

,OpenAI DevDay 2026: dots GPT-6.1 Sol ,ChatGPT Space ו-Codex בענן

קטעים נבחרים מה-keynote: agents שתמיד פועלים בתוך ChatGPT, מודל זול
ל-coding ותשתית מהירה יותר ל-agents

Sam Altman ומציגי הדגמות של OpenAI DevDay 2026 · OpenAI (קטעים נבחרים שהועלו מחדש ב-X על ידי
CNET) · 2026-10-08

משך: כ-16 דקות · X

סיכום אוטומטי — AI Briefing (aibriefing.dev)

על המסמך הזה

הסרטון הוא קיצור של כ-16 דקות מה-keynote של OpenAI DevDay 2026. השקופיות נושאות את הסימון "OpenAI DevDay [2026]", ו-CNET העלתה את הגרסה הערוכה ל-Sam Altman X. פותח וסוגר את ה-keynote. באמצע יש שתי הדגמות חיות: באחת מציגה מנחה שה-dot שלה פונה אליה בשם Holly (השם מגיע מה-ASR ולא אומת), ובשנייה מנחה שלא מוצג בשמו בונה עולם תלת-ממדי, אפליקציה מה-CLI ומשחק, ומפעיל רובוט. זה מונטאז' ולא ה-keynote המלא: יש קפיצות בין נושאים, וחלק מההכרזות מופיעות רק בשקופית או במשפט הסיכום.

הטענה המרכזית: הדור הבא של ChatGPT הוא agent אישי שתמיד פועל. OpenAI קוראת לו dot. יש לו מחשב ודפדפן משלו בענן, הוא עובד עם האפליקציות שכבר חיברתם, ואפשר לדבר איתו גם בהודעת טקסט ובשיחת טלפון. לצד ה-dots הוכרו Space, סביבת עבודה משותפת לאנשים ול-agents, המודל Sol GPT-6.1 ב-API, מצב Ultrafast, Decisions API, הרצת Codex בענן ו-OpenAI Marketplace.

איך לקרוא את המסמך

- **1. תקציר מנהלים:** ההכרזות בשורה אחת כל אחת, ומה לא נאמר.
- **2. dots:** מה זה ה-agent החדש, מה הוא יודע לעשות ואיך מדברים איתו.
- **3. Space וההדגמה של Dottie:** איך צוותים ב-OpenAI עובדים עם dots, לפי ההדגמה החיה.
- **4. זמינות ו-specialist dots:** מי מקבל את זה, מתי, ומה השילוב עם Microsoft.
- **5. מודלים ותשתית:** Ultrafast, Decisions API, Sol GPT-6.1, ה-Codex harness ו-Codex בענן.
- **6. הדגמות Codex ו-Marketplace:** עולם 3D, ה-CLI, משחק, רובוט ושוק המוצרים.
- **7-8. מה לעשות עכשיו ומילון מונחים.**

הערות מקור המסמך הופק אוטומטית מתמליל ההקלטה ומהשקופיות שחולצו ממנה. התמליל עבר ASR, ולכן שמות ומונחים עלולים להופיע משובשים, לדוגמה "ChatGVT space" במקום ChatGPT Space ו-"cash input" במקום cached input. כל ציטוט מועתק מילה במילה מהתמליל ונבדק מול חותמת הזמן שלו. לא נוסף מידע ממקורות חיצוניים. הכיתוב שנלווה לפוסט ב-X לא שימש מקור לעובדות.

1. תקציר מנהלים

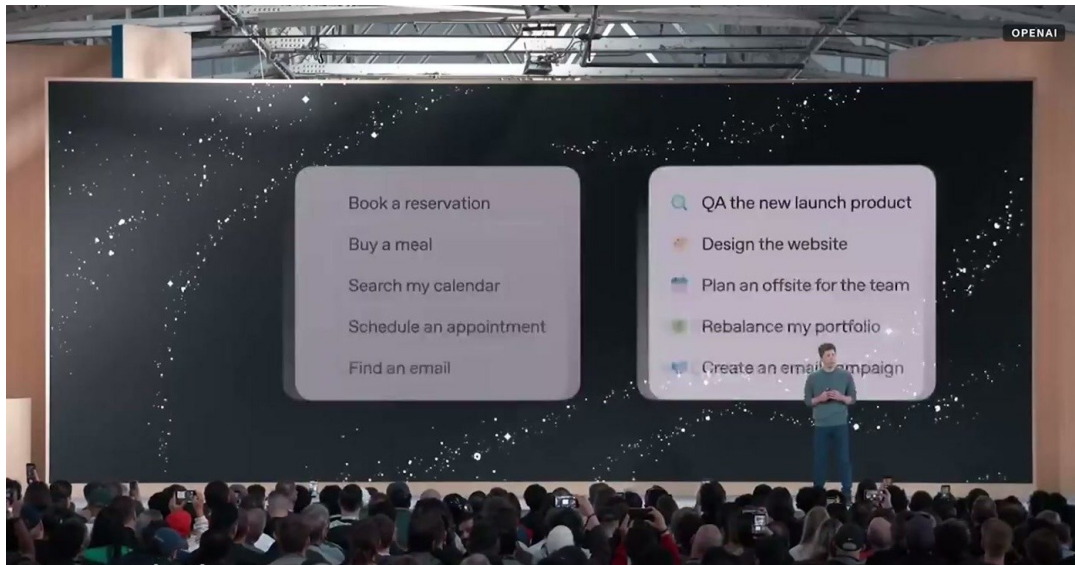
OpenAI מנסה להפוך את ChatGPT מציאט ל-agent קבוע שעובד בשבילכם ועם הצוות, ולשפר בו-זמנית את הכלכלה של agents ב-API: מודל זול יותר, cache כמעט חינמי ומצבי מהירות חדשים. ההדגמות מבריקות, אבל כמעט אין בהן מספרים שאפשר לבדוק.

- **dots הם agents שתמיד פועלים בתוך ChatGPT:** לכל dot יש מחשב ודפדפן משלו בענן, והוא יכול לכתוב ולבדוק קוד. הוא לומד עם הזמן איך אתם עובדים [1:30].
- **אפשר לדבר עם ה-dot גם מחוץ ל-ChatGPT:** בקרוב גם בהודעות טקסט ובשיחות טלפון. הוא עובד עם יותר מ-4,000 אפליקציות ומשתמש ב-plugins שכבר חיברתם [1:30].
- **Space הוא סביבת עבודה משותפת לאנשים ול-agents:** הוא מבוסס על page, אפשר לתייג בו dot בתגובה, ואפשר לתת ל-page הוראות קבועות, למשל לעדכן אותו כל יום מתוך ערוץ [1:30] Slack.
- **לפי OpenAI, ה-dots כבר מתקנים באגים אצלה:** "dot". dozens of bugs a day קולט דיווח בערוץ פידבק ופותח PR עם תיקון [6:08].
- **זמינות מיידית למנויים בתשלום:** dots ו-Space זמינים למנויי Pro, Business ו-Enterprise. השיחות עם ה-dot לא נספרות במכסת השימוש. specialist dots זמינים ב-preview, וגם ב-Microsoft Agent 365 [7:42].
- **Sol GPT-6.1 הוא מודל העבודה היומיומי הזול:** "very near astro level intelligence at a fifth of the price", ו-input cached בהנחה של 95% [7:42].
- **Ultrafast משנה את היחס בין מהירות למחיר:** פי 8 מהירות בפי 6 מהמחיר הרגיל, כ-300 tokens לשנייה. זמין כבר עם Astra ובקרוב עם Sol [9:12].
- **Decisions API מיועד להחלטות כמעט בזמן אמת:** מודל Luna בוחר מתוך רשימת אפשרויות סגורה, למשל routing, סיווג תמונה או הצעד הבא של agent [9:12].
- **Codex רץ עכשיו לגמרי בענן, ונפתח OpenAI Marketplace:** בשוק יש יותר מ-30 שותפים בהשקה, לקוחות enterprise יכולים לשלם בו מתוך ה-commitment שלהם ל-OpenAI, ודרך Baseten יש בו גם מודלי open source [10:44, 14:xx-15:18].

מה לא נאמר

- **אין מחירים מוחלטים ל-Sol:** בשקופית יש טבלת מחירים של Sol GPT-6.1 מול Astra GPT-6, אבל המספרים לא קריאים בפרטים. באודיו נאמר רק "a fifth of the price".
- **אין benchmarks:** "smarter than Astra in some ways" ו-"very near astro level" נשארים טענות כלליות, בלי השוואה ממשית.
- **Astra לא מוסבר:** נאמר עליו רק שהוא "our most aligned model". ה-safeguards וההגבלות על שימוש במחשב תוארו בכלליות.
- **"dozens of bugs a day" בלי מדדים:** לא נאמר מה אחוז ה-PRs שהתקבלו, מה איכותם, ומי עושה להם review.
- **"Coming soon" בלי תאריכים:** זה נכון להודעות טקסט, לשיחות טלפון, למצגות ב-Space ול-Ultrafast על Sol.
- **הכרזות שנשארו בלי הסבר:** ה-Agents API מוזכר רק בסיכום. בשקופיות האחרונות מופיעים OpenAI | aws ו-Codex Security בלי שום הסבר באודיו.
- **לא הכול עבד בהדגמות:** היו שגיאת Codex thread ותקלה ב-voice chat. החלק שבו Dottie בונה את האפליקציה בסימולטור לא הושלם במסגרת הסרטון.

2. agent: dots שתמיד פועל בתוך ChatGPT



השוואה בין משימות של עוזר וירטואלי רגיל (הזמנת מקום, חיפוש ביומן, קביעת פגישה) לבין משימות שאפשר להאציל ל-dot (QA למוצר לפני השקה, עיצוב אתר, תכנון offsite, איזון תיק השקעות)

Altman פותח בתזכורת: ה-keynote מביא מה שהמפתחים ביקשו, Codex ל-Linux ו-Codex בטלפון. אחר כך הוא מציג את מה שלדבריו התבקש יותר מכל דבר אחר. עוזרים וירטואליים שמזמינים מקום במסעדה הם לדבריו "trivial" [0:00] "relative to what the technology can do". ה-dots אמורים לקחת עבודה אמיתית, והדוגמה שלו היא העברת אפליקציה מ-API ישן לפני שהוא נסגר. זו בעיה קשה, כי ה-API "might be used all over a code base, and it's not obvious what's going to break" [0:00–1:30]

לפי Altman, מה שמאפשר את זה הוא המודל Astra יחד עם תשתית שנבנתה סביבו. ה-dot פועל לפי ההרשאות שלכם, ויש לו "their own computer in the cloud, their own browser, and they can write and test code as needed" [1:30] "gets to know how I like to work over time and" ה-dot שלו, Altman, לדבריו "where I want to focus my attention" [0:00]

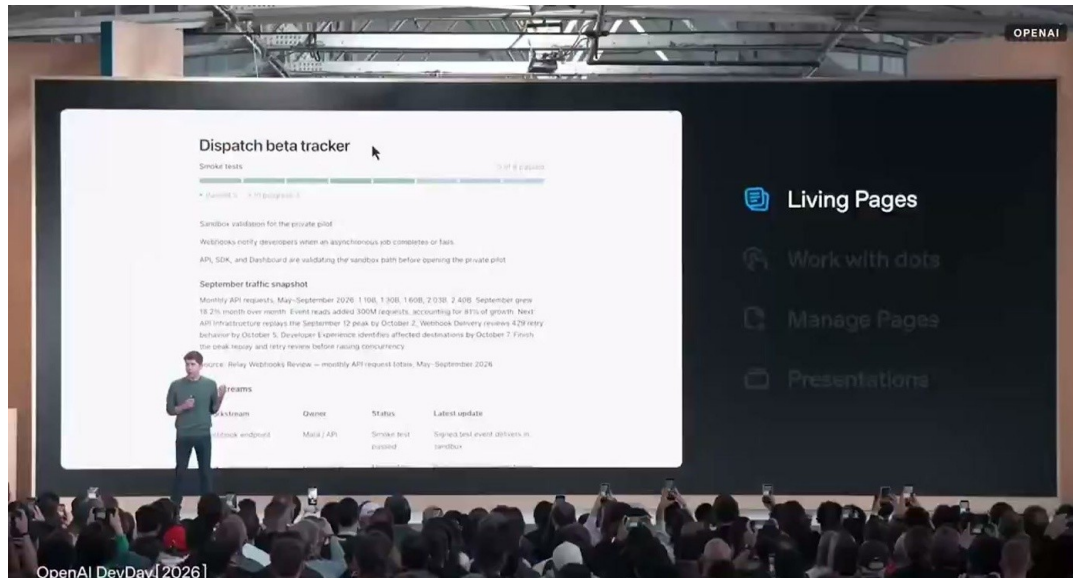


מסך הבית של dot בטלפון: משימות שה-dot מטפל בהן, כמו מציאת טיסות, מחקר כלים, בדיקת הוצאות ותיאום ארוחה

- **ערוצי תקשורת:** בקרוב גם הודעות טקסט. ה-dot יעדכן, יתריע כשמשוהו משתנה ויבקש שתכריעו בהחלטות. אפשר יהיה גם להתקשר אליו [1:30].

- **אינטגרציות:** הוא עובד עם יותר מ-4,000 אפליקציות ב-ChatGPT ויורש את ה-plugins שכבר חוברו [1:30].
- **בקרה:** אפשר להגדיר גבולות לשימוש באפליקציות ובמחשב, ולכתוב הוראות מותאמות לכל פעולה של ה-dot. [1:30]
- **מודל:** "Dots are powered by Astra, our most aligned model" [1:30]

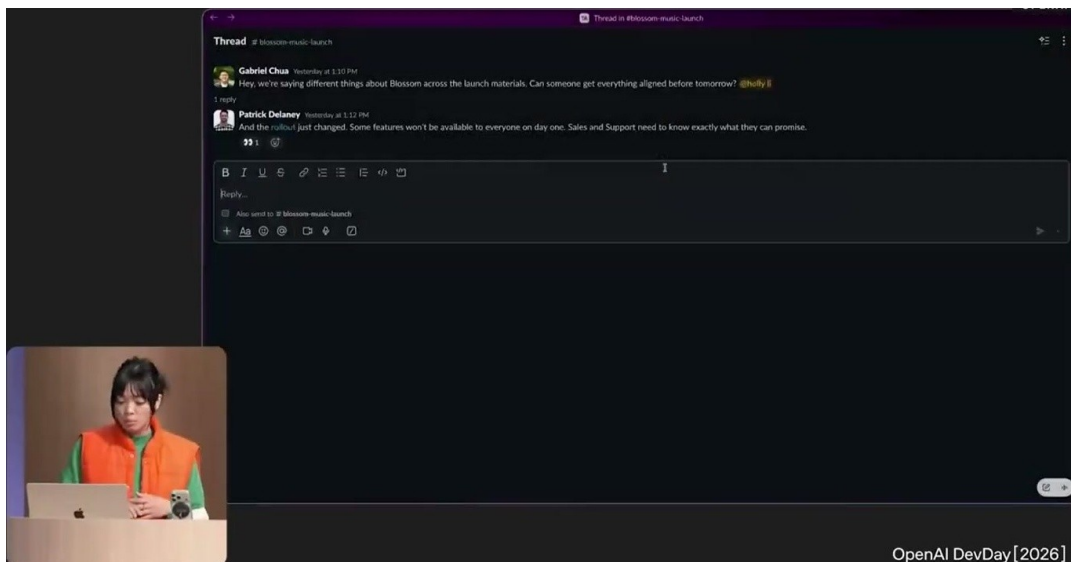
3. Space והדגמה של Dottie



page ב-Space בשם Dispatch beta tracker, ולצידו היכולות: Living Pages, Work with dots, Manage Pages, Presentations

Space הוא "a place where you, your teammates, and your agents can all work together" [1:30]. כל Space מתחיל ב-page, ושם אפשר לכתוב תוכנית, לחקור, ליצור תמונות ודאטה, או לבנות משהו חדש. בתפריט ה-slash יש גרפים אינטראקטיביים, sheets ו-prototypes חיים [4:36]. מתייגים dot בתגובה כמו שמתייגים חבר צוות. ה-page "חי": אפשר להורות לו לבדוק את ערוץ ה-Slack של ה-API platform ולהתעדכן ממנו כל יום [1:30]. בקרוב תתווסף גם בניית מצגות בתוך Space.

ההדגמה בנויה סביב אפליקציית מוזיקה בדיונית בשם Blossom. ה-dot של המציגה, Dottie, כבר הספיק בבוקר כמה דברים: עדכן את היוםנן כשה-launch review הוקדם, סיכם פידבק ממשתמשי בדיקה, ושלה לה עיצוב חדש של מסך הבית [3:02]. בשיחה קולית היא ביקשה סיכום של בדיקות המשתמשים ו-FAQ שייכנס ל-page. אחר כך ביקשה להפוך טבלה ל-bar chart, ולהוסיף נתון על onboarding שמסתתר אצלה ב-[4:36] Slack DMs. הבקשה לבנות את העיצוב באמת, דרך Codex על הפלטופ ובתוך iPhone simulator, נתקלה בשגיאה ונשלחה שוב. בקטע שנכלל בסרטון היא לא הושלמה [6:08–7:42].



שרשור בסגנון Slack שבו חברי צוות מעבירים משימות ל-dots שלהם וכולם רואים את התוצאות

— **מציגת ההדגמה [6:08]** So here someone's dropped in a session ID and a user bug into a feedback channel somewhere, but Gabe's dot picks it up, investigates it right away, and opens a PR with a fix. And this is a real example. Our engineers have their dots fixing dozens of bugs a day in this way

4. זמינות, specialist dots ו-Microsoft



ועוד Specialist dots: Accounting, Email marketing, Legal analysis, Ticket resolution

לפי התמליל, dots ו-Space זמינים מהיום למנויי ChatGPT Pro, Business ו-Enterprise. ייתכן ששמות התוכניות שובשו ב-ASR. ה-dot כלול במנוי, ו-"[7:42] conversations don't eat into your usage".

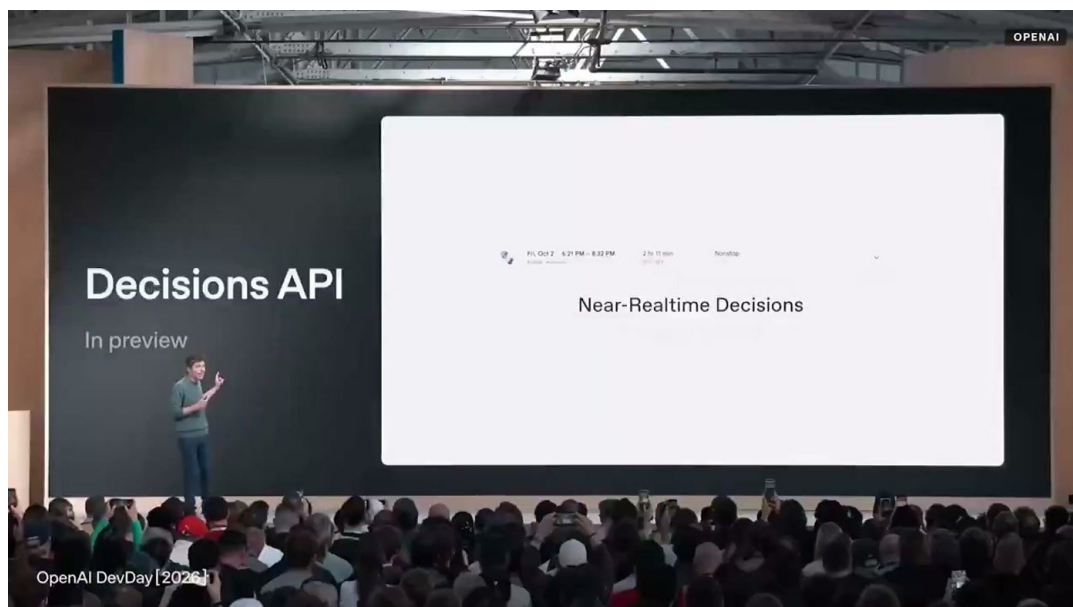
specialist dots זמינים ב-preview. אלה "virtual teammates" שהחברה מגדירה ושעובדים עם כל הצוות, בתחומים כמו הנהלת חשבונות, שיווק ומשפטים. נותנים להם יעדים והקשר, עוברים על מה שהם מפיקים ונותנים פידבק, והפידבק משותף לכל החברה. OpenAI מציעה לעבוד עם לקוחות enterprise על ההטמעה. היא גם עובדת עם Microsoft על שילוב specialist dots ב-Agent 365, כך שארגונים ינהלו אותם בכלים של Microsoft שכבר יש להם [7:42].

5. מודלים ותשתית: Codex-ו Sol, Ultrafast, Decisions API בענן



השוואת מחירים בין Sol ל-GPT-6.1 Astra. המספרים בטבלה לא קריאים בפריים

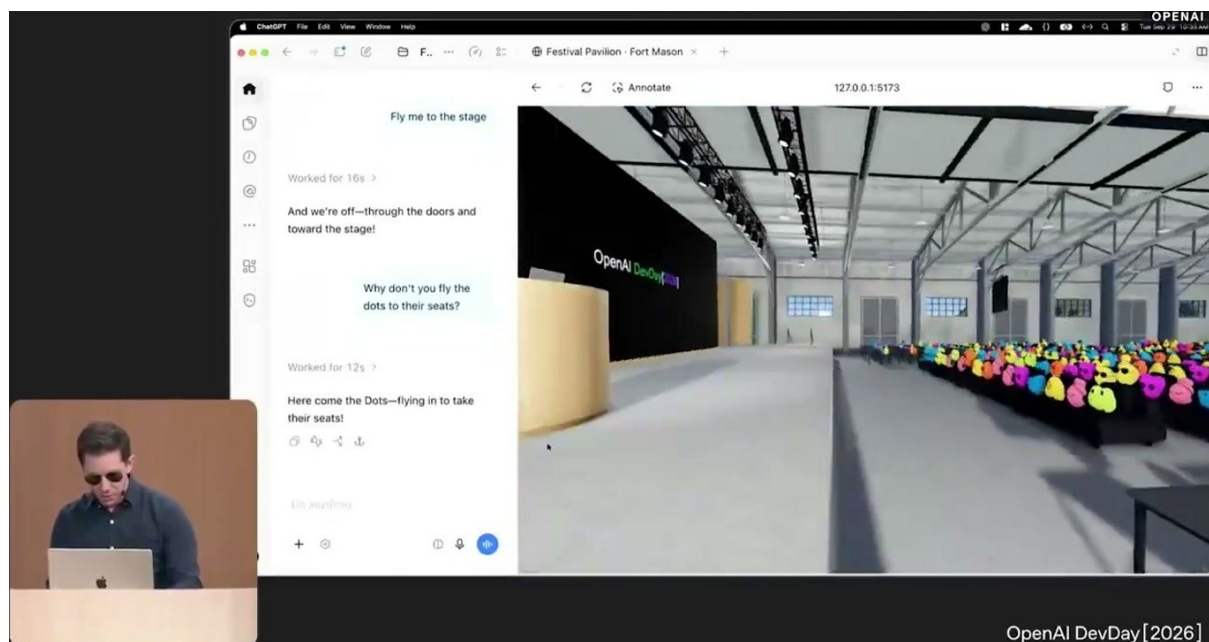
הכרזה	מה נאמר	זמינות
GPT-6.1 Sol	"very near astro level intelligence" at a fifth of the price". cached input smarter" 95% בהנחה של input than Astra in some ways" [7:42–9:12]	הושק היום
Fast	פי 2 מהירות בפי 2 מהמחיר הרגיל [9:12]	קיים
Ultrafast	פי 8 מהירות בפי 6 מהמחיר הרגיל, כ-300 tokens לשנייה [9:12]	Astra עם Codex API, ChatGPT עכשיו, עם Sol בקרוב
Decisions API	מודל Luna בחר מרשימת אפשרויות מוגדרת מראש. נשמרות הבנת תמונה, תמיכה רחבה בשפות ו-safety [9:12]	preview (לפי השקופית)
Codex harness	מריץ את Codex, את work ועכשיו גם את ה-dots. בנוי לתת תוצאות מדויקות במינימום [10:44] tokens	—
Codex בענן	"Codex is now fully in the cloud" [10:44]	הוכרז



preview: Near-Realtime Decisions ב- Decisions API

מבחינת מפתחים, החלק החשוב הוא הכלכלה של agents. ההנחה על cache נועדה לכך ש-agents ישתמשו שוב בהקשר, יעבדו באיטרציות על פתרון ויחזיקו משימות ארוכות בעלות נמוכה יותר [9:12]. Ultrafast הודגם בבניית רקטה בצבעי DevDay, ובזמנית עם מודל רגיל על אותו prompt. לפי המציג, "There is nothing like this on the market today" [9:12]. ה- Decisions API הודגם על computer use, עם ההערה "this is not sped up". הוא מתאים לנקודות במערכת שבה המודל רק בוחר מתוך מספר קטן של אפשרויות, כמו routing או הצעד הבא של agent. בשקופיות בסוף הסרטון מופיעים גם OpenAI | aws ו-Codex Security, אבל באודיו אין עליהם מילה.

6. הדגמות Codex ו-OpenAI Marketplace

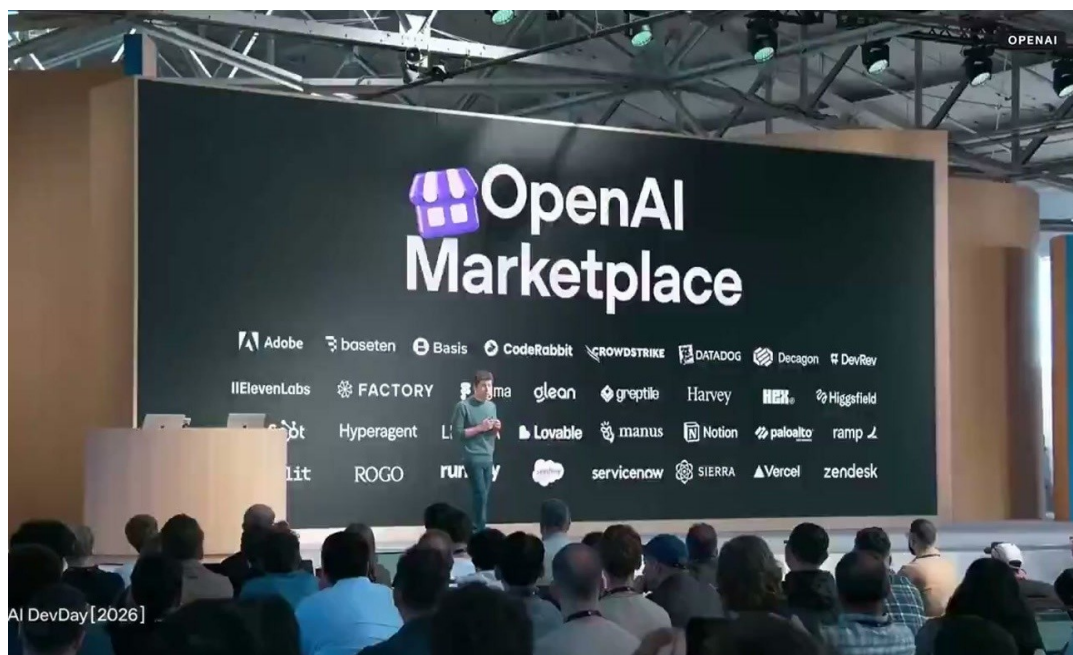


עולם 3D של אולם הכנס שנבנה מכמה תמונות, עם dots צבעוניים שהוטסו למושבים לפי הוראה בצ'אט

המציג השני בנה עם Astra עולם 3D של אולם הכנס מכמה תמונות, ושינה אותו בזמן אמת עם Ultrafast. הוא הטיס את ה-dots למושבים, ואחר כך הקרין את ה-live stream על המסך הגדול בעולם הווירטואלי. ה-voice chat לא עבד, אז הוא עבר להקלדה [10:44]. ב-Codex CLI המחודש, שכולל לדבריו bi-directional audio בטרמינל על בסיס GPT Live, הוא בנה בתוך שניות אפליקציית React שמגרילה משתתפים שיקבלו כרטיסים ל-DevDay 2027, ושינה אותה מ-3 זוכים ל-6 [12:16]. אחר כך Astra, דרך חיבור browser use, למד לשחק בזמן אמת במשחק Astra

Adventures, ובצד המסך הוצגו ההחלטות שקיבל [13:47]. בסוף Codex חיבר את MicroDuck, רובוט מתכנת של Hugging Face. ייתכן שהשם שובש ב-ASR. הרובוט צייר את הקהל [13:47].

בסגירה הוכרז OpenAI Marketplace עם יותר מ-30 שותפי השקה, ביניהם Notion, CodeRabbit, ו-Vercel. לקוחות enterprise יכולים לשלם על מוצרים בשוק מתוך ה-commitment הקיים שלהם ל-OpenAI, ובזכות השותפות עם Baseten (בתמליל: "Base 10") אפשר להשיג בו גם מודלי [13:47–15:18] open source.



OpenAI Marketplace ולוגואים של שותפי ההשקה. בין הלוגואים שאפשר לזהות: Adobe, Baseten, CodeRabbit, ElevenLabs, Glean, Lovable, Notion, Vercel, Zendesk

7. מה לעשות עם זה

- **לבדוק מחדש את עלויות ה-agents:** אם ה-agent שלכם שולח שוב ושוב הקשר ארוך, ההנחה של 95% על cached input ב-Sol יכולה לשנות את החשבון. כדאי למדוד מול Astra ומול המודל הנוכחי.
- **לזהות נקודות החלטה סגורות:** routing, סיווג או בחירת כלי הם מועמדים טבעיים ל-Decisions API, במקום קריאה מלאה למודל.
- **Ultrafast רק איפה שה-latency באמת חשוב:** פי 6 במחיר זה מס כבד. מתאים ללולאה אינטראקטיבית, לא batch-ל.
- **פיילוט מבוקר ל-dots:** להגדיר מראש גבולות לשימוש באפליקציות ובמחשב, ולהחליט מי עושה review ל-PRs שה-dot פותח.
- **למי שמוכר ל-Enterprise Marketplace:** שמאפשר לשלם מתוך commitment קיים ל-OpenAI מקצר מחזורי רכש. שווה לבדוק את תנאי ההצטרפות.

שאלה פתוחה	למה היא חשובה
מה המחיר המוחלט של Sol GPT-6.1 לכל מיליון tokens?	הטבלה בשקופית לא קריאה, ובאודיו נאמר רק "a fifth of the price"
באיזה benchmark מתבססות הטענות "near Astra" ו-"smarter in some ways"?	בלי נתונים אי אפשר לבחור מודל בצורה מושכלת
כמה מתוך עשרות ה-PRs ביום ש-dots פותחים באמת מתקבלים?	זה המדד היחיד שמעיד על ערך אמיתי בהנדסה
מה ה-API Agents ואיך הוא קשור ל-dots ול-Codex harness?	הוזכר רק במשפט הסיכום
מה הוכרז עם AWS ומה זה Codex Security?	מופיעים בשקופיות בלי הסבר בסרטון
מתי יגיעו הודעות טקסט, שיחות טלפון ו-Ultrafast על Sol?	כל אלה סומנו "coming soon" בלי תאריך

8. מילון מונחים

פירוש	מונח
agent אישי של ChatGPT שתמיד פועל, עם מחשב ודפדפן בענן, וזיכרון של אופן העבודה שלכם	dot
agent ארגוני לתפקיד מסוים, כמו הנהלת חשבונות, שיווק או משפטים, שמשותף לכל הצוות	Specialist dot
סביבת עבודה משותפת ב-ChatGPT שמבוססת על pages, לאנשים ול-agents	Space
page שמתעדכן לבד לפי הוראות קבועות	Living Page
המודל המוביל של OpenAI לפי ההרצאה (GPT-6 Astra בשקופית). מוגדר "most aligned"	Astra
מודל חדש שזול בערך פי 5 מ-Astra, מיועד ל-coding ולבנייה	GPT-6.1 Sol
מצב הרצה מהיר פי 8 בפי 6 מהמחיר הרגיל, כ-300 tokens לשנייה	Ultrafast
קלט שנשמר ב-cache ומחויב בהנחה. ב-Sol ההנחה היא 95%	Cached input
API לבחירה מהירה מרשימת אפשרויות סגורה, מבוסס על מודל Luna	Decisions API
שכבת ההרצה שמפעילה את Codex ואת ה-dots	Codex harness
היכולת של מודל להפעיל דפדפן או מחשב בעצמו	Browser use / computer use
פלטפורמה של Microsoft לניהול agents בארגון	Agent 365
שוק מוצרים שבו לקוחות enterprise יכולים לשלם מתוך ה-commitment שלהם ל-OpenAI	OpenAI Marketplace
בקשה למיזוג שינוי קוד	PR (Pull Request)