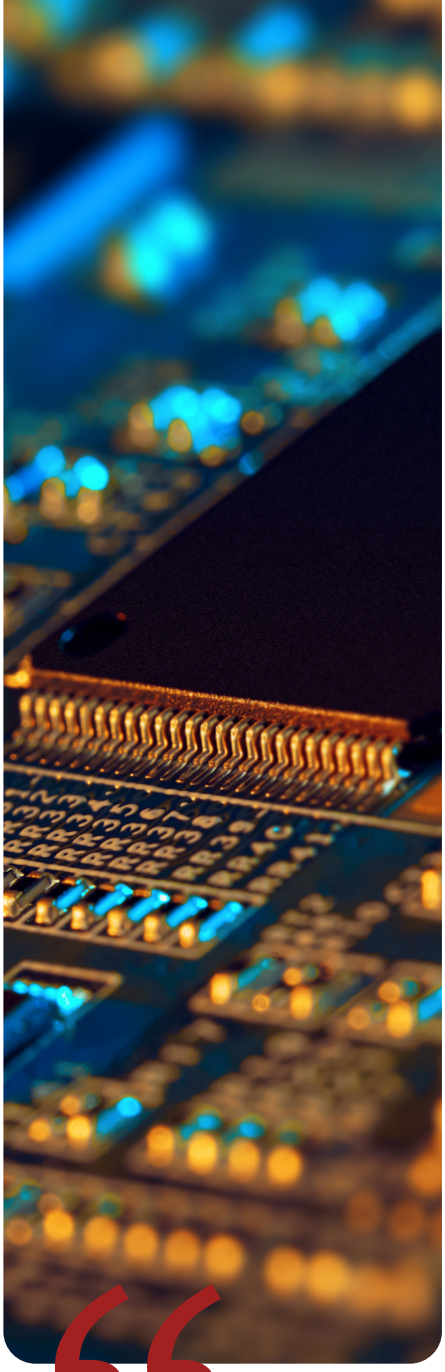
strategy&
Part of the PwC network

From automation to advantage

How Arabic-first AI and local talent can redefine GCC business services



Across the GCC, artificial intelligence is reshaping how core operations are delivered. Finance cycles are shortening, HR interactions are increasingly automated, and procurement is becoming more digitized. Business services – once back-office functions – are evolving into enterprise platforms with efficiency, insight, and resilience.

However, this momentum is hitting a structural ceiling. Most AI systems still operate primarily in English. In a region where Arabic underpins contracts, customer interaction, and public administration, this limits scale and value.

The next productivity leap will not come from more automation, but from Arabic-first AI systems paired with local talent positioned to govern and scale them.

AI is not yet fluent enough

The region has made significant progress in consolidating finance, HR, IT, procurement, and customer operations into centralized business services. Expectations now extend beyond efficiency to quality at scale. AI is the key, yet it remains limited by language and context.

Large language models are trained predominantly on English data. In major datasets such as Common Crawl's CC100, which is used to train mainstream LLMs, english accounts for roughly 82GB of text compared to just 5.4GB in Arabic, despite more than 400 million Arabic speakers.[1] Dialect diversity, code-switching, and linguistic complexity widen the gap between capability and performance.

The impact is evident across the GCC. Pilots demonstrate potential but struggle to scale. Chatbots require frequent human intervention, document-processing tools misinterpret text, and automation often merely shifts effort without eliminating it. Generic Arabic-capable models achieve only around 30-50% accuracy on Arabic financial text.[2]

This is not due to AI capability, but a limitation of enterprise-grade Arabic fluency. More mature ecosystems such as China have reduced this gap through investment in native-language models, data and deployment.

How Arabic-first AI changes the economics

When models are not natively fluent, organizations compensate through supervision, rework, and parallel manual processes, eroding the advantage of automation. More than 20% of Arabic chatbot queries still require human handover.[3]

Arabic-native LLMs are still nascent, and mostly limited to front-end conversational layers. AI has struggled to move into end-to-end workflow automation, where impact compounds across processes.

> The next productivity leap will not come from more automation, but from Arabic-first AI systems paired with local talent positioned to govern and scale them.

Across the GCC, artificial intelligence is reshaping how core operations are delivered. Finance cycles are shortening, HR interactions are increasingly automated, and procurement is becoming more digitized. Business services – once back-office functions – are evolving into enterprise platforms with efficiency, insight, and resilience.

However, this momentum is hitting a structural ceiling. Most AI systems still operate primarily in English. In a region where Arabic underpins contracts, customer interaction, and public administration, this limits scale and value.

The next productivity leap will not come from more automation, but from Arabic-first AI systems paired with local talent positioned to govern and scale them.

AI is not yet fluent enough

The region has made significant progress in consolidating finance, HR, IT, procurement, and customer operations into centralized business services. Expectations now extend beyond efficiency to quality at scale. AI is the key, yet it remains limited by language and context.

Large language models are trained predominantly on English data. In major datasets such as Common Crawl's CC100, which is used to train mainstream LLMs, english accounts for roughly 82GB of text compared to just 5.4GB in Arabic, despite more than 400 million Arabic speakers.[1] Dialect diversity, code-switching, and linguistic complexity widen the gap between capability and performance.

The impact is evident across the GCC. Pilots demonstrate potential but struggle to scale. Chatbots require frequent human intervention, document-processing tools misinterpret text, and automation often merely shifts effort without eliminating it. Generic Arabic-capable models achieve only around 30-50% accuracy on Arabic financial text.[2]

This is not due to AI capability, but a limitation of enterprise-grade Arabic fluency. More mature ecosystems such as China have reduced this gap through investment in native-language models, data and deployment.

How Arabic-first AI changes the economics

When models are not natively fluent, organizations compensate through supervision, rework, and parallel manual processes, eroding the advantage of automation. More than 20% of Arabic chatbot queries still require human handover.[3]

Arabic-native LLMs are still nascent, and mostly limited to front-end conversational layers. AI has struggled to move into end-to-end workflow automation, where impact compounds across processes.

Improved language alignment can facilitate a shift that is already underway. IBM's watsonx Orchestrate, for example, automates finance journal entries to generate approximately $600,000 in annual savings, reduces HR costs by around 40 percent, and enables a tenfold acceleration of supplier onboarding. AI operates here as an integrated execution layer across workflows.

The outcome would be a different operating model. Business services platforms that process and act in Arabic, shaped by talent at the right control points, shift from cost centers to strategic assets. The GCC has the opportunity to define this model for the rest of the world, devising a system where AI is most effectively governed, adapted, and trusted in real-world operations.

[1] Common Crawl, CC100 dataset is a major dataset used to train mainstream LLMs such as LLaMA (Meta); https://commoncrawl.github.io/cc-crawl-statistics/plots/languages

[2] Domain Adaptation for Arabic Machine Translation: Financial Texts as a Case Study; https://www.mdpi.com/2076-3417/14/16/7088

[3] Bilingual AI-Driven Chatbot for Academic Advising; https://thesai.org/Publications/ViewPaper?Volume=13&Issue=8&Code=IJACSA&SerialNo=8

[4] https://www.mckinsey.com/capabilities/quantumblack/our-insights/how-businesses-can-close-chinas-ai-talent-gap; https://global.chinadaily.com.cn/a/202408/19/WS66c281e0a31060630b923a83.html

---

Wissam Abdelsamad
Partner
wissam.abdelsamad@strategyand.pwc.com

Gaurav Pagaria
Principal
gaurav.pagaria@strategyand.pwc.com

www.strategyand.pwc.com/me