

Vinuni ra mắt V-Bench - thước đo năng lực của các mô hình ai dành cho người Việt

09:41 06/07/2026

Hà Nội, ngày 06/07/2026 - Trung tâm Nghiên cứu AI - Trường Đại học VinUni chính thức giới thiệu V-Bench - bộ công cụ phi lợi nhuận được phát triển nhằm phục vụ cộng đồng và thúc đẩy sự phát triển AI có trách nhiệm tại Việt Nam. Dựa trên bộ dữ liệu chuyên sâu hơn 40.000 câu hỏi và tác vụ, được thẩm định bởi đội ngũ chuyên gia uy tín trong và ngoài nước, V-Bench có thể đánh giá toàn diện năng lực Tiếng Việt của các mô hình ngôn ngữ lớn, từ đó đưa ra hệ tham chiếu khách quan, độc lập, làm cơ sở cho



The banner features a dark blue background. On the left, the V-BENCH logo is at the top. Below it, the title 'Thước đo năng lực cho AI tiếng Việt' is displayed in white and yellow. A subtitle explains its purpose: 'Đánh giá toàn diện năng lực LLM tiếng Việt – từ tri thức học thuật, văn hóa vùng miền đến an toàn AI và hành vi agent phức tạp.' Two categories are listed: 'Năng lực tri thức' (Knowledge) with sub-points 'Tri thức - Văn hóa - An toàn AI' and 'Năng lực agent' (Agent) with sub-points 'Hành vi Agent - Tool Use - Systems'. Three statistics are shown: '40.000+' questions, '15' models, and '5' groups. On the right, there is a graphic of a glowing blue lotus flower inside a sphere, surrounded by floating data cards. The VinUniversity logo is at the bottom left.

V-BENCH

Thước đo năng lực cho AI tiếng Việt

Đánh giá toàn diện năng lực LLM tiếng Việt – từ tri thức học thuật, văn hóa vùng miền đến an toàn AI và hành vi agent phức tạp.

Năng lực tri thức
Tri thức - Văn hóa - An toàn AI

Năng lực agent
Hành vi Agent - Tool Use - Systems

40.000+
Câu hỏi đánh giá

15
Mô hình đã đánh giá

5
Nhóm dữ liệu

VINUNIVERSITY
Center for AI Research

Được khởi xướng bởi Trung tâm Nghiên cứu AI - Trường Đại học VinUni, với sự tham gia cộng tác hoàn toàn phi lợi nhuận của hội đồng khoa học gồm 18 chuyên gia người Việt đầu ngành AI trong và ngoài nước, V-Bench là một trong những công cụ đo lường khả năng xử lý tiếng Việt của các mô hình ngôn ngữ lớn (LLM - Large Language Model) toàn diện nhất hiện nay. Không chỉ dựa trên những kiến thức học thuật, bộ dữ liệu hơn 40.000 câu hỏi và tác vụ của V-Bench được xây dựng theo cả chiều rộng lẫn chiều sâu, bao phủ nhiều lĩnh vực và phản ánh đầy đủ các yếu tố về văn hóa, ngôn ngữ, chủ quyền, an toàn thông tin và đặc thù địa phương.

V-Bench chia bộ công cụ theo 2 hạng mục lớn với 5 nhóm đánh giá chuyên biệt, bao gồm: Văn hóa ngầm định - đánh giá khả năng thấu hiểu các tầng nghĩa văn hóa ẩn sâu trong ngôn ngữ Việt như phong tục, tín ngưỡng, lịch sử; Đa dạng vùng miền - kiểm tra hiểu biết về phương ngữ Bắc - Trung - Nam, cách xưng hô và văn phong đặc trưng từng vùng; An toàn và chủ quyền số - đảm bảo tuân thủ các nguyên tắc an toàn thông tin quốc gia, chống thông tin sai lệch và bảo vệ chủ quyền số Việt Nam; Ứng dụng thực tiễn - đánh giá hiệu quả trong giáo dục, y tế cộng đồng và các lĩnh vực ứng dụng cụ thể của Việt Nam; Năng lực agentic AI - đánh giá khả năng lập kế hoạch, sử dụng công cụ, RAG và hành động tự chủ trong môi trường Việt Nam.

Nhờ đó, V-Bench không chỉ đo lường khả năng trả lời câu hỏi mang tính học thuật, mà còn đánh giá sự thấu hiểu văn hóa, ngôn ngữ vùng miền và khả năng hành động của AI trong bối cảnh Việt Nam, từ đó cho ra kết quả khách quan, toàn diện, sát với thực tế, có tính tham chiếu và giá trị định hướng cao cho cộng đồng. Đặc biệt, kết quả đánh giá bởi V-Bench được bảo chứng bởi hội đồng chuyên gia uy tín cùng phương pháp luận nghiêm ngặt và bộ dữ liệu có quy mô đồ sộ, khoa học, công phu nhất hiện nay.

Được xây dựng theo các kỹ thuật tương đồng và phổ biến trong nghiên cứu LLM quốc tế, V-Bench cũng có khả năng tương thích cao về thuật ngữ và quy trình với hệ sinh thái benchmark quốc tế, giúp các nhà phát triển LLM toàn cầu có thể tham gia đánh giá mà không gặp rào cản về quy ước kỹ thuật.



“

V-Bench mong muốn góp phần định nghĩa lại tiêu chuẩn AI tại Việt Nam và xây dựng một hệ tham chiếu AI cho quốc gia.

GS.TS. Dương Nguyên Vũ
Phó Hiệu trưởng, Giám đốc Khoa học Trung tâm Nghiên cứu AI
- Trường Đại học VinUni



GS.TS. Dương Nguyên Vũ - Phó Hiệu trưởng, Giám đốc Khoa học Trung tâm Nghiên cứu AI Trường Đại học VinUni,

GS.TS. Dương Nguyên Vũ - Phó Hiệu trưởng, Giám đốc Khoa học Trung tâm Nghiên cứu AI Trường Đại học VinUni, cho biết: “Với bộ dữ liệu toàn diện cùng sự góp sức hoàn toàn khách quan, phi lợi nhuận của cộng đồng chuyên gia uy tín người Việt trên toàn cầu, V-Bench mong muốn góp phần định nghĩa lại tiêu chuẩn AI tại Việt Nam và xây dựng một hệ tham chiếu AI cho quốc gia. Kết quả đánh giá của V-Bench sẽ mang tới góc nhìn toàn cảnh, giúp các tổ chức, doanh nghiệp khám phá chiều sâu năng lực của từng mô hình AI, từ đó lựa chọn mô hình tối ưu cho từng bối cảnh sử dụng và nhu cầu thực tế”.

Được cung cấp miễn phí tại địa chỉ <https://vbench.ai/>, V-Bench hiện đã tiến hành đánh giá và công bố kết quả tham chiếu của 15 mô hình LLM tiêu biểu hiện nay. Trong thời gian tới, V-Bench sẽ tiếp tục được phát triển để bổ sung các tiêu chí đánh giá mới dựa trên hình ảnh và âm thanh như: chữ Nôm, biển hiệu vùng miền, biểu đồ, di sản văn hóa, nhận diện và hiểu giọng nói ba miền, hiểu ngữ cảnh văn hóa qua video ngắn, hiểu tài liệu dài tiếng Việt trên 100.000 token như luật, hợp đồng, sách giáo khoa và trích dẫn chính xác từ văn bản pháp quy, công báo...

Là một trong những dự án nghiên cứu AI tiêu biểu của Trung tâm Nghiên cứu AI - Trường Đại học VinUni, V-Bench là minh chứng cho khát vọng bứt phá, nâng tầm vị thế trí tuệ Việt trên bản đồ học thuật toàn cầu của VinUni, đồng thời khẳng định sứ mệnh tiên phong, tinh thần cống hiến vì cộng đồng và phụng sự khoa học của Tập đoàn Vingroup./.

Tuấn Minh