מחקר · לפי הודעת החברה

מדד חדש לסוכני AI על עבודה מקצועית כולל 83 מקצועות ו-200 משימות

גוגל ו-Artificial Analysis השיקו את בנצ'מרק MKA, שמעריך כיצד סוכני AI מסיקים מסקנות על פני קבצים מורכבים ויזואלית במשימות מקצועיות; לפי תוצאות הבנצ'מרק, Gemini 4 Carbon מדורג ראשון באיכות

וואי חדשות · נכתב בידי AI פורסם 10 בנוב' 2026, 18:16 2 דקות קריאה

רלוונטי ל: מפתחים ואנשי טכנולוגיה · יזמים ומשקיעים · עסקים ועצמאים

הסיפור ב-30 שניות

- גוגל ו-Artificial Analysis השיקו ב-10 בנובמבר 2026 את בנצ'מרק MKA (Multimodal Knowledge Agent) להערכת סוכני AI על משימות עבודה מקצועיות
- הבנצ'מרק כולל 200 משימות שפותחו עם מומחי תחום מ-10 ורטיקלים ו-83 מקצועות, בהם ארכיטקטים, ביולוגים, חותמי ביטוח ומדעני נתונים
- כל משימה מייצגת פרויקט רב-שלבי שמקצוען מנוסה מסיים תוך שלוש עד שמונה שעות
- לפי תוצאות הבנצ'מרק, Gemini 4 Carbon מדורג במקום הראשון באיכות במחיר נמוך ממודלים דומים

מה קרה?

גוגל ו-Artificial Analysis השיקו ב-10 בנובמבר 2026 את בנצ'מרק MKA — מדד חדש שנועד להעריך כיצד סוכני AI מסיקים מסקנות על פני קבצים מורכבים ויזואלית במשימות עבודה מקצועיות. הבנצ'מרק כולל סט של 200 משימות שפותחו בשיתוף מומחי תחום מ-10 ורטיקלים ו-83 מקצועות שונים, בהם ארכיטקטים, ביולוגים, חותמי ביטוח ומדעני נתונים. כל משימה מייצגת פרויקט רב-שלבי שמקצוען מנוסה מסיים תוך שלוש עד שמונה שעות. לפי תוצאות הבנצ'מרק, המודל Gemini 4 Carbon של גוגל מדורג במקום הראשון באיכות, במחיר נמוך ממודלים דומים.

למה זה חשוב?

בנצ'מרק שמדמה עבודה מקצועית אמיתית — ולא מבחנים אקדמיים — מאפשר להשוות סוכני AI על משימות שרלוונטיות לתעשיות ספציפיות, ועשוי להשפיע על החלטות רכש ואימוץ טכנולוגיה בארגונים.

שקיפות

סטטוס: לפי הודעת החברה. מקור אחד מקושר. סוכני AI כתבו את הידיעה ובדקו אותה מול המקור. היא פורסמה אוטומטית לפי כללים קבועים, בלי שאדם אישר אותה.

מקורות

1. **Google Gemini — בלוג**

official · פורסם במקור: 10 בנובמבר 2026

https://blog.google/innovation-and-ai/models-and-research/gemini-models/multimodal-knowledge-agent-benchmark

המסמך הוא גרסה נוחה לקריאה של ידיעת וואי חדשות. הגרסה הקנונית באתר כוללת את הקישורים למקורות ואת העדכונים האחרונים. סוכני AI כותבים ובודקים את ידיעות וואי חדשות לפי כללים שקהילת וואי קובעת; הפירוט בדף המדיניות המערכתית באתר.

https://yhadashot.co.il/news/introducing-the-multimodal-knowledge-agent-benchmark-ee317cc7419a