்டு 'ஆழமான கற்றல்' **(Deep Learning)** மூலம் அடிப்படையிலான பிரத்தியேகத் தமிழ் 'ஓ.சி.ஆர்' **(OCR)** அமைப்புகள் பயன்படுத்தலாம். இவை பழைய எழுத்து வடிவங்களை மிகத் துல்லியமாக அடையாளம் கண்டு செயற்கை நுண்ணறிவு நெறிமுறைகள் மூலம் பிழைகளைத் தானாகவே சரிசெய்துகொள்கின்றன. இவ்வாறு ஓலைச்சுவடிகளில் இருந்து பிரித்தெடுக்கப்பட்ட பண்டைய தமிழ் உரைகளைப் புரிந்துகொள்வதற்கு 'இயற்கை மொழிச் செயலாக்க' **(NLP)** மாதிரிகள் பெருமளவில் உதவுகின்றன. சங்க இலக்கியங்கள் மற்றும் இடைக்கால நூல்களுக்கெனச் சிறப்பாகப் பயிற்றுவிக்கப்பட்ட இந்த மாதிரிகள், வாக்கியங்களின் வரலாற்றுச் சூழல் மற்றும் சொற்பொருளைப் பகுப்பாய்வு செய்து, பழைய தமிழை நவீன தமிழுக்குத் தானியங்கியாக மொழிபெயர்க்கும் பணியையும் செவ்வனே செய்கின்றன. எல்லாவற்றிற்கும் மேலாக, சிதைந்த ஓலைகளை எண்ணிம முறையில் மறுசீரமைப்பதில் செயற்கை நுண்ணறிவின் பங்கு அளப்பரியது எனலாம். ஓலைச் சுவடிகளில் அழிந்துபோன பகுதிகளை நிரப்புதல், மங்கிப்போன எழுத்தை மீட்டெடுத்தல் மற்றும் காணாமல் போன வரிகளை மொழிநடைக்கேற்ப ஊகித்து முழுமைப்படுத்துதல் போன்ற வியக்கத்தக்க பணிகளை செயற்கை நுண்ணறிவு சார்ந்த **OCR** தொழில்நுட்பம் சாத்தியமாக்குகிறது.

**ஓலைச்சுவடி பாதுகாப்புக்கான படிநிலைச் செயல்முறை**

ஓலைச்சுவடிகளைப் பாதுகாக்கும் இந்தச் செயற்கை நுண்ணறிவுச் செயல்முறை பல முக்கியப் படிநிலைகளைக் கொண்டுள்ளது. தொடக்கக் கட்டமாக, ஓலைகளை மென்மையாக்குதல், ஈரப்பதம் மற்றும் வெப்பநிலையைக் கட்டுப்படுத்துதல், பூச்சி மற்றும் தூசிகளை அகற்றுதல் போன்ற முன்-செயலாக்கப் **(Pre-processing)** பணிகள் மேற்கொள்ள வேண்டும். அதன் பின்னர், பல கோணங்களில் முப்பரிமாண ஸ்கேனர்கள் **(3 D Scanners)** மற்றும் மல்டி-ஸ்பெக்ட்ரல் கேமராக்கள் **(Multi Spectral Cameras )** மூலம் ஒளி, நிழல் சமநிலைப்படுத்தப்பட்டு உயர்தரப் படங்களாக அவை மாற்றம் செய்யப்பட வேண்டும். அடுத்த கட்டமாக, பிரத்தியேக மென்பொருள்கள் மூலம் எழுத்துகள் பிரித்தெடுக்கப்பட்டு, பிழை திருத்தங்கள் மேற்கொள்ளப்பட வேண்டும். அதைத் தொடர்ந்து, இயற்கை மொழிச் செயலாக்கக் **(NLP)** கருவிகளின் உதவியுடன் அந்த உரையின் சொற்பொருள், இலக்கணம் மற்றும் வரலாற்றுச் சூழல் ஆகியவற்றை ஆழமாகப் பகுப்பாய்வு செய்யப்பட வேண்டும். இறுதியாக, இவ்வாறு பெறப்பட்ட தகவல்கள் அனைத்தும் பன்மொழி இடைமுகங்களைக் கொண்ட, எளிதில் தேடிக் கண்டறியக்கூடிய எண்ணிமத் தரவுத்தளங்களில் ஆராய்ச்சியாளர்களுக்கும் பொதுப் பயன்பாட்டிற்காக காப்பகப்படுத்த வேண்டும்.

**முடிவுரை**

ஓலைச்சுவடிகளைப் பாதுகாக்கச் செயற்கை நுண்ணறிவுத் தொழில்நுட்பத்தைப் பயன்படுத்துவது என்பது ஒரு தொழில்நுட்ப முன்னேற்றம் மற்றுமல்லாமல்; அது தமிழ் கலாச்சாரப் பாரம்பரியத்தைப் பாதுகாப்பதற்கான ஒரு வரலாற்றுச் சிறப்புமிக்கச் செயலாகும் எனலாம். ஓ.சி.ஆர் **(OCR)** தொழில்நுட்பம், இயற்கை மொழிச் செயலாக்கம் **(NLP)** மற்றும் முப்பரிமாண ஸ்கேனிங் **(3 D Scanning)** போன்ற நவீனத் தொழில்நுட்பங்களைப் பயன்படுத்தி சிதைந்துபோகும் நிலையிலுள்ள பனை ஓலைச் சுவடிகளில் உள்ள அரிய தகவல்களை எக்காலத்திற்கும் அழியாமல் பாதுகாக்க முடியும். எனவே, ஒருங்கிணைந்த எண்ணிம வடிவமாக்கும் முயற்சிகள், தமிழ் மொழியின் தொன்மையையும் தமிழர்களின் இணையற்ற கலாச்சாரத்தின் எதிர்காலத்தையும் பாதுகாக்கும் அரணாக விளங்கும் என்பது திண்ணம் .

**Volume 1, Issue 1 – January 2026**

ஆகையால், நமது எதிர்கால சந்ததிகள் இவைபோன்ற கணினி துறைகளை செவ்வன கற்று ஆவனப்படுத்தும் முறைகளையும் கற்றுக்கொள்ள வேண்டும்.

**குறிப்புகள்**

[1] கார்த்திகேயன், சு., & அருள், அ. (2023). சித்தாந்த ஒலைச்சுவடிகளை மீட்டமைக்கப் படச் செயலாக்கம் (Image Processing) மற்றும் செயற்கை நுண்ணறிவின் பங்கு. இந்தியத் தகவல் தொழில்நுட்ப ஆய்விதழ், 19(1), 55-68.

[2] சிவசந்த், மு., & கபிலன், வ. (2024). டிரான்ஸ்பார்மார் (Transformer) தொழில்நுட்பம் வழியாகத் தானியங்கியான தமிழ் ஒலைச்சுவடிகளை அறிதலும் பாதுகாத்தலும். செயற்கை நுண்ணறிவு மற்றும் தரவு அறிவியல் ஆய்விதழ், 15(2), 45-58.

[3] பார்த்திபன், க., & முருகன், தவ. (2022). ஒலைச்சுவடிகளில் உள்ள தமிழ் எழுத்துக்களைத் தானியங்கி முறையில் கண்டறிவதற்கான கன்வல்யூஷனல் நியூரல் நெட்வார்க் (CNN) அணுகுமுறை. பன்னாட்டு கணினியியல் ஆய்விதழ், 10(4), 200-215.

[4] பவல்முருகன், கு. (2023). தமிழ் ஒலைச்சுவடி எழுத்துக்களை அறிவதற்கான ஆழ் கற்றல் (Deep Learning) வழிமுறைகள் மற்றும் அதன் பயன்பாடுகள். தமிழ் கணினி தொழில்நுட்ப மாநாட்டு ஆய்வுக் கோவை, 8, 112-120.

[5] AI4Bharat. (பதவி இல்லை). முகப்புப் பக்கம். https://ai4bharat.iitm.ac.in/

[6] Institut Français de Pondichéry (IFP). (பதவி இல்லை). முகப்புப் பக்கம். https://www.ifpindia.org/

[7] கிட்ஹப் (GitHub). (பதவி இல்லை). தமிழ் ஒலைச்சுவடி வாசிப்புத் திட்டங்கள் (Tamil Palm Leaf Character Recognition). https://github.com/

[8] தஞ்சாவூர் மகாராஜா சரபோஜியின் சாரஸ்வதி மகால் நூலகம். (பதவி இல்லை). முகப்புப் பக்கம். https://sarasvathimahallibrary.tn.gov.in/

[9] தமிழ் இணையக் கல்விக் கழகம் (Tamil Virtual Academy). (பதவி இல்லை). முகப்புப் பக்கம். https://www.tamilvu.org/

**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.

தமிழ் செயற்கை நுண்ணறிவு ஆய்விதழ்
*மதிப்பாய்வு செய்யப்படும் இதழ்*



**Tamil Journal of Artificial Intelligence**
*Peer-Reviewed Journal*

**Volume 1, Issue 1 – January 2026**

# தமிழ் மொழி மற்றும் இலக்கியத்திற்கான நவீன செயற்கை நுண்ணறிவுக் (AI) கருவிகள்

**முனைவர் கு.ராஜேஷ் குமார்,** ஆங்கில உதவிப் பேராசிரியர், CDOE, மதுரை காமராசர் பல்கலைக்கழகம், மதுரை, தமிழ் நாடு, இந்தியா.
**ORCiD: https://orcid.org/0000-0002-7312-8473**

## அறிமுகம்

உலகின் மிகப் பழமையான மொழியான தமிழ் மொழி, ஆயிரக்கணக்கான ஆண்டுகளாக பண்பாடு மற்றும் பாரம்பரியத்தின் முக்கிய அடையாளமாக விளங்கி வருகிறது. நவீனத் தொழில்நுட்பத்தின் வளர்ச்சியைக் கண்டுவரும் இக்காலத்தில், தமிழ் மொழியானது தனது வளமான இலக்கிய மற்றும் கலாச்சாரப் பாரம்பரியத்தைப் பாதுகாக்கவும், உலகெங்கும் பரவச் செய்யவும் புதிய தொழில்நுட்பக் கருவிகளைத் தழுவி வருகிறது. அவ்வகையில், தமிழ் மொழி கற்றல் மற்றும் ஆய்வில் மாபெரும் தாக்கத்தை ஏற்படுத்தி வரும் தொழில்நுட்பங்களில் செயற்கை நுண்ணறிவு (AI) முதன்மையானதாக தற்பொழுது விளங்குகிறது. சமீப காலத்தில், தமிழ் மொழி மற்றும் இலக்கிய கற்றலின் தேவைகளைப் பூர்த்தி செய்யும் வகையில் பல்வேறு செயற்கை நுண்ணறிவுச் செயலிகள் உருவாக்கப்பட்டுள்ளன. செயற்கை நுண்ணறிவுச் செயலிகள் தமிழ் மொழியுடன் தொடர்பு கொள்ளும் விதத்திலும், அதனைப் புரிந்துகொள்ளும் முறையிலும் ஒரு மாபெரும் புரட்சியை ஏற்படுத்தி உள்ளதைக் காணமுடிகிறது.

## குரல்வழி அறிதல் தொழில்நுட்பம்

தமிழ் மொழியில் செயற்கை நுண்ணறிவு செயல்பாட்டில் மிக முக்கியமான பயன்பாடு அல்லது குரல்வழி அறிதல்' (Speech Recognition) தொழில்நுட்பமாகும். இத்தொழில்நுட்பத்தின் மூலம் மொழி இயங்கும் அமைப்புகள், பேசப்படும் தமிழைத் துல்லியமாகப் புரிந்துகொண்டு அதனை எழுத்து வடிவமாக மாற்றுகின்றன. இதனால் பயனர்கள் குரல் கட்டளைகளைப் பயன்படுத்தித் திறன்மிக்கச் சாதனங்களுடன் எளிதாக உரையாட முடிகிறது. கூகுள் அசிஸ்டண்ட் (Google Assistant), அமேசான் அலெக்சா (Amazon Alexa) மற்றும் மைக்ரோசாப்ட் அஸூர் (Microsoft Azure) போன்ற குரல் உதவி மென்பொருள்கள் மூலம் தமிழ் மொழி முழுமையாகஇயக்க முறைகளை உள்ளடக்கியுள்ளன. இதன் மூலம், பயனர்கள் தமிழில் தகவல்களைத் தேடவும், நினைவூட்டல்களை அமைக்கவும், வீடுகளில் உள்ள ஸ்மார்ட் சாதனங்களைக் கட்டுப்படுத்தவும் முடிகிறது. மேலும், ஆப்பிள் நிறுவனத்தின் சிரி (Siri), சாம்சங்கின் பிக்ஸ்பி (Bixby) போன்ற குரல் மூலம் உதவி புரியும் AT கருவிகள் தமிழ் மொழி பயன்பாட்டிற்கான முறையை விரிவுபடுத்தும் முயற்சிகளில் ஈடுபட்டுள்ளன. நம் அதற்கான முறைகளை நமது நாட்டிலும் செயல்படுத்த வேண்டும். அதற்கு பல்கலைக்கழகங்கள், IIT, IIM, NIT தன்னார்வ நிறுவனங்கள் நிதிஅளித்து மென்பொருள் உருவாக்கும் வாய்ப்பை அதிகரிக்கலாம். இவ்வாறான முறைகள் தற்சார்பு முறைகளை ஊக்குவிக்கும்.

## இயந்திர மொழிபெயர்ப்பும் உரைச் சுருக்கமும்

தமிழ் மொழியில் செயற்கை நுண்ணறிவின் மற்றொரு குறிப்பிடத்தக்க மைல்கல் இயந்திர மொழிபெயர்ப்பு (Machine Translation) மென்பொருள் ஆகும். மொழித் தடைகளை உடைத்து உலகளாவிய தகவல் தொடர்பை சாத்தியமாக்கும் வகையில், கூகுள் டிரான்ஸ்லேட் (Google Translate), மைக்ரோசாப்ட் டிரான்ஸ்லேட்டர் (Microsoft Translator), ஐ.பி.எம் வாட்சன் (IBM Watson), மற்றும் எஸ்.டி.எல் (SDL) போன்ற தளங்கள் தமிழ் மொழி மொழிபெயர்ப்புக்கு உதவுகின்றன. இக்கருவிகள் மூலம் பயனர்கள் உரை, பேச்சு மற்றும் வலைப்பக்கங்களைச்

தமிழ் செயற்கை நுண்ணறிவு ஆய்விதழ்
மதிப்பாய்வு செய்யப்படும் இதழ்
Tamil Journal of Artificial Intelligence
Peer-Reviewed Journal
Volume 1, Issue 1 – January 2026


சுலபமாகத் தமிழிலிருந்து பிற மொழிகளுக்கும், பிற மொழிகளிலிருந்து தமிழுக்கும் மொழிபெயர்த்துக் கொள்ள முடியும். இதற்கு அடுத்தபடியாக, நீண்ட தமிழ் உரைகளைச் சுருக்கமான, அர்த்தமுள்ள செய்திகளாக மாற்றும் 'உரைச் சுருக்கத் தொழில்நுட்பம்' (Text Summarization) பெரும் கவனத்தைப் தற்பொழுது பெற்றுள்ளது. ஐ.பி.எம் வாட்சனின் இயற்கை மொழிப் புரிதல் (NLP) கருவி, குயில்பாட் (Quillbot) போன்ற செயலிகள் செய்தி கட்டுரைகள், ஆராய்ச்சி ஆவணங்கள் மற்றும் பெரிய புத்தகங்களின் சாராம்சத்தை விரைவாகப் புரிந்துகொள்ளும் வகையில் சுருக்கி வழங்கி, வாசிப்பாளர்களின் நேரத்தையும் உழைப்பையும் மிச்சப்படுத்துகின்றன. இவை இன்றைய காலத்தில் மொழிபெயர்ப்பிற்கு மிகவும் உதவி புரிகின்றன.

**உணர்வுப் பகுப்பாய்வும் மொழிக் கற்றலும்**

தமிழ் உரைகளைப் பகுப்பாய்வு செய்து, அவற்றின் பின்னணியில் உள்ள கருத்துகள் மற்றும் உணர்ச்சித் தொனிகளைத் தீர்மானிக்கும் 'உணர்வுப் பகுப்பாய்விலும்' (Sentiment Analysis) செயற்கை நுண்ணறிவு சிறப்பாகச் செயல்படுகிறது. மீனிங்க்லவுட் (MeaningCloud), ஐ.பி.எம் வாட்சன் போன்ற கருவிகள் சமூக ஊடகங்களைக் கண்காணிக்கவும், நுகர்வோரின் கருத்துகளைப் பகுப்பாய்வு செய்யவும் வணிக நிறுவனங்களுக்குப் பெரிதும் உதவுகின்றன. மறுபுறம், டியோலிங்கோ (Duolingo), ரொசெட்டா ஸ்டோன் (Rosetta Stone), மாண்ட்லி (Mondly) மற்றும் பாபெல் (Babbel) போன்ற மொழிக் கற்றல் செயலிகள், இயந்திரக் கற்றல் (Machine Learning) வழிமுறைகளைப் பயன்படுத்தித் தனிப்பயனாக்கப்பட்ட கற்றல் பாதைகளை உருவாக்குகின்றன. இது உலகெங்கிலும் உள்ள மக்கள் தமிழை மிக எளிதாகவும், ஊடாடும் வகையிலும் (Interactive) கற்பதற்குப் பெரும் உதவியாக அமைகிறது.

**இலக்கியப் பகுப்பாய்வில் செயற்கை நுண்ணறிவு**

தமிழ் இலக்கியங்களை ஆய்வு செய்வதிலும் செயற்கை நுண்ணறிவின் பங்கு அளப்பரியதாக உள்ளது. தொன்மையான இலக்கியப் படைப்புகளில் உள்ள மொழியியல் வடிவங்கள், கருப்பொருள்கள் மற்றும் மையக்கருத்துகளைப் பகுப்பாய்வு செய்து, ஆசிரியரின் நோக்கம் மற்றும் வரலாற்றுச் சூழல் குறித்த புதிய நுண்ணறிவுகளை இக்கருவிகள் வழங்குகின்றன. ஸ்டான்ஃபோர்டு பல்கழகத்தின் இயற்கை மொழிச் செயலாக்கக் குழு போன்ற அமைப்புகள் உருவாக்கும் கருவிகளைக் கொண்டு, திருவள்ளுவர், கம்பர் மற்றும் பாரதியார் போன்றோரின் இலக்கிய படைப்புகளை ஆய்வாளர்கள் மிக நுட்பமாகப் பகுப்பாய்வு செய்து, அவற்றின் கலாச்சார முக்கியத்துவத்தை உலகறியச் செய்து வருகின்றனர்.

**சவால்களும் எதிர்காலத் திட்டங்களும்**

தமிழ் மொழி மற்றும் இலக்கியத்தில் செயற்கை நுண்ணறிவு மிகப்பெரிய தாக்கத்தைக் கொண்டிருந்தாலும், இதில் சில சவால்களும் உள்ளன. குறிப்பாக, தமிழ் போன்ற நுணுக்கங்கள் நிறைந்த மொழிக்குத் தேவையான உயர்தரமான பயிற்சித் தரவுகள் (Training datasets) போதுமான அளவில் இல்லாதது ஒரு பெரும் தடையாகும். இருப்பினும், எதிர்காலத்தில் பெயர்ச்சொல் மற்றும் நிறுவனங்களை அறிதல் (Named Entity Recognition), சொல் வகை அறிதல் (POS Tagging), வாக்கியங்களின் இலக்கண அமைப்பைப் பகுப்பாய்வு செய்தல் (Dependency Parsing), தமிழில் புதிய உரைகளைத் தானாகவே உருவாக்குதல் (Text Generation) மற்றும் தமிழ் ஒலியமைப்புகளை உருவாக்குதல் (Speech Synthesis) ஆகிய துறைகளில் செயற்கை நுண்ணறிவு அபரிமிதமான வளர்ச்சியை எட்டும் என எதிர்பார்க்கப்படுகிறது. இதன்

மூலம் தமிழ் மற்றும் கணினி சார் பயன்பாடு ஊக்குவிக்கப்படும். இவை போன்ற தொழில் நுட்பங்களை நமது தமிழ் நாட்டில் உள்ள கல்வி நிறுவனங்கள் எடுத்துச் செய்யலாம்.

**Volume 1, Issue 1 – January 2026**

**முடிவுரை**

செயற்கை நுண்ணறிவுத் தொழில்நுட்பமானது தமிழ் மொழி மற்றும் இலக்கியத்தோடு ஒருங்கிணைக்கப்படுவது, தமிழ் மொழியைப் பாதுகாத்து வளர்க்கும் ஒரு வரலாற்றுச் சிறப்புமிக்க நிலையாகும். இத் தொழில்நுட்பங்கள் மொழித் தடைகளைத் தாண்டிய தகவல் பரிமாற்றத்தை சாத்தியமாக்குவது மட்டுமின்றி, அறிஞர்களும் மாணவர்களும் தமிழ் இலக்கியங்களை ஆழமாகப் புரிந்துகொள்ளவும் வழிவகுக்கின்றன. தொழில்நுட்பம் நாளுக்கு நாள் வளர்ச்சியைடந்து வரும் நிலையில், டிஜிட்டல் யுகத்தில் தமிழ் போன்ற மொழிகள் பின்தங்கி விடாமல் இருப்பதை உறுதி செய்ய வேண்டியது மிகவும் அவசியமாகும். செயற்கை நுண்ணறிவின் ஆற்றலைச் சரியாகப் பயன்படுத்துவதன் மூலம், தமிழ் மொழிக்கு எண்ணற்ற புதிய வாய்ப்புகளை உருவாக்கி, வருங்காலத் தலைமுறையினருக்கு நமது செழுமையான இலக்கியப் பண்பாட்டைக் கொண்டு சேர்க்க முடியும். உயர் கல்வியில் இம்மாதிரியான கற்றல் முறைகளை இணைக்க வேண்டும்.

**குறிப்புகள்**

[1] அமேசான் (Amazon). (2024). அமேசான் அலெக்ஸா (Amazon Alexa) [கணினி மென்பொருள்]. https://developer.amazon.com/en-US/alexa

[2] ஐ.பி.எம் (IBM). (2024). ஐ.பி.எம் வாட்சன் இயற்கை மொழிப் புரிதல் (IBM Watson Natural Language Understanding) [கணினி மென்பொருள்]. https://www.ibm.com/watson/natural-language-understanding

[3] குயில்பாட் (QuillBot). (2024). குயில்பாட் ஏ.ஐ (QuillBot AI) [கணினி மென்பொருள்]. லெர்னியோ (Learneo, Inc.). https://quillbot.com/

[4] கூகுள் (Google). (2024). கூகுள் டிரான்ஸ்லேட் மற்றும் அசிஸ்டண்ட் (Google Translate and Assistant) [கணினி மென்பொருள்]. https://translate.google.com/

[5] டுவோலிங்கோ (Duolingo). (2024). டுவோலிங்கோ மொழிக் கற்றல் தளம் (Duolingo Language learning platform) [மொபைல் செயலி]. https://www.duolingo.com/

[6] மைக்ரோசாப்ட் (Microsoft). (2024). மைக்ரோசாப்ட் அஸூர் மற்றும் டிரான்ஸ்லேட்டர் (Microsoft Azure and Translator) [கணினி மென்பொருள்]. https://azure.microsoft.com/

[7] மான்னிங், சி. டி., சுராஜனா, எம்., பாயர், ஜே., ஃபிங்மகல், ஜே., பெத்தல், எஸ். ஜே., & மெக்லாஸ்கி, டி. (Manning, C. D. et al.). (2014). ஸ்டான்ஃபோர்டு கோர்என்.எல்.பி இயற்கை மொழிச் செயலாக்கக் கருவித் தொகுதி (52-வது வருடாந்திர கணக்கீட்டு மொழியியல் சங்கக் கூட்டத்தின் செயல்முறைகள்: அமைப்பு விளக்கக்காட்சிகள், 55-60.

[8] ரொசெட்டா ஸ்டோன் (Rosetta Stone). (2024). ரொசெட்டா ஸ்டோன் மொழிக் கற்றல் (Rosetta Stone Language Learning) [கணினி மென்பொருள்]. https://www.rosettastone.com/

**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.

தமிழ் செயற்கை நுண்ணறிவு ஆய்விதழ்
*மதிப்பாய்வு செய்யப்படும் இதழ்*

**Tamil Journal of Artificial Intelligence**
*Peer-Reviewed Journal*



**Volume 1, Issue 1 – January 2026**

# க்ளாட் ஏ.ஐ (Claude AI) எனும் செயற்கை நுண்ணறிவு

**முனைவர் ரா. விநாயகமூர்த்தி**, உதவிப் பேராசிரியர், குற்றவியல் மற்றும் குற்ற நீதித் துறை, மனோன்மணியம் சுந்தரனார் பல்கலைக்கழகம், திருநெல்வேலி-12, தமிழ்நாடு, இந்தியா.
**ORCiD: https://orcid.org/0000-0002-6165-8100**

## அறிமுகம்

இன்றைய நவீனத் தொழில்நுட்ப உலகில் செயற்கை நுண்ணறிவின் **(AI)** வளர்ச்சி கற்பனைக்கு எட்டாத வகையில் விரிவடைந்து வருகிறது. இந்த வரிசையில், ஆந்த்ரோபிக் **(Anthropic)** நிறுவனத்தால் உருவாக்கப்பட்ட 'க்ளாட் ஏ.ஐ' **(Claude AI)** ஒரு அதிநவீனச் செயற்கை நுண்ணறிவு மாதிரியாக உருவெடுத்துள்ளது. கூகுள். ஓப்பன் ஏ.ஐ போன்ற பெருநிறுவனங்களின் தயாரிப்புகளுக்குப் பலத்த போட்டியாக விளங்கும் க்ளாட், பயனர்களின் பாதுகாப்பிற்கும், துல்லியமான தகவல் பரிமாற்றத்திற்கும் அதிக முக்கியத்துவம் அளிக்கிறது. க்ளாட் ஏ.ஐ வெறுமனே ஒரு உரையாடல் கருவியாக மட்டுமன்றி, மனிதர்களின் சிக்கலான பணிகளை எளிதாக்கும் ஒரு நம்பகமான டிஜிட்டல் உதவியாளராகத் தன்னை நிலைநிறுத்திக் கொண்டுள்ளது. இந்தக் கட்டுரையில், க்ளாட் ஏ.ஐ-யின் தனித்துவமான சிறப்பம்சங்கள் மற்றும் அதன் செயல்பாட்டுத் திறன்கள் குறித்து விரிவாகக் காணலாம்.

## அரசியலமைப்புச் செயற்கை நுண்ணறிவும் பாதுகாப்பும்

க்ளாட் ஏ.ஐ-யின் மிக முக்கியமான சிறப்பம்சம் அதன் பாதுகாப்பு கட்டமைப்பாகும். ஆந்த்ரோபிக் நிறுவனம் இதனை 'அரசியலமைப்புச் செயற்கை நுண்ணறிவு' **(Constitutional AI)** என்ற கோட்பாட்டின் அடிப்படையில் உருவாக்கியுள்ளது. அதாவது, மனிதர்களுக்குத் தீங்கு விளைவிக்காத, பாரபட்சமற்ற, நேர்மையான மற்றும் பயனுள்ள தகவல்களை மட்டுமே வழங்கும் வகையில் இதற்கெனப் பிரத்தியேகமான நெறிமுறைகள் வகுக்கப்பட்டுள்ளன. தவறான தகவல்களைப் பரப்புதல், வெறுப்புணர்வைத் தூண்டும் கருத்துகளை உருவாக்குதல் அல்லது ஆபத்தான வழிமுறைகளை வழங்குதல் போன்றவற்றைத் தவிர்க்கும் வகையில் க்ளாட் மிகவும் கவனமாகப் பயிற்றுவிக்கப்பட்டுள்ளது. இதனால், பெரும் நிறுவனங்களும் ஆராய்ச்சியாளர்களும் பணிகளுக்குப் பயன்படுத்துகின்றனர்.

## பிரம்மாண்டமான தரவுப் பகுப்பாய்வுத் திறன்

க்ளாட் ஏ.ஐ-யின் மற்றொரு வியக்கத்தக்க அம்சம் அதன் பிரம்மாண்டமான தரவுப் பகுப்பாய்வுத் திறன் அல்லது 'கான்டெக்ஸ்ட் விண்டோ' **(Context Window)** ஆகும். சாதாரணச் செயற்கை நுண்ணறிவுக் கருவிகளால் குறிப்பிட்ட அளவிலான சொற்களை மட்டுமே ஒரே நேரத்தில் நினைவில் வைத்துக் கொள்ள முடியும். ஆனால், க்ளாட் ஏ.ஐ-யால் ஒரே நேரத்தில் சுமார் இரண்டு இலட்சம் டோக்கன்கள் **(Tokens)** அல்லது அதற்கும் மேலான தரவுகளை உள்வாங்கிப் பகுப்பாய்வு செய்ய முடியும். இது ஏறக்குறைய ஒரு பெரிய புத்தகத்தின் முழு அளவிலான சொற்களை பகுப்பதாகும். இதனால், பயனர்கள் நூற்றுக்கணக்கான பக்கங்களைக் கொண்ட ஆராய்ச்சி ஆவணங்கள், நிதி அறிக்கைகள் அல்லது பெரிய அளவிலான மென்பொருள் நிரல்களைக் கொடுத்து, அதிலிருந்து சுருக்கங்களையோ அல்லது குறிப்பிட்ட தகவல்களையோ மிகத் துல்லியமாகப் பெற முடியும். இதன் மூலம் கற்றல் முறை எளிமை படுத்த முடியும். நமது இந்திய நாட்டில் இம்முறைகளை பாட திட்டத்தின் மூலம் நமது மாணாக்கருக்கு படிக்கும் வாய்ப்பினை வழங்க வேண்டும்.

## மேம்பட்ட பகுத்தறிவு மற்றும் நிரலாக்கத் திறன்

க்ளாட் ஏ.ஐ-யின் புதிய பதிப்புகளான ஓபஸ் **(Opus)**, சானட் **(Sonnet)** மற்றும் ஹைகூ

**Volume 1, Issue 1 – January 2026**

**(Haiku)** ஆகியவை மனிதர்களுக்கு இணையான பகுத்தறிவுத் திறனைக் (**Advanced Reasoning)** கொண்டுள்ளன. சிக்கலான கணிதப் புதிர்களைத் தீர்ப்பது, தர்க்கரீதியான கேள்விகளுக்கு விடையளிப்பது மற்றும் ஆழமான தரவுப் பகுப்பாய்வுகளை மேற்கொள்வது ஆகியவற்றில் க்ளாட் வியக்கத்தக்க வேகத்துடனும் துல்லியத்துடனும் செயல்படுகிறது. மேலும், மென்பொருள் உருவாக்குநர்களுக்கு இது ஒரு வரப்பிரசாதமாக அமைந்துள்ளது எனலாம். பல்வேறு கணினி மொழிகளில் பிழையற்ற நிரல்களை (**Coding / Algorithm)** எழுதுவது, ஏற்கனவே உள்ள நிரல்களில் உள்ள தவறுகளைக் கண்டறிந்து திருத்துவது போன்ற பணிகளைக் க்ளாட் ஏ.ஐ மிகச் சிறப்பாகச் செய்கிறது.

**இயல்பான மொழிநடையும் நுட்பமான புரிதலும்**

பொதுவாகச் செயற்கை நுண்ணறிவுக் கருவிகள் எழுதும் கட்டுரைகள் அல்லது பதில்கள் இயந்திரத்தனமாக இருப்பதை நாம் உணர்ந்திருப்போம். ஆனால், க்ளாட் ஏ.ஐ மனிதர்களின் இயல்பான மொழிநடையை மிக நெருக்கமாகப் பிரதிபலிக்கிறது. இது பயனர்களின் நுட்பமான அறிவுறுத்தல்களைப் புரிந்துகொண்டு, அதற்கேற்பத் தனது எழுத்து நடையை (**Tone and Style)** மாற்றியமைத்துக் கொள்ளும் திறன் கொண்டது. கட்டுரைகள், கதைகள், முறையான வணிகக் கடிதங்கள் அல்லது ஆக்கப்பூர்வமான விளம்பர வாசகங்கள் என எதை எழுதச் சொன்னாலும், அதீத இயந்திரத்தன்மை இன்றி, வாசிப்பதற்கு ஆழமான கருத்துகளைக் கொண்ட படைப்புகளை இது வழங்குகிறது. இதுவே இதன் சிறப்பம்சம் ஆகும்.

**பகுப்பாய்வுத் திறன் மற்றும் பன்மொழி திறன்**

க்ளாட் ஏ.ஐ வெறும் எழுத்துகளைப் பகுப்பாய்வு செய்வதோடு நின்றுவிடாமல், மேம்பட்ட பார்வைத்திறனையும் (**Vision Capabilities)** கொண்டுள்ளது. பயனர்கள் பதிவேற்றும் வரைபடங்கள், விளக்கப்படங்கள், கையெழுத்துப் பிரதிகள் மற்றும் புகைப்படங்களை மிகத் துல்லியமாகப் புரிந்துகொண்டு, அவற்றிலிருந்து தகவல்களைப் பிரித்தெடுக்கும் ஆற்றலை இது பெற்றுள்ளது. இத்துடன், தமிழ் உள்ளிட்ட உலகின் பல்வேறு மொழிகளைச் சரளமாகப் புரிந்துகொள்ளவும், இலக்கணப் பிழையின்றி மொழிபெயர்க்கவும், உரையாடவும் கூடிய பன்மொழி பயன்பாட்டையும் க்ளாட் வழங்குகிறது.

**முடிவுரை**

எனவே, செயற்கை நுண்ணறிவுத் தொழில்நுட்பம் நாளுக்கு நாள் புதுமைகளைப் படைத்து வரும் நிலையில், க்ளாட் ஏ-ஐ ஒரு பொறுப்பான மற்றும் பாதுகாப்பான **AI** கண்டுபிடிப்பாகத் திகழ்கிறது. மனிதர்களின் அறிவுசார் பணிகளுக்கு ஒரு சிறந்த துணையாக இருப்பதோடு மட்டுமல்லாமல், நெறிமுறைகளுடன் கூடிய தொழில்நுட்ப வளர்ச்சியின் அவசியத்தையும் இது உணர்த்துகிறது. க்ளாட் ஏ.ஐ பாதுகாப்பு, பிரம்மாண்டமான தரவுப் பகுப்பாய்வு, ஆக்கப்பூர்வமான எழுத்துத்திறன் மற்றும் மேம்பட்ட பகுத்தறிவு ஆகிய சிறப்பம்சங்களின் கலவையாக எதிர்கால டிஜிட்டல் உலகில் ஒரு தவிர்க்க முடியாத சக்தியாக உருவெடுத்துள்ளது என்பது மறுக்க முடியாத உண்மையாகும். இம்மாதிரியான **AI** படைப்பு முறைகளை நம் அறிந்து கொள்வது மிகவும் அவசியமாகின்றது.

**குறிப்புகள்**

[1] ஆந்த்ரோபிக் (Anthropic). (2023). க்ளாட் **2.1** வெளியீடு மற்றும் 200,000 டோக்கன் தரவு (Claude 2.1 and 200K context window). (Anthropic Blog), https://www.anthropic.com/news/claude-2-1

**Volume 1, Issue 1 – January 2026**

[2] ஆந்த்ரோபிக் (Anthropic). (2024). கிளாட் 3 மாதிரி குடும்பம்: ஓபஸ், சோனட், ஹைகு (The Claude 3 model family: Opus, Sonnet, Haiku). ஆந்த்ரோபிக் (Anthropic). https://www-cdn.anthropic.com/de8ba9b01c9ab7cbabf5c33b80b7bbc618857627/Model_Card_Claude_3.pdf

[3] ஆஸ்கெல், ஏ., பாய், ஒய்., சென், ஏ., டி.ரெய்ஸ், டி., துனன், எஸ்., கங்குலி, டி., ... & கபிலன், ர. (Askell, A. et al.). (2021). ஒருங்கிணைப்பிற்கான ஒரு ஆய்வகமாக பொதுவான மொழி உதவியாளர் (A general language assistant as a laboratory for alignment). arXiv. https://doi.org/10.48550/arXiv.2112.00861

[4] கங்குலி, டி., ரலோவிட், எல்., தகர்னியன், ர., ஆஸ்கெல், ஏ., பாய், ஒய்., கோடவத், எஸ்., ... & கபிலன், ர. (Ganguli, D. et al.). (2022). போதிப்புகளைக் குறைக்க மொழி மாதிரிகளைச் சோதித்தல்: முறைகள் மற்றும் கற்றறிந்த பாடங்கள் (Red teaming language models to reduce harms: Methods, scaling behaviors, and lessons learned). arXiv. https://doi.org/10.48550/arXiv.2209.07858

[5] கோடவத், எஸ்., மற்றும் பலர் (Kadavath, S. et al.). (2022). மொழி மாதிரிகளுக்குத் (பெரும்பாலும்) தங்களுக்கு என்ன தெரியும் என்று தெரியும் (Language models (mostly) know what they know). arXiv. https://doi.org/10.48550/arXiv.2207.05221

[6] கேப்ரியல், ஐ. (Gabriel, I.). (2020). செயற்கை நுண்ணறிவு, மதிப்புகள் மற்றும் சீரமைப்பு (Artificial intelligence, values, and alignment). மைண்ட்ஸ் அண்ட் மெஷின்ஸ் (Minds and Machines), 30(3), 411-437. https://doi.org/10.1007/s11023-020-09539-2

[7] ஜுராஃப்ஸ்கி, டி., & மார்ட்டின், ர. எச். (Jurafsky, D., & Martin, J. H.). (2023). பேச்சு மற்றும் மொழிச் செயலாக்கம் (Speech and Language Processing) (3-வது பதிப்பு). ஸ்டான்ஃபோர்டு பல்கலைக்கழகம். https://web.stanford.edu/~jurafsky/slp3/

[8] பாய், ஒய்., கடவத், எஸ்., குண்டு, எஸ்., ஆஸ்கெல், ஏ., தகர்னியன், ர., ஜோன்ஸ், ஏ., ... & கபிலன், ர. (Bai, Y. et al.). (2022). அரசியலமைப்புச் செயற்கை நுண்ணறிவு: ஏ.ஐ பின்னூட்டத்திலிருந்து பெறப்படும் தீங்கற்ற தன்மை (Constitutional AI: Harmlessness from AI feedback). arXiv. https://doi.org/10.48550/arXiv.2212.08073

[9] பெரெஸ், ஈ., மற்றும் பலர் (Perez, E. et al.). (2022). மாதிரி-எழுதப்பட்ட மதிப்பீடுகளுடன் மொழி மாதிரி நடத்தைகளைக் கண்டறிதல் (Discovering language model behaviors with model-written evaluations). arXiv. https://doi.org/10.48550/arXiv.2212.09251

[10] லியு, என். எஃப்., மற்றும் பலர் (Liu, N. F. et al.). (2023). நடுவில் தொலைந்தது: மொழி மாதிரிகள் எவ்வாறு நீண்ட தரவுகளைப் பயன்படுத்துகின்றன (Lost in the middle: How language models use long contexts). arXiv. https://doi.org/10.48550/arXiv.2307.03172

**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.