IA locale pour les réunions : Hedy peut désormais fonctionner entièrement sur votre appareil

Hedy 3.2 vous permet d'exécuter toute la chaîne de traitement IA sur votre propre appareil. Résumés, notes, réponses de chat et suggestions en direct se font en local, afin que rien n'atteigne un serveur pour le traitement IA.

Publié par Julian Pscheid · 23 avril 2026 · Mis à jour 20 août 2026
Lire cet article en ligne: https://www.hedy.ai/fr/post/local-ai-meetings-hedy-3-2/

Une avocate et son client lors d'une consultation privée dans un cabinet juridique, avec un MacBook entre eux entouré d'un doux hologramme cyan et violet suggérant un traitement IA effectué sur l'appareil

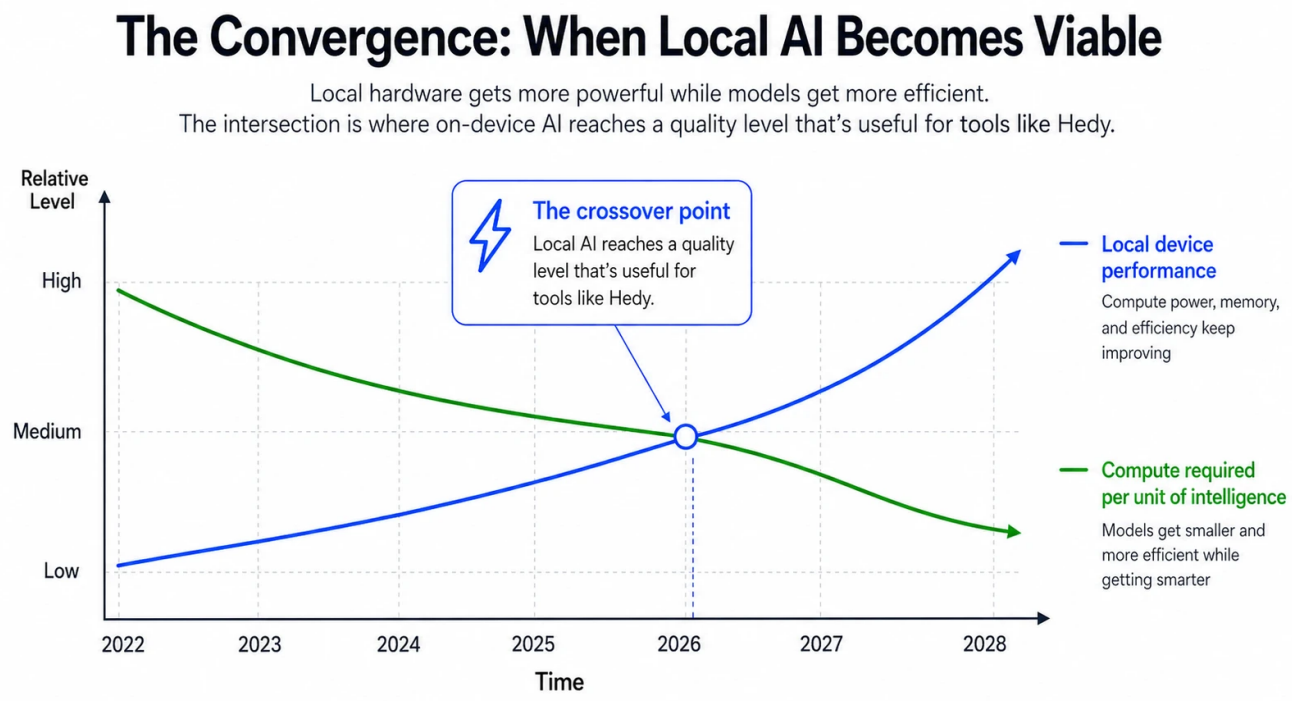
Avec Hedy 3.2, toute notre chaîne de traitement IA peut désormais fonctionner sur votre propre machine. Quand vous l'activez, les transcriptions de réunion, les résumés, les notes détaillées, les réponses de chat et les suggestions en direct se font tous sur l'appareil qui a capté l'audio. Rien de vos conversations n'est envoyé à un serveur. Voici ce qui change quand votre IA de réunion fonctionne ainsi.

La reconnaissance vocale fonctionne sur votre appareil depuis le lancement de Hedy. Les enregistrements audio sont toujours restés sur l'appareil eux aussi. Vos conversations n'ont jamais été utilisées pour entraîner des modèles d'IA. Nous avons été intentionnels sur la confidentialité dès le départ.

Mais il restait toujours un élément que nous ne pouvions pas encore ramener sur l'appareil : le travail d'IA lui-même. La partie qui lit votre transcription, rédige votre résumé, prend vos notes détaillées, répond à vos questions sur une réunion et propose des suggestions pendant que vous y participez. Ce travail

devait se faire sur des serveurs, parce que les modèles capables de le faire correctement étaient trop volumineux pour tourner sur un ordinateur portable ou un téléphone.

Cela évolue vite. Les appareils deviennent toujours plus puissants. Les modèles d'IA deviennent à la fois plus petits et plus intelligents. Il y a quelques mois, ces deux tendances se sont croisées pour nous. Les modèles qui tiennent sur un ordinateur portable ou un iPhone récent sont désormais assez solides pour gérer l'analyse de Hedy avec un niveau de qualité utile pour de vraies réunions.



Graphique conceptuel montrant la performance des appareils locaux qui augmente et la puissance de calcul nécessaire par unité d'intelligence qui diminue, avec un point de croisement situé autour de 2026 où l'IA sur l'appareil devient viable pour des outils comme Hedy

Illustratif — les courbes représentent la tendance générale, pas des mesures précises.

Avec Hedy 3.2, vous pouvez donc activer Local AI Processing et exécuter toute notre chaîne de traitement IA sur votre propre appareil. Résumés, notes, réponses de chat, suggestions. Tout se passe sur votre machine.

Pourquoi l'IA sur l'appareil compte

Pendant la majeure partie des dernières années, l'IA a été quelque chose qu'une poignée de grandes entreprises exploitent pour vous. Vous envoyez vos données à leurs serveurs, leurs modèles les traitent, puis les résultats reviennent. Ce modèle a des avantages évidents, et il a rendu possible une grande partie de ce que nous avons pu construire. Il a aussi un coût structurel : l'IA la plus utile a besoin des données les plus personnelles, et ces données vivent ailleurs que là où elles ont été générées.

L'IA locale inverse cette logique. Votre conversation reste là où elle a eu lieu. Le travail se fait sur le même appareil que celui qui a capté l'audio. Rien de la réunion n'atteint un serveur, sauf si vous activez Cloud Sync ou partagez explicitement la session, et même dans ce cas, le traitement IA reste local.

Si vous gardez Cloud Sync désactivé, rien de votre conversation n'arrive sur nos serveurs, à moins que vous ne partagiez explicitement cette session.

Qui bénéficie d'une IA de réunion privée sur l'appareil

Certains utilisateurs de Hedy l'activeront et n'y repenseront plus. D'autres attendaient quelque chose comme cela depuis des années. Voici qui, selon nous, en bénéficie le plus.

• Coachs et consultants dont les conversations avec les clients impliquent de fortes attentes de confidentialité. Ils utilisaient Hedy pour la préparation et les appels internes. Ils peuvent maintenant l'utiliser pendant le travail client sans que quoi que ce soit quitte l'ordinateur portable.

• Avocats qui utilisent Hedy pour les appels internes, mais jamais pour les conversations avec les clients. Le secret professionnel de l'avocat a une forme précise, et « nous promettons de traiter vos données avec soin » ne correspond pas à cette forme. Des données qui ne bougent pas, oui.

• Patients qui se rendent à des rendez-vous médicaux et veulent un récapitulatif clair de ce que leur médecin a dit, sans que leur discussion de santé se retrouve sur un serveur tiers. Avec l'IA locale, le récapitulatif peut se faire sur le même téléphone que celui qui a capturé la conversation.

• Journalistes qui travaillent sur des sujets sensibles et évitent entièrement les outils cloud. Ils peuvent capturer une interview, obtenir une transcription et discuter avec la réunion, sans que quoi que ce soit atteigne un serveur.

• Toute personne en dehors des États-Unis qui ne veut pas que ses conversations soient stockées sur des serveurs américains. Nous avons ajouté la résidence des données dans l'UE plus tôt cette année. L'IA locale va encore plus loin : les données ne résident sur les serveurs d'aucune entreprise.

• Travailleurs à distance avec un mauvais wifi en avion ou dans des zones rurales sans couverture. Hedy fonctionne désormais entièrement hors ligne. Ouvrez votre ordinateur portable pendant un vol, ayez une vraie conversation et obtenez le résumé avant l'atterrissage.

• Les curieux de confidentialité qui n'ont pas de profession réglementée ni de modèle de menace précis. Ils préféraient simplement l'idée que l'outil qui écoute leurs réunions n'envoie pas l'audio ailleurs. Ce n'était pas possible avant. Maintenant, ça l'est.

Ce qui relie ces personnes, c'est que Hedy leur était déjà utile en principe, mais que le modèle de données ne correspondait pas à leurs contraintes. L'IA locale supprime cette contrainte.

Si vous évaluez les preneurs de notes en général plutôt que de vous renseigner sur le nôtre, nous avons rédigé un guide distinct sur l'utilisation privée des notes de réunion IA (/fr/post/how-to-use-ai-meeting-notes-privately/) qui explique ce qui quitte votre appareil à chaque étape et comment vérifier les affirmations de n'importe quel fournisseur.

La version honnête

Nous voulons être clairs sur ce que c'est et ce que ce n'est pas, parce que nous préférons que vous commenciez avec des attentes réalistes plutôt que de vous sentir déçu plus tard.

Local AI Processing est optionnel et désactivé par défaut. Cloud AI reste plus rapide, produit toujours de meilleurs résultats et fonctionne sur toutes les plateformes prises en charge par Hedy. Si vous n'avez pas de raison précise de vouloir un traitement sur l'appareil, l'option cloud offre actuellement une meilleure expérience.

Un résumé qui semble instantané dans le cloud peut prendre de 30 secondes à plusieurs minutes en local, selon votre matériel et le modèle que vous choisissez. Les modèles plus petits sont bons pour les résumés courts, mais se laissent piéger par les conversations longues ou nuancées. Les modèles plus grands se rapprochent de la qualité cloud, mais ont besoin d'un vrai matériel pour bien fonctionner. Et nous ne basculons pas silencieusement vers le cloud quand quelque chose échoue en local. Vous avez choisi le local pour une raison, et une nouvelle tentative discrète auprès de nos serveurs trahirait cette

intention. Vous verrez une erreur à la place.

L'IA locale est prise en charge sur les Mac Apple Silicon, les PC Windows équipés d'un GPU compatible, les iPhone 15 Pro et ultérieurs et les iPad M-series ou A17 Pro. Android et le Web sont sur la feuille de route, mais ne sont pas encore prêts. La diversité du matériel Android et les contraintes liées à l'exécution de modèles dans un navigateur rendent difficile la livraison d'une expérience cohérente aujourd'hui.

Vous choisissez le modèle qui correspond à votre matériel. Nous proposons trois niveaux de qualité, allant de modèles compacts qui tiennent sur un téléphone, à des options intermédiaires qui fonctionnent bien sur la plupart des ordinateurs portables modernes, jusqu'à des modèles plus grands qui approchent la qualité cloud sur du matériel capable. Le sélecteur de modèle vous montre ce qui convient avant le téléchargement.

Où cela nous mène

L'IA locale dans Hedy 3.2 est un point de départ, pas un produit fini. Les modèles continueront de s'améliorer. Le matériel grand public continuera de gagner en capacité. L'écart entre la qualité locale et la qualité cloud continuera de se réduire. Nous étendrons la prise en charge à davantage de plateformes à mesure que la technologie le permettra.

Cela fait partie d'un mouvement plus large qui mérite d'être reconnu. Pendant des années, les plus grands progrès de l'IA sont venus d'entreprises exécutant des modèles toujours plus grands sur des fermes de serveurs toujours plus grandes. Un changement plus discret se produit en parallèle : les modèles à poids ouverts deviennent plus intelligents et plus petits tous les quelques mois. Le matériel capable de les exécuter se trouve dans la poche et sur le bureau de la plupart des gens. L'écart de capacité entre l'IA de réunion cloud et l'IA de réunion sur l'appareil se réduit rapidement.

La vision d'ensemble compte davantage qu'une seule version. Nous pensons que les prochaines années de l'IA seront définies par un passage d'un monde où une poignée d'entreprises exploitent l'IA pour vous, à un monde où vous pouvez exécuter votre propre chaîne de traitement, sur votre propre appareil, avec vos propres données, de bout en bout. Que vous choisissiez de le faire ou non, avoir cette option est ce qui donne à la technologie sa forme appropriée. Cela maintient l'équilibre des pouvoirs.

Hedy est construit pour être à l'avant-garde de ce changement. L'IA locale est la première étape concrète. Il y en aura d'autres.

Pour l'activer, mettez à jour vers Hedy 3.2 et allez dans les paramètres, sous Speech & AI. Le bouton est libellé Local AI Processing. Choisissez un modèle adapté à votre matériel, et c'est prêt.

Pour les détails d'ingénierie — les modèles que nous avons choisis, la façon dont ils tiennent sur Mac, Windows et iPhone, et ce que l'inférence locale coûte en latence — consultez notre plongée technique sur l'IA locale (/fr/post/local-ai-engineering-deep-dive-hedy-3-2/) .

Si vous l'essayez, nous serions ravis de savoir ce que vous en pensez.

Hedy AI · Coaching IA en direct pour les conversations importantes

Essayez Hedy gratuitement: https://www.hedy.ai/fr/downloads/

https://www.hedy.ai/fr/post/local-ai-meetings-hedy-3-2/