論文とETIC 17

# あなたを知るAIに、仕事を任せる

個人AIの設計と、見守る手間の測り方




Daiki Morimoto / Unknown Lab
ORCID: 0009-0009-6444-750X
Version 1.0 / 公開日 2026-11-07
DOI: 10.5281/zenodo.24035478
未査読 / 専門職としての著者確認済み
独立専門レビュー未実施 / AI支援執筆
ライセンス: CC BY 4.0

## 30秒で読む

Wuら[1]が提案するのは、利用者の好み、価値、約束を覚え続け、そのつど説明し直さなくても仕事を任せられる個人AIの設計である。持続する利用者モデルを仕事ごとの分身が引き継ぎ、任せる範囲は領域ごとに決め、本番の前には利用者の文脈で練習する。進歩は、利用者が納得する結果を、より少ない確認と手直しで得られるかどうかで測るべきだとされる。未査読の立場論文による設計提案であり、実装、被験者、効果の測定は含まれない。

## 出張の手配と、数週間前の約束

ある人がAIに出張の手配を頼む。その人は数週間前、同じ日の夕方にある保護者面談へ出ると約束していたが、依頼文にその話はない。依頼ごとに答えるだけのアシスタントなら、予定表を空けて出張を入れてしまうかもしれない。Wuらはこの型の例を挙げている[1, 3.2節]。持続する個人エージェントなら、まず面談の後に出られる便を探し、どちらの約束も動かせなければ衝突を利用者へ戻す。目の前の依頼に現れない約束を守れることに、著者らは継続性の価値を見ている。

問題はAIの実行能力より手前にある。調べものや手配を任せても、利用者は曖昧な目標を詰め、解釈を直し、結果が十分かを判断し続けなければならない[1, 1節]。依頼のたびに前提を説明し、結果を点検し直すなら、任せた分の手間は大きくは減らない。著者らは、目標を立てて改め、どの判断を任せるかを選び、状況が変われば方向を変えるという人の主体性を保ったまま、意図に沿った結果を得る手間を減らすことを、個人エージェントの進歩の基準に据える。

## 論文の性格

主題の論文[1]は、2026年11月2日にarXivへ投稿されたv1（CC BY 4.0）で、査読を経ていない立場論文である。中身は設計図、委任の段階案、評価指標の定義、付録の任意の形式化から成る。人を対象にした標本、介入期間、試験、ベンチマーク、測定された作業削減はなく、著者ら自身の実装も示されていない。既存製品への言及[1, 3.1節]は位置づけのための例示であり、本稿はそれらの機能を確認していない。以下で扱う数値は、式の性質と本稿の仮の計算に限られる。資料ごとの根拠の種類は補足Aの表1にまとめた。

## 四つの部品が一つの委任をつくる

論文の提案は、四つの部品の組合せとして読むと分かりやすい。

第一の部品は、持続する利用者モデルである[1, 2.1節]。利用者の価値、判断の傾向、制約、固有の知識を、日々の振る舞いから更新しながら持ち続ける。著者らはこれを利用者中心のオペレーティングシステムの核にたとえ、単なる記憶の置き場を超えて、その人の判断の仕方を表すモデルとして描く。保護者面談の約束がこの核に残っていれば、後の出張依頼でも参照できる。

第二の部品は、仕事ごとの分身（clone）である。仕事用、アシスタント用、社交用、利用者が定義する用途の分身が同じ利用者モデルを受け継ぎ、それぞれの道具、アプリ、他者のエージェントにつながる。分身という語は人の複製を指さず、一つの利用者表現を場面ごとに投影したものを指す。そのため、出張を手配する分身も面談の約束を知った状態で動ける。著者らは、この層から、作業状態を時間やアプリをまたいで持ち越す仕事の継続、推定した意図に基づく先回りの支援、新しい道具を取り込む能力の拡張が生じるとする。

第三の部品は、任せる範囲の段階づけである[1, 2.2節]。著者らが描くL0からL3までの4段階では、L0とL1で人が実行や要所の判断に深く関わり、L2でアバターが定型的な判断を引き受けて人は少数のチェックポイントで介

入する。L3では委任された領域に限ってアバターが運用と判断を最後まで担い、人は事後の監査や方針の更新を行う。段階は領域ごとに決まり、どの段階でも利用者は目標を改め、委任を取り消す権限を持つ。出張の例でいえば、遅い便を探すところまでは任せ、二つの約束がぶつかったときの判断は利用者に戻す線引きがこれにあたる。L0〜L3は状態を記述する区分であり、L3に達することが常に望ましいとは書かれていない（詳細は補足Bの表3）。

第四の部品は、リハーサル工学である[1, 4.2節]。プロンプト、コンテキスト、ハーネスといった既存の工夫は、推論のその場で、エージェントが何を見て何をしてよいかを決める。リハーサルは回をまたいで利用者固有のフィードバックを取り込むもので、手順は四つある。

1. 利用者が許可した操作の履歴と好みを集める。
2. シミュレーション、反実仮想の展開、管理されたサンドボックスで、代表的な場面を再現する。
3. 明示された監督方針のもとで、候補の行動を評価する。
4. 限られた自律の範囲で、記憶、技能の選び方、判断方針を更新する。

本稿による当てはめの例を置く（原論文の実験ではない）。許可された過去の予定表と出張依頼をサンドボックスで再生し、面談や通院のような先約がある日に、候補の手配が「先約と衝突する場合は利用者に戻す」という方針を守れたかを評価する。守れなかった場面からは、「平日夕方の学校行事は動かしにくい」という記憶を限られた範囲で更新する。本番の依頼に出会う前に、その人の予定表の上で失敗しておけることが、この発想の要点である。ただし、論文はリハーサルによる改善を実測しておらず、方針のずれ、プライバシー、過学習をリスクとして挙げている。

四つを合わせると、核が覚え、分身が引き継ぎ、範囲が任せ方を決め、リハーサルが本番の前に核と方針を直す循環になる。冒頭の場面で求められていた、説明し直さずに任せて確認すべきところだけを確認する状態を、論文はこの循環が機能したときに得られるものとして描いている。

## 既存の個人化・記憶との違い

Wuらの整理では、既存の個人化手法の多くは、学習時に嗜好を学ぶか生成時に調整することで、その場の出力を利用者に寄せてきた。記憶システムは主に文脈を延ばし、経験をためてきた[1, 2.1節]。著者らは、記憶に頼る方式はエージェントの仕事では脆くなりうると指摘する[1, 4.1節]。記録の網羅は不完全で、好みは文脈に左右され、利用者の行動そのものにも一貫しない部分があるからである。明示的なプロフィールや権限設定を使う製品もあるが、それらを分身、環境、時間をまたいで一貫したモデルへまとめることは課題として残るという。

この論文が目指すのは、行動を生む判断の構造を学び、仕事をまたいで持ち運べる安定した判断の土台にすることである。本稿は、記憶、個人化、権限管理、評価、安全を利用者のための判断の委任という一つの研究課題に束ね、その進歩の測り方まで提案した点を論文の貢献と読む。既存手法より優れることを比較実験で示した論文ではない。

## 進歩を、成果と見守る手間で測る

設計に合わせて、著者らは評価の物差しとしてAgentic Oversight Index（AOI）も提案している[1, 2.3節]。まず、エージェントが届けた価値から計算や道具の費用を引き、この正味の価値をtanhで正規化した値Uにする。Uは−1と1の間に収まり、両端の値はとらない。そこから、人が見守る手間Pに重みλを掛けた値を差し引いたものがAOIである。見守る手間には、介入の回数、監督の時間、求められる判断の複雑さ、説明責任や監査の作業が含まれ、人だけで同じ仕事をした場合の負担を1とした比で表す。1を超えれば、人だけで行うより手間が増えたことになる。このためAOIは1未満にとどまるが有限の下限はもたず、手間が大きければ−1を下回る。

比率でなく差の形をとるのは、価値を手間で割ると、手間が0に近づくときに値が発散するからである。正味の価値を固定すれば、手間が増えるほど評価は必ず下がり、正味の価値が負のエージェントにも手間が減点として効く。届けた価値には満足や適切な信頼といった体験の価値を任意で含

めうるが、人の見守りは手間の項だけで数え、エージェント側の費用とは分けてある。

ただし、順位は重みで決まる。本稿の仮の計算（補足D）で、成果が高いが確認も多い案と、成果はやや低いが確認が少ない案を比べると、見守り1単位の重みを0.5から0.2へ変えるだけで順位が入れ替わる。普遍的な重みは検証されておらず、指標の信頼性と妥当性も確かめられていない。どの手間を負担として数えるかも定義次第である。式と記号は補足Dの表4に置いた。

## 示されたこと、試されていないこと

論文が示したのは、設計図、委任の区分、AOIの数学的性質である。前節で見たAOIの性質は定義から導かれるもので、利用者で確かめた結果ではない。定量的な結論として言えるのはこの範囲に限られる。

安全面では二つの層が提案されている[1, 4.3節]。技術的な安全の層は、意図とのずれの検知、高リスク時の確認、最小権限、サンドボックス、取り消せない操作の実行前の検証、チェックポイントと可能な操作の巻き戻し、プロンプトインジェクションの検知、記憶や認証情報の保護から成る。もう一つの層であるエージェント保険は、監査可能な実行記録と紛争手続きに支えられた補償の構想である。著者ら自身、補償は正しい実行を保証せず、金銭で取り戻せない害を防ぐには技術的な制御が引き続き必要だとしている。監査できる記録と、どの段階でも委任を取り消せることが、提案全体の前提にある。本稿は法的・保険上の評価は行わない。

未解決の点として、本稿は二つを挙げる。第一に、付録の形式化は利用者の同一性の一貫性と安定性に許容値を置くが、それをどう測り、値をどう決めるかは示していない（補足C）。第二に、これは本稿の解釈だが、行動には好みだけでなく制約も映る。いつも遅い便を選ぶ人は、朝の時間を好んでいるのかもしれないし、朝に子どもを送る用事があるだけかもしれない。転職や病気で好みが正当に変わったとき、安定を求める仕組みがその変化をずれとして押し戻すおそれもある。論文はフィードバックによる

修正を想定しているが、制約と好みを見分ける手がかりや、利用者が自分のモデルを直す経路はまだ具体化されていない。

Unknown Labは、未検証の提案として次の比較を挙げる。同じ道具と課題のもとで、明示的なプロフィールと記憶だけを使う単純な方式と、より豊かな利用者モデルを持つ方式を比べ、成果と各種の手間を分けて記録する。そのうえで、学習に使っていない場面での成績と、好みが変わったときに修正が反映される速さを見る。単純な方式で同等の成果と手間が得られるなら、追加のモデルを使う理由は弱まる。ただし、差が出なかった結果や少数の事例は同等であることの証明にならないため、許容する差を事前に決め、十分な数の課題で比べる必要がある。

## ケアの近くで手間を数えると

AOIの「人の手間」をケアの文脈に置くと、何を負担と数えるかが難しくなる。これを具体化するため、MINT研究[2]を補助的に読む。MINTはTelegram上で動く規則ベースのストレス対処チャットボットで、大規模言語モデルでもアバターの実装でもない。開発者と練習の設計者を含む、著者である研究者5人が、2025年6月の28日間、4週にわたって決められた順序でMINTを自ら使い、日記やワークショップをもとに6つのテーマを導いた。第4週に予定通知をすべて外すと、5人中4人が2日以内にほぼ使わなくなり、残る1人は呼吸法と歩行を自分で続けたと語った。

順序が固定され、時期や疲れの影響が重なり、自己報告に基づくため、この結果から通知撤去の因果効果や臨床効果は言えない。それでも評価の論点ははっきりする。練習や記録に使った時間をすべて見守りの手間に数えると、身につけるための時間まで減点になる。第4週の利用の減少も、自分で対処できるようになった結果か、離れていった結果かを、利用回数だけでは見分けられない。

Wuらの式は体験の価値を含めうるが、ケアの場面でその価値をどう測り、誰の目標に基づいて線を引くかは残る。二つの研究は対象も方法も異なり、互いを検証するものではない。方法と6テーマの詳細は補足Eに置いた。

## ETIC 2.0から考える

ETIC 2.0[3]は、作業、人、環境、関係のつながりの中に技術を位置づけ、ケアを設計するための公開された枠組みである。Wuらは、一回ごとの依頼よりも、仕事と時間をまたいで続く利用者の目標と約束を設計の単位に置き、成果と同じ物差しの中で人の手間を数えようとした。保護者面談の約束を守る出張手配は、技術がその人の生活の作業と関係の中に入って初めて価値を持つ例として読める。Unknown Labは、この点がETIC 2.0の見方と重なると考える。ただし、論文はETIC 2.0を検証しておらず、この重なりは枠組みの妥当性を示すものではない。

Unknown Labは、ケアの場面では手間を減らすことを進歩とみなす物差しに慎重さが要ると考える。MINT研究の経験に照らすと、利用者が支援に費やす時間には、確認や手直しのように避けたい負担と、練習や関係づくりのように本人が望む参加が混ざっている。どちらに当たるかは利用者の目標によって変わり、同じ10分の記録が、ある人には負担であり、別の人には自分を振り返る大切な時間でありうる。合計点だけを見ると、見守りが減った一方で本人が重視する成果が悪化した場合でも、評価が上がって見えることがある。

そこでUnknown Labは、AOIのような複合スコアを導入する前の段階として、既存のプロフィールや記憶の仕組みを使い、成果、見守りの負担、本人が価値を置く参加を分けて記録しておくことを提案する。重みは課題ごとに利用者と決め、その根拠とともに残す。原論文はすでに体験の価値を任意で含めうるため、この手順は新しい変数を加えず、用意された項の中で測り方を明確にするものにとどまる。この方法が有効であることは、原論文も本稿も示していない。

## 結論：この個人AI構想が差し出すもの

Wuらの論文が差し出すのは、その人の好みと約束を覚え続け、説明し直さずに仕事を任せられる個人AIの設計図である。持続する利用者モデル、それを引き継ぐ分身、領域ごとに決めて取り消せる委任の範囲、本番前のリハーサルを組み合わせ、進歩は、利用者が納得する結果をより少ない確認と手直しで得られるかどうかで測られる。既存の個人化や記憶を委任という一つの課題に束ね、手間を評価の中心に置いた点に、この提案の独自性がある。

まだ試されていないのは、その全体である。実装も被験者もなく、利用者モデルが約束や好みを正しく保てるか、リハーサルが本番での失敗を減らすか、AOIの重みと指標が信頼できるかは、いずれも測定されていない。ケアに近い場面では、減らすべき手間と本人が望む参加の線引きも残る。次に確かめるべきは利用者が何に気づき、何を確かめなければならないかであり、その手がかりになる論文を補足Fに挙げた。

# 補足

## 補足A 資料と根拠の種類

### 表1 本稿の資料と根拠の種類

| 資料 | 根拠の種類 | 本稿での使い方 |
|---|---|---|
| Wuら[1] | 未査読の立場論文。設計図、段階案、評価式とその数学的性質 | 主題。仕組みと評価の考え方を読む |
| Bolpagniら[2] | 査読誌掲載の協働的オートエスノグラフィー。著者5人、28日間、定性分析が中心 | 評価上の論点を具体化する補助例 |
| Morimoto[3] | Unknown Labの公開枠組み | 論評の参照点。実証根拠には使わない |
| Mittalら[4] | プレプリント。書誌と要旨のみ確認 | 次に読む候補 |

出典：各文献の書誌情報と本文。分類と用途は本稿の判断。

## 補足B アバターの構成要素と委任の段階

### 表2 アバターの構成要素と役割

| 構成要素 | 論文での役割 | 読む際の注意 |
| --- | --- | --- |
| 利用者モデルと記憶 | 価値、判断の傾向、制約、利用者固有の知識を持ち続ける核 | 表現方法と精度は未提示 |
| 知覚・計画・決定・実行の循環 | 環境の信号を、利用者に沿った行動へ変える | 各部の方式は未指定 |
| 仕事・アシスタント・社交・カスタムの分身 | 同じ利用者モデルを継承し、文脈ごとの道具や他者のエージェントへつながる | 人の複製を意味しない |
| 仕事の継続、先回りの支援、能力の拡張 | 分身の層から生じるシステム能力 | 先回りは推定した意図に依存する。誤推定時の扱いが課題 |

出典：[1] 2.1節と図2を本稿が要約。注意欄は本稿の解釈。

論文は、分身が他の利用者のエージェントと会議の時間を交渉する場面も描き、交渉に必要な範囲を超えて好みを明かさない方法を未解決の問いとして残している[1, 3.2節]。

表3 委任の成熟度L0〜L3（著者らの提案）

| 段階 | エージェントが担うこと | 人が担うこと |
|---|---|---|
| L0 Human-for-the-loop | 断片的な下位作業の自動化 | 実行と判断の両方に関わる |
| L1 Human-in-the-loop | 長めの作業手順の実行 | 目標の具体化、振り分け、確認といった要所の判断と、責任の大半 |
| L2 Human-on-the-loop | 持続する同一性・記憶・判断方針を備えた核が、定型的な判断と段取りを引き受ける | 調整された少数のチェックポイントでの介入。実行時の防護と監査記録が前提 |
| L3 Human-out-of-the-loop | 委任された領域に限り、運用と判断を最後まで担う | 事後の監査、紛争の解決、方針の更新 |

出典：[1] 2.2節と図3を本稿が要約。どの段階でも、利用者は目標を修正し委任を取り消す権限を保持する。検証済みの標準ではない。

研究の焦点は、段階とともに移るとされる。L0とL1では実行の質、L2ではリスクを抑えた判断の委任、L3では高い責任を伴う自律の制度的な保証が中心になる。ある領域がどの段階にあるかを判定する数値基準は、示されていない。

論文の表1は、プロンプト、コンテキスト、ハーネス、リハーサルの4種の工学を、信号、適用時点、仕組み、目的、リスクの観点で並べた概念的な比較であり、性能の比較結果は含まない。本文のリハーサルの当てはめを実際に評価するなら、学習に使っていない期間の依頼で試すことと、模擬した相手の反応と現実の相手の反応のずれを確かめることが必要になる。

## 補足C 利用者の同一性の形式化

付録Aは、個人化を任意で形式化している[1, 付録A]。利用者uの同一性の表現はIu=(zu, Ku)と書かれる。zuは価値観、判断の仕方、リスクへの態度といった比較的安定した構造、Kuは利用者固有の知識と経験であり、どう考えるかと何を知っているかを分けて持つ設計である。

zuは直接観察できないため、場面xtと行動otの記録から推定する（zu=f({(xt, ot)})）。論文はこれを、認知の逆モデリングと呼ぶ。長く一緒に働く同僚が、どの会議を断り、どれを受けるかを見て、「午前を守る人だ」と察する過程に近い。

学習の目標は、方針πが利用者の実際の行動に高い確率を与えること、すなわちE[log π(o | x, lu)]の最大化である。このとき、同一性の不一致度Cを許容値ε以下に抑える制約を課す。これにより、場面ごとの場当たり的な最適化を防ぐ。さらに、利用者とアバターの双方が時間とともに変わるため、両者の表現の隔たりDをδ以下に保つ安定条件も置かれる。

C、D、ε、δの操作的な定義と、値の決め方は示されていない。なお、ここで隔たりを表すDと、補足DのAOIで判断の複雑さを表すDは、論文中で同じ文字が別の意味に使われている。

## 補足D AOIの式と仮の計算

AOIは、課題の分布τにおけるアバターAについて、次のように定義される[1, 2.3節, 付録B]。

- U(A, τ)=tanh((V−C_sys)/s_τ)
- AOI(A, τ)=U−λP（λ>0）

人の負担はH=w_f·F+w_t·T+w_d·D+w_a·Qで、重みはいずれも非負である。これを、人だけで行う場合の負担H_human（>0）で割り、P=H/H_humanとする。

表4 AOIの構成要素

| 記号 | 意味 | 読む際の注意 |
|---|---|---|
| V | 利用者に届いた価値（課題の成功、質、感じられた有用性）。付録では、満足や適切な信頼という任意の体験価値も含めうる | 処理量だけを数える設計ではない |
| C_sys | 計算、道具の使用、学習にかかるエージェント側の費用 | 人の見守りを含めず、二重計上を避ける |
| $s_T$ | 課題ごとの正の基準尺度。Vと同じ価値の単位 | 決め方は課題に依存する |
| U | tanhで−1と1の間に収めた正味の価値 | 正味の価値の順序は保たれる |
| H | 介入頻度F、監督時間T、判断の複雑さD、説明責任· 監査の作業Qを、非負の重みで足し合わせた人の負担 | 項目同士が重なりうる |
| P | 人だけで行う場合の負担 $H_{human}$（>0）に対するHの比 | 0は見守り不要、1は人だけと同等、1超は人だけより重い |
| λ | 見守り1単位に割り当てる減点 | 普遍的な値は未検証 |

出典：[1] 2.3節と付録B。注意欄は本稿の解釈。

定義から、Uを固定するとPが1単位増えるごとにAOIはλだけ下がり、正味の価値が負の場合でも同じであることが導かれる。Pが0ならAOIはUに一致する。Uは−1と1の間（両端を含まない）にあるため、AOIは1未満に収まるが有限の下限はなく、Pが大きければ−1を下回る。付録Bの価値や費用の曲線は例示であり、データに当てはめたものではない。

以下は、重みの働きを見るための本稿の仮の算術である。研究データでも、効果量でも、製品の比較でもない。

- 案A：成果が高いが、確認も多い（U=0.80, P=0.60）
- 案B：成果はやや低いが、確認が少ない（U=0.60, P=0.10）

λ=0.50なら、A=0.50、B=0.55で、Bが上になる。λ=0.20なら、A=0.68、B=0.58で、Aが上になる。両者が並ぶのはλ=0.40のとき（A=B=0.56）である。順位は見守り1単位をどれだけ重く見るかで決まるため、λやwには課題と利用者に即した根拠が要る。

項目の重なりにも注意が要る。一回の確認操作が介入回数Fを増やし、監督時間Tに加わり、監査作業Qにも記録されることがある。回数、時間、説明責任は正当に別々の側面でありうるが、同じ一つの行為が三つの項目に寄与するので、各項目が何を測るかを先に定義すれば意図しない重複を避けやすい。AOIの横にU、P、FからQまでの内訳を並べれば、合計点だけで順位を読み違える危険も減る。AOIの信頼性と妥当性は確かめられておらず、臨床上の効果や省力効果も推定されていない。

## 補足E MINT研究の方法と6テーマ

MINT研究[2]は、JMIR Human Factorsに2026年10月27日に掲載された協働的オートエスノグラフィーで、2025年12月のプレプリントと同じ研究である。協働的オートエスノグラフィーは、研究者が自らの経験を体系的かつ共同で記録し、解釈する方法である。

MINTが届ける練習は、CBTの技法に基づく5分未満の5種類（認知的再評価、呼吸の調整、前向きな気分の誘導、笑顔、最良の自分の想像）である。個人化に使われたのはVARK質問票で本人が答えた提示形式の好みだけで、それに合わせて図、音声、文章、体を動かす指示のいずれかで届ける。この使い方は、好みに合わせた提示が学習に効くことを示すものではない。ボットは、Node.js、grammY、MongoDB、定時実行のcronで作られた規則ベースの有限状態システムである。

表5 MINT研究の設計と主な観察

| 項目 | 内容 | 解釈上の限界 |
| --- | --- | --- |
| 参加者 | 著者5人全員。デジタルメンタルヘルスの研究者で、開発責任者（RP2）と練習の設計者（RP4）を含む | 外部の利用者や患者はいない。同質で、設計を知る集団 |
| 期間と順序 | 2025年6月の28日間。第1週は好みの形式、第2週は好みでない形式、第3週は練習を自分で選ぶ案内のみ、第4週は予定通知をすべて廃止 | 順序は固定で、無作為化された個人内比較ではない |
| データと分析 | 日記、共有の振り返り記録、生態学的瞬間評価（EMA）、3人の装着型機器、約90分の対面ワークショップ2回（録音・文字起こし）。6段階のリフレクシブ・テーマ分析（研究者が自らの立場と解釈を省察しながら主題をつくる方法） | 付録のコードの出現回数は分析の痕跡であり、一般の利用者における頻度や評定者間信頼性を示すものではない |
| 主な観察 | 通知の全廃後、5人中4人が2日以内にほぼ使用をやめた。RP4は呼吸と歩行を自分で続けたと語った | 順序、時期、疲れが重なっており、通知撤去の因果効果は特定できない。RP4の例は自己報告 |
| ストレス尺度 | 開始時と終了時にPSS-10を記録 | 対照がなく、ストレス低減の効果は推定できない |

出典：[2] 方法・結果・限界の節、表1・表2、付録2を本稿が要約。

表6 MINT研究の6テーマと設計上の示唆

| テーマ | 観察の例 | 著者らの設計上の示唆 |
|---|---|---|
| 時機と受け入れやすさ | 自分で確保した休息の時間に通知が割り込んだ。落ち込んだ日には「最良の自分」の想像がかえってつらく、呼吸法は気分を問わず使いやすかった | 生理的な興奮だけに頼らず、支援を受け取る余力がある時機を推定する |
| 応答と継続性 | 多くのデータを入力しても役立つ形で返ってこず、宿題のように感じた。週ごとの振り返りを望む声があった | 過去の入力を覚えて応える継続性を、中核機能として扱う |
| 既存の対処との統合 | 走る、犬の散歩、車内で歌う、読書のほうが速く効いた | 練習を、既存の対処の幅を広げるものとして設計する |
| 関わらないことの意味 | ストレスが最も高い時期ほど記録が途絶えた | 沈黙を、適応と評価に使える意味のある信号として扱う |
| 通知への依存 | 通知の全廃後、自発的な使用が急に減った | 促されての利用から自発的な対処へ移る道筋を設け、通知を徐々に減らす |
| 日常の技術が基準になる | 装着型機器の値は実感と合わないことが多かった。何人かは汎用の大規模言語モデルに気持ちを書き、覚えていて応える点を評価した | 装着型の計測は文脈の補助に使い、継続性への期待に応える構成を探る |

出典：[2] 結果の節、表3、付録2。各テーマの頻度は推定していない。

著者らは、これらのテーマ自体を新しい発見とは位置づけておらず、同じ現象を内側から細かく見た点に価値を置いている。

解釈には注意が要る。沈黙の原因の分析は振り返りの自己報告に基づき、著者らは感情の過負荷と他の用事や疲れとを見分けられないことを認めている。付録2の通知撤去に関する小テーマには5人全員が声を寄せているが、これは語った人数であり、使用をやめた人数は本文に記された4人である。

汎用LLMの自発的な利用は、検証された比較でも治療の代わりに使うことの推奨でもなく、著者らは反すうを長引かせうる面にも触れている。

研究の限界として、参加者は研究者だけの同質な少人数で開発者が参加者を兼ね、期間は4週間、試作は一つだった。第4週の通知の一斉撤去は現実的な運用を模したものではなく、著者らも利用の減少を通知への依存だけに帰すことはできないとしている。生活上の成果を長期にわたって独立に評価する指標はなく、運営の人手、保守、費用対効果も数量化されていない。

資金提供と利益相反は、いずれもないと申告されている。生の定性データは公開されておらず、匿名化した一部がデータ共有の合意のうえで請求に応じて提供される。付録1は練習とEMAのJSONや日程表CSVの所在としてOSFを示しているが、本稿は入手の可否と利用条件を確認しておらず、仕組みを再現できたとはいえない。

## 補足F 次に読む論文

Mittalら[4]は、利用者が期待した進め方からエージェントが外れる「委任の失敗」を利用者の視点から分類した体系MANDATEを提案し、失敗の深刻さと気づきやすさについての利用者調査も報告している。AOIの介入頻度、監督時間、監査作業を定義するには、利用者が何に気づき、何を確かめなければならないかを把握する必要があり、この分類はその候補を具体化する次の一歩になりうる。

本稿が確認したのは書誌と要旨だけで、全文の方法と結果は監査していない。未査読のv1であり、分類の妥当性や採用状況も確認されていない。

# 参考文献

[1] Wu Z, Ma X, Yao Y, Huang Z, Hu H, Lu Y, Wang T, Yang Y, Xiong W, Cheng P, Zhang Z. Avatar Intelligence: Amplifying Human Agency Through Persistent Personal Agents. arXiv:2611.03092v1 [cs.AI]; 2026年11月2日投稿（v1を2026年11月7日に確認）. 未査読. CC BY 4.0. https://arxiv.org/abs/2611.03092v1（全文HTML: https://arxiv.org/html/2611.03092v1 ）

[2] Bolpagni M, Piffer SA, Pozza M, Maminimini M, Gabrielli S. Insights on Designing the MINT (Micro-Interventions for Navigating Tension) Chatbot System for Stress Management: Collaborative Autoethnography. JMIR Human Factors. 2026;13:e89876. 2026年10月27日公開. doi:10.2196/89876. PMCID: PMC13717341. CC BY 4.0. https://doi.org/10.2196/89876（出版社ページ: https://humanfactors.jmir.org/2026/1/e89876/ ）

[3] Morimoto D. エコテクノロジカル・インテグラティブ・ケア(ETIC)2.0, Version 2.1改訂版 - AI時代の作業療法とケア生態系設計. Unknown Lab; 2026. doi:10.5281/zenodo.21848456. https://doi.org/10.5281/zenodo.21848456

[4] Mittal S, Kim SSY, Counts S, Neville J, Das D, Laban P. Babysitting AI Agents: A User-Centered Taxonomy and Study of Agentic Delegation Failures. arXiv:2610.34748v1 [cs.HC]; 2026年10月26日投稿. 未査読. https://arxiv.org/abs/2610.34748v1