

TII Launches Falcon-Emirati Alongside New Arabic AI Models for Speech and Visual Text Understanding

Falcon-Emirati is a 7-billion-parameter model purpose-built for Emirati vocabulary, expressions and cultural context

Model achieves 84.83% on a native Emirati-language benchmark, outperforming every Arabic and multilingual open-source model evaluated

Falcon-ASR delivers leading Emirati and Arabic speech transcription while supporting accessible, multilingual voice services for the UAE's diverse communities

Falcon-OCR-Arabic delivers leading text-extraction performance from images through a lightweight model

Abu Dhabi, UAE, 6 October 2026 – The Technology Innovation Institute (TII), the applied research arm of Abu Dhabi's Advanced Technology Research Council (ATRC), today announced the launch of three new artificial intelligence models that bring Emirati language and culture, as well as multimodality, within reach of developers, researchers and everyday users.

Falcon-Emirati is a 7-billion-parameter language model purpose-built for Emirati Arabic, capturing the dialect's vocabulary, expressions, idioms and cultural context. It launches alongside **Falcon-ASR (Automatic Speech Recognition)**, a compact 1.6-billion-parameter model that converts spoken Standard and Emirati Arabic, English, French, Spanish and Portuguese into written text, and **Falcon-OCR-Arabic (Optical Character Recognition)**, a lightweight, efficient and powerful model designed to extract Arabic text and structured content from images and documents.

H.E. Faisal Al Bannai, Adviser to the UAE President and Secretary General of ATRC, said: "Our language belongs in the future we are building. With Falcon-Emirati, we are putting our knowledge and expertise behind that belief. There is real pride in building technology that understands our people. And there is real freedom in having the capability here in the UAE to develop it further, set our own priorities and turn our ambitions into something people can use."

The models create a more natural interface between people and AI across text, speech and visual content, strengthening Arabic in AI. Falcon-Emirati enables interaction in the UAE's local dialect, while Falcon-ASR supports transcription across languages widely spoken by Emiratis, residents and international communities. Potential applications include multilingual customer service, meeting and interview transcription, subtitling, accessibility tools and voice-enabled digital services.

Arabic is spoken by hundreds of millions of people, but it is not a single, uniform spoken form. Modern Standard Arabic is used primarily in formal writing, education and news, while everyday

conversation takes place through distinct regional dialects. In the UAE, Emirati Arabic carries its own vocabulary, rhythms, idioms and cultural references, particularly across poetry, proverbs and anecdotes, whose meaning cannot always be understood through literal translation.

Like many Arabic dialects, it remains underrepresented in the data used to train AI systems. As a result, a model trained primarily on Modern Standard Arabic may understand every word of an Emirati sentence while still missing its intended meaning.

Dr. Najwa Aaraj, Chief Executive Officer of TII, said: “Sovereign AI capability must reflect the language used in daily life. Emirati Arabic carries distinctive expressions, cultural references and ways of communicating. Falcon-Emirati and Falcon-ASR will help ensure that the next generation of AI understands not only Arabic but how Emirati communities actually speak it.”

Built on Falcon-H1-Arabic, Falcon-Emirati was developed by TII researchers using native Emirati content, Modern Standard Arabic material covering Emirati culture and heritage, and synthetic data informed by Emirati-specific linguistic resources. The model scored 84.83% on Alyah, a native Emirati Arabic benchmark covering everyday language, figurative expressions, heritage knowledge and poetry, outperforming every Arabic and multilingual open-source model evaluated. Falcon-Emirati will be available through a [Falcon Chat platform](#).

Falcon-ASR extends TII's Arabic AI capabilities into speech, converting spoken Emirati Arabic, Modern Standard Arabic, English, French, Spanish and Portuguese into written text. Despite its compact 1.6-billion-parameter size, the model delivered leading accuracy across public Arabic and internal Emirati speech benchmarks, outperforming larger systems included in the evaluation. On Emirati speech, it surpassed the performance of a 30-billion-parameter multimodal model.

Falcon-ASR also records when each word is spoken within an audio file, supporting applications including accurate subtitling, searchable recordings, meeting and interview transcription, accessibility tools and other voice-enabled services.

Falcon-OCR-Arabic extends the Falcon family into Arabic optical character recognition and document understanding. The lightweight model extracts text and structured information from images and documents, including tables, mathematical formulas and different document sections. Its capabilities can support applications including document digitization, searchable archives and automated information extraction.

Dr. Hakim Hacid, Chief Researcher of TII's Artificial Intelligence and Digital Research Center, said: “These capabilities empower developers, researchers and communities across the region to build on Emirati language and culture, while expanding what is possible for Arabic AI more broadly. Across Falcon-Emirati, Falcon-ASR and Falcon-OCR-Arabic, we have applied targeted specialization to language, speech and document understanding, addressing areas where Arabic AI still has significant room to advance.”

The launch further expands TII's sovereign AI capabilities across language and modalities reflecting the languages and communities those technologies are designed to serve.

###

About Technology Innovation Institute (TII):

The Technology Innovation Institute (TII) is the dedicated applied research pillar of Abu Dhabi's Advanced Technology Research Council (ATRC). TII is a pioneering global research and development center that focuses on applied research and new-age technology capabilities. The Institute has nine dedicated research centers in advanced materials, autonomous robotics, cryptography, AI and digital science, directed energy, quantum, secure systems, propulsion and space, and renewable and sustainable energy. By working with exceptional talent, universities, research institutions, and industry partners from all over the world, TII connects an intellectual community and contributes to building an R&D ecosystem that reinforces the status of Abu Dhabi and the UAE as a global hub for innovation.

For more information, visit www.tii.ae

For media inquiries, please contact:

joe.sigolo@tii.ae
tii@edelman.com