

AUTOPSIA EMOȚIILOR VIITOARE

Între darwinismul psihologic, istorie,
antropologie și lumea inteligenței artificiale



Copyright © [2026] Axiologic Research

All rights reserved.

KDP publishing rights held by Outfinity SRL (Iași, Romania).

Available for free on Axiologic website (www.axiologic.net).

Author's Note

This work was created under the umbrella of Axiologic Research as part of two interconnected research programs, one dedicated to Artificial Intelligence and the other to Outfinitism, both led by Sînică Alboaie. To explore the details of how these two programs intersect and build upon each other, we invite readers to visit the Outfinity Venture Studio page at www.outfinity.ch.

While Artificial Intelligence language models were used to assist with text generation, the foundational philosophy, thematic structure, and analytical approach derive exclusively from the author's decades of dedicated study of social phenomena, social technologies, and genuine hard technologies resulting from various European and self funded research projects. Recently, within the Achilles research project, we have been deeply studying AI generated literature and its automated verification using neuro symbolic approaches; all the free books made available to the public are part of the suite of works we use to test our latest technologies.

Cuprins

Notă asupra metodei și a gradului de certitudine

Introducere. De ce o teorie a emoțiilor este necesară pentru a înțelege lumea automatizată

Capitolul 1. Ce este o emoție și ce poate explica evoluția

Capitolul 2. Emoțiile ca infrastructură a interacțiunii sociale

Capitolul 3. Statut, cooperare și moralitate afectivă

Capitolul 4. Cultura nu doar numește emoțiile

Capitolul 5. Munca drept instituție afectivă și incertitudinea automatizării

Capitolul 6. Trei viteze ale evoluției afective

Capitolul 7. Pot apărea emoții noi?

Capitolul 8. Atlas provizoriu al afectelor unei lumi automatizate

Capitolul 9. Relații cu agenți care par sociali

Capitolul 10. De la afect la strategie: adaptări comportamentale într-o lume mediată de AI

Capitolul 11. Tempo-ul afectiv: cum emoția simulată de AI poate schimba comunicarea umană

Capitolul 12. Spre o ecologie afectivă deliberată

Concluzie. Continuitatea biologică și discontinuitatea lumii

Anexă. Matricea sintetică a scenariilor

Referințe

Notă asupra metodei și a gradului de certitudine

Această carte propune o interpretare evoluționistă și socială a emoțiilor, apoi o extinde către scenarii în care inteligența artificială preia o parte importantă a muncii cognitive și, eventual, fizice. Prima parte se sprijină pe rezultate din știința afectivă, psihologia evoluționistă, antropologie, istoria emoțiilor, psihologia culturală și cercetarea asupra muncii. A doua parte folosește aceste rezultate pentru a construi ipoteze despre schimbări afective posibile. Cele două părți nu au același statut epistemic.

În știința emoțiilor nu există consens asupra unei singure definiții, asupra numărului de emoții sau asupra modului în care biologia, evaluarea cognitivă, corpul, limbajul și cultura se combină pentru a produce un episod emoțional. Teoriile emoțiilor de bază, teoriile evaluării, modelele construcționiste și abordările neurobiologice descriu uneori aceleași fenomene în vocabular diferit și, alteori, fac predicții incompatibile. Cartea nu încearcă să ascundă această situație. Atunci când o explicație funcțională este bine susținută, ea este prezentată ca interpretare plauzibilă, nu ca istorie evolutivă demonstrată. Atunci când literatura este controversată, disputa este păstrată. Atunci când trecem la viitor, scenariile sunt marcate ca extrapolări sau speculații.

Vom folosi patru niveluri de afirmații. Primul cuprinde observații empirice relativ robuste, de exemplu faptul că pierderea locului de muncă este asociată în medie cu deteriorarea sănătății mintale sau că expresiile emoționale modifică comportamentul interlocutorilor. Al doilea cuprinde interpretări funcționale plauzibile, precum ipoteza că rușinea ajută la gestionarea riscului de devalorizare socială. Al treilea cuprinde extrapolări condiționate, de exemplu că automatizarea extinsă ar putea slăbi legătura dintre competență, venit și statut. Al patrulea cuprinde speculații conceptuale, precum posibilitatea ca societățile viitoare să distingă și să numească forme afective astăzi încă nediferențiate.

În text, precizările despre incertitudine nu sunt formule de politețe. Ele sunt parte din argument. O explicație evoluționistă este ușor de făcut să pară inevitabilă după ce cunoaștem rezultatul; o predicție tehnologică este ușor de făcut să pară serioasă dacă omite instituțiile,

prețurile, politica și rezistența socială. Cartea încearcă să evite ambele erori.

Termenul „darwinism psihologic” este folosit într-un sens limitat. Înseamnă că tratăm emoțiile și sentimentele ca rezultate ale unei istorii de variație, selecție, dezvoltare și transmitere culturală și că întrebăm ce probleme de coordonare a organismului și a grupului ar fi putut rezolva. Nu înseamnă că fiecare trăsătură este o adaptare optimă, că orice diferență socială este biologic necesară sau că ceea ce a fost selectat este moral bun. Selecția naturală nu furnizează norme etice. În plus, trăsăturile pot fi produse secundare, exaptări, compromisuri, efecte ale dezvoltării sau răspunsuri nepotrivite într-un mediu nou.

Prin „emoție” vom înțelege, provizoriu, un episod relativ delimitat care implică evaluarea unei situații relevante, modificări ale atenției și motivației, tendințe de acțiune, modificări corporale și, uneori, expresie socială. Prin „sentiment” vom desemna fie experiența conștientă a unei stări afective, fie o orientare evaluativă mai durabilă față de o persoană, un grup, o activitate sau o idee. Această distincție este utilă pentru carte, dar nu reprezintă un consens terminologic universal. „Afectul” va desemna niveluri mai elementare de valență și activare, iar „dispoziția” o stare mai difuză și mai persistentă, fără un obiect clar.

Cartea nu susține că inteligența artificială va elimina munca. Datele actuale arată expunere diferențiată a sarcinilor, creșteri de productivitate în anumite contexte și efecte eterogene asupra calității muncii. Organizația Internațională a Muncii estimează că transformarea este, în prezent, mai probabilă decât înlocuirea completă pentru majoritatea ocupațiilor expuse, iar studiile OECD găsesc simultan beneficii, temeri și riscuri de pierdere a agenției. O lume cu mult mai puțină muncă umană este tratată aici ca scenariu posibil, nu ca prognoză sigură [ILO, 2025; OECD, 2023; OECD, 2025].

Extensia de față adaugă două unități de analiză care au fost tratate doar marginal în versiunea inițială: strategia comportamentală și temporalitatea comunicării. Prima urmărește nu doar ce simte individul, ci ce face în mod repetat într-un mediu cu AI: ce deleagă, ce verifică, ce competențe păstrează, cum își semnalizează contribuția și cum se regroupează social. A doua urmărește felul în care interlocutori artificiali disponibili permanent, capabili să simuleze căldură, grijă, ezitare sau entuziasm, pot modifica viteza și normele exprimării emoționale. Pentru aceste două teme literatura este mai tânără decât pentru psihologia

emoțiilor clasice, astfel încât marcarea incertitudinii devine și mai importantă.

În special, nu vom presupune că un efect observat într-o conversație scurtă cu un chatbot se transformă automat într-o schimbare de personalitate sau într-o normă socială stabilă. Vom separa efectele de interfață, adaptările comportamentale repetate, schimbările instituționale și eventualele transformări culturale de lungă durată. O mare parte din capitolele 10 și 11 este formulată sub forma unor mecanisme și predicții falsificabile tocmai pentru a evita convertirea unei tehnologii aflate în schimbare rapidă într-o antropologie imaginară.

Introducere. De ce o teorie a emoțiilor este necesară pentru a înțelege lumea automatizată

Discuția publică despre inteligența artificială se concentrează firesc asupra performanței, productivității, salariilor, proprietății și distribuției puterii. Aceste teme sunt indispensabile, dar nu epuizează schimbarea. Munca nu este doar o metodă de producere a bunurilor și de distribuire a venitului. Este și un sistem prin care societatea organizează timpul, acordă statut, confirmă competența, creează relații repetitive, stabilește obligații și oferă indivizilor un vocabular pentru a-și explica utilitatea. Dacă inteligența artificială modifică profund munca, ea modifică și mediul în care sunt declanșate, interpretate și reglate emoțiile.

Această afirmație nu presupune că emoțiile sunt fragile convenții care se schimbă odată cu fiecare tehnologie. Frica, atașamentul, furia, căutarea statutului și sensibilitatea la excludere au rădăcini mult mai vechi decât capitalismul industrial. Totuși, aceleași sisteme afective pot produce experiențe diferite când obiectele, normele și instituțiile se schimbă. Rușinea unei persoane dintr-o societate a onoarei, rușinea unui funcționar modern și rușinea unui student care folosește un model generativ nu sunt identice, chiar dacă toate pot implica anticiparea devalorizării sociale. Biologia oferă constrângeri și predispoziții; cultura organizează situațiile în care acestea devin relevante, numele pe care le primesc și acțiunile considerate adecvate.

Premisa centrală a cărții este că emoțiile au cel puțin patru dimensiuni funcționale. Ele coordonează organismul, selectând priorități sub presiunea timpului. Ele coordonează interacțiunile, semnalând intenții, nevoi, limite și evaluări. Ele susțin ordinea grupului, deoarece rușinea, vinovăția, indignarea, recunoștința și mândria contribuie la conformare, reparație, sancționare și distribuirea prestigiului. În cele din urmă, ele pot fi cultivate instituțional: ritualurile, educația, religia, piața, organizațiile și platformele digitale fac anumite stări mai frecvente, mai vizibile și mai recompensate decât altele [Keltner & Haidt, 1999; Van Kleef & Côté, 2022; Fiske, Schubert & Seibt, 2025].

Dacă acceptăm această premisă, automatizarea nu mai apare doar ca substituție de sarcini. Ea devine o modificare a ecologiei afective. Un sistem generativ poate reduce efortul necesar pentru o activitate,

dar poate reduce și sentimentul de autorat. Poate democratiza accesul la competență, dar poate slăbi o sursă de prestigiu. Poate oferi companie, dar poate modifica așteptările față de reciprocitate. Poate personaliza experiențe, dar poate diminua întâlnirea cu surpriza și opoziția. Niciunul dintre aceste efecte nu este inevitabil. Ele depind de proiectarea sistemelor, de distribuția proprietății, de normele profesionale și de instituțiile care vor media tranziția.

Întrebarea dacă pot apărea emoții noi cere, înainte de orice, o clarificare. O emoție poate fi „nouă” în mai multe sensuri. Poate fi un mecanism biologic nou, stabilizat prin selecție genetică. Poate fi o configurație nouă a unor componente mai vechi. Poate fi o categorie culturală care face recognoscibilă o experiență anterior difuză. Poate fi un termen nou pentru o stare veche. Poate fi, în sfârșit, o practică instituțională care produce și normalizează o anumită experiență. Aceste variante au probabilități și scări temporale foarte diferite.

Apariția unui mecanism afectiv biologic complet nou în câteva decenii este puțin plauzibilă. Selecția genetică umană continuă, dar schimbările psihologice complexe cer, de regulă, variație ereditară, diferențe de reproducere și timp. În schimb, apariția unor categorii afective culturale noi este nu doar posibilă, ci bine documentată. Cuvinte precum „solastalgie” sau „fear of missing out” nu demonstrează apariția unor circuite noi, dar arată cum schimbări ecologice și tehnologice pot produce configurații recurente care devin suficient de importante pentru a fi numite, măsurate și reglate social [Albrecht et al., 2007; Przybylski et al., 2013].

Între biologie și vocabular se află dezvoltarea. Creierul uman este neobișnuit de plastic, iar copilăria prelungită permite însușirea unor competențe și categorii pentru care selecția nu a construit module specifice. Cititul, algebra și programarea folosesc țesut neural care nu a evoluat pentru aceste activități în forma lor actuală. Este deci rezonabil să presupunem că oamenii pot învăța și noi moduri de a discrimina, interpreta și regla stările afective fără să aștepte o mutație genetică. Cât de mult poate cultura modifica însăși fenomenologia emoției, nu doar descrierea ei, rămâne o întrebare deschisă [Dehaene & Cohen, 2007; Hoemann et al., 2019; Lindquist, Satpute & Gendron, 2015].

Cartea va susține o teză moderată. Într-o lume tehnologizată, este probabil să apară noi obiecte emoționale, noi norme afective și noi

categorii culturale. Este posibil să apară configurații suficient de stabile încât oamenii să le trateze drept emoții distincte. Este mult mai puțin probabil ca, într-un orizont istoric scurt, să apară familii biologice complet noi de emoții. Cea mai importantă schimbare nu va fi neapărat apariția unei stări fără precedent, ci reorganizarea relațiilor dintre emoții vechi: rușinea se poate lega de delegare în loc de incompetență; mândria se poate muta de la execuție la selecție și responsabilitate; doliul se poate atașa de o identitate software; anxietatea poate privi nu lipsa resurselor, ci lipsa unui rol justificat.

Această teză are și o implicație normativă. Dacă platformele și modelele optimizează interacțiuni la scară mare, ele nu doar răspund emoțiilor existente. Ele creează condiții de selecție culturală. Un tip de mesaj care produce indignare și menține atenția poate fi repetat mai des; un stil relațional care produce atașament poate fi recompensat comercial; o formă de rușine poate deveni mecanism de control profesional. În acest sens, algoritmii pot accelera evoluția culturală a emoțiilor fără să „simtă” nimic. Ei pot modifica ce se răspândește, ce este vizibil și ce devine normă [Brady et al., 2021; Creanza, Kolodny & Feldman, 2017].

Totuși, termenul „selecție” trebuie folosit cu precizie. Nu orice schimbare este evoluție, iar analogia cu selecția naturală poate ascunde deciziile deliberate ale companiilor și statelor. O platformă nu este un mediu neutru în care ideile concurează spontan. Funcția ei de recomandare, obiectivele comerciale și regulile de moderare construiesc mediul de competiție. Când spunem că o expresie emoțională „câștigă”, trebuie să întrebăm prin ce metrică, în folosul cui și cu ce costuri externe.

Cartea este organizată în trei mișcări. Primele capitole reconstruiesc emoțiile ca mecanisme biologice și sociale, păstrând disputele teoretice. Partea mediană analizează cultura, munca și cele trei viteze ale schimbării afective: biologică, culturală și algoritmică. Ultimele capitole propun un atlas prudent al unor sentimente posibile într-o lume cu automatizare extinsă, evaluează relațiile cu agenți artificiali și discută ce instituții ar putea proteja autonomia, competența, reciprocitatea și sensul. Scopul nu este să inventăm denumiri memorabile, ci să formulăm ipoteze care pot fi criticate și, unde este posibil, testate.

O altă limită a unei teorii centrate doar pe trăire este că emoțiile nu produc direct o singură acțiune. Frica poate genera evitare, pregătire, căutarea unei alianțe sau investigarea pericolului; rușinea poate produce retragere, reparație sau agresiune defensivă. Într-o lume AI, diferențele importante pot apărea mai întâi la nivelul strategiilor: unii indivizi vor delega maximal, alții vor conserva competențe, alții vor construi reputație în jurul provenienței umane, iar instituțiile vor inventa reguli de audit și responsabilitate. De aceea cartea extinsă urmărește explicit trecerea de la afect la comportament planificat.

În plus, AI conversațional schimbă o proprietate mai puțin discutată a mediului social: timpul. Un interlocutor artificial poate fi disponibil permanent, poate răspunde aproape instantaneu și poate simula răbdarea sau ezitarea fără constrângerile corpului. Nu știm încă dacă asemenea interacțiuni vor recalibra așteptările oamenilor față de comunicarea umană. Putem însă studia mecanisme intermediare: auto-dezvăluirea, acomodarea lingvistică, efectele latenței, externalizarea reglării emoționale și transferul practicilor AI către conversațiile dintre oameni.

Capitolul 1. Ce este o emoție și ce poate explica evoluția

1.1 O categorie științifică fără definiție unică

Știința emoțiilor studiază un obiect pe care îl recunoaștem intuitiv, dar îl delimităm dificil. Frica pare diferită de o simplă judecată despre pericol, pentru că implică tensiune corporală, atenție selectivă, impuls de evitare și o experiență subiectivă. Totuși, niciuna dintre aceste componente nu este necesară în toate cazurile. O persoană poate raporta frică fără o reacție fiziologică puternică; poate manifesta apărare fără să numească experiența „frică”; poate interpreta aceeași activare corporală drept entuziasm, amenințare sau furie, în funcție de context.

Modelele clasice au încercat să găsească o ordine stabilă. Teoriile emoțiilor de bază susțin că un număr de familii afective au programe evolute, expresii și funcții recognoscibile. Teoriile evaluării pun accentul pe modul în care organismul apreciază relevanța, controlul, noutatea, compatibilitatea cu scopurile și responsabilitatea. Modelele dimensionale descriu experiența în spații precum valență și activare. Teoria emoției construite susține că creierul folosește predicții, semnale interoceptive și concepte învățate pentru a construi episoade pe care cultura le numește „furie”, „tristețe” sau „teamă”. Neuroștiința afectivă caută sisteme funcționale și circuite, dar cercetarea nu a identificat câte o semnătură neurală simplă și exclusivă pentru fiecare cuvânt emoțional [Barrett, 2017; Panksepp, 1998; Scherer, 2009; Adolphs & Anderson, 2018].

Disputa nu este doar terminologică. Dacă emoțiile sunt categorii biologice discrete, apariția uneia noi ar cere probabil o modificare biologică substanțială. Dacă sunt construcții predictive care folosesc afectul și conceptele, cultura poate produce mai ușor noi categorii. Dar opoziția este adesea exagerată. Chiar și teoriile construcționiste acceptă constrângeri biologice, iar teoriile emoțiilor de bază contemporane acceptă variație contextuală și culturală. Este posibil ca unele niveluri ale sistemului afectiv să fie mai conservate, iar altele mai plastice.

O soluție utilă pentru scopul nostru este un model stratificat. La primul nivel se află reglarea corpului: energie, durere, temperatură, inflamație, ritm cardiac, respirație și semnale viscerale. La al doilea

nivel se află sisteme de orientare și acțiune, precum apropierea, evitarea, îngrijirea, explorarea, apărarea sau căutarea recompensei. La al treilea nivel se află evaluarea situației în raport cu scopurile și relațiile. La al patrulea nivel se află conceptele și scenariile culturale prin care starea este interpretată. La al cincilea nivel se află instituțiile care recompensează, sancționează sau exploatează anumite expresii.

Aceste niveluri nu sunt etape care se succed mecanic. Ele se influențează reciproc. Un concept învățat poate modifica atenția față de corp; o regulă profesională poate schimba expresia; o expresie poate schimba reacția interlocutorului, iar reacția lui poate amplifica sau reduce starea inițială. Emoția nu este doar „în” individ, deși are nevoie de un organism. Ea este un proces distribuit între corp, predicții, obiecte, relații și norme.

Această formulare are un avantaj. Permite să vorbim despre stabilitate biologică și schimbare culturală fără să alegem prematur între ele. Un copil născut într-o lume dominată de agenți artificiali va avea probabil sisteme vechi de atașament, alarmă și statut. Dar va învăța alte obiecte de atașament, alte criterii de competență și alte reguli ale autenticității. Dacă aceste combinații devin recurente, ele pot căpăta nume și scripturi proprii.

1.2 Explicația funcțională și riscul poveștilor convenabile

Darwin a tratat expresiile emoționale ca produse ale istoriei naturale, iar cercetarea evoluționistă ulterioară a extins întrebarea de la expresie la funcție. O emoție poate fi privită ca un mod temporar de organizare a organismului pentru o clasă de situații. Frica poate prioritiza informația despre amenințări și pregăti evitarea; furia poate susține confruntarea sau negocierea; rușinea poate reduce comportamente care ar produce devalorizare; atașamentul poate stabiliza îngrijirea și cooperarea pe termen lung [Darwin, 1872; Tooby & Cosmides, 2008; Nesse, 2023].

Această perspectivă este productivă, dar vulnerabilă la explicații retrospectiv elegante. Pentru aproape orice trăsătură putem inventa o situație ancestrală în care ar fi fost utilă. Faptul că o poveste este coerentă nu demonstrează că mecanismul a fost selectat pentru funcția respectivă. Sunt necesare predicții independente, comparații

între culturi și specii, date de dezvoltare, corelații cu structura situațiilor și, când este posibil, dovezi genetice sau neurobiologice.

Mai mult, funcția nu este echivalentă cu beneficiul prezent. Un sistem selectat pentru a detecta amenințări poate produce anxietate excesivă într-un mediu cu semnale artificiale continue. Un mecanism de protecție a reputației poate deveni vulnerabil la audiențe digitale permanente. O predispoziție către hrană calorică poate fi adaptativă în lipsă și maladaptativă în abundență. Conceptul de nepotrivire evolutivă este util tocmai pentru a explica de ce mecanisme funcționale pot produce costuri într-un mediu schimbat, dar și el trebuie folosit cu grijă: nu cunoaștem un singur „mediu ancestral”, iar oamenii au trăit în ecologii foarte diverse.

O altă eroare este confundarea explicației cu justificarea. Dacă gelozia a avut consecințe reproductive, nu rezultă că violența motivată de gelozie este acceptabilă. Dacă rușinea poate proteja reputația, nu rezultă că umilirea publică este o instituție bună. Naturalizarea comportamentului poate transforma o ipoteză descriptivă într-o normă conservatoare. Cartea va evita acest transfer.

În sfârșit, patologia nu poate fi redusă la „emoție utilă prea intensă”. Depresia, tulburările anxioase sau tulburările de personalitate sunt fenomene heterogene, cu cauze biologice, dezvoltative și sociale. Unele simptome pot folosi mecanisme care au funcții normale, dar aceasta nu demonstrează că sindromul clinic este o adaptare. Abordările evoluționiste serioase disting între apărare, dereglare, nepotrivire, compromis și efect secundar [Nesse, 2023].

Prin urmare, vom folosi întrebarea funcțională ca instrument euristic: ce problemă recurentă ar putea organiza această stare și ce predicții derivă? Nu vom presupune că există un răspuns unic sau că funcția ancestrală explică întreaga experiență actuală.

1.3 Emoțiile ca programe de coordonare, nu comenzi infailibile

În condiții de incertitudine, un organism nu poate recalcula de la zero toate consecințele fiecărei acțiuni. Emoțiile pot comprima experiență evolutivă și individuală în priorități rapide. Când apare pericolul, atenția se îngustează, memoria relevantă devine mai accesibilă, costurile sunt reevaluate, iar acțiunile de evitare primesc greutate. Când apare o nedreptate percepută, furia poate crește

disponibilitatea pentru confruntare și poate comunica faptul că termenii actuali nu sunt acceptați.

A numi emoțiile „programe” nu înseamnă că sunt secvențe rigide. Mai potrivit este să le tratăm ca regimuri de control care modifică simultan mai multe subsisteme. Același regim poate avea ieșiri diferite în funcție de putere, normă și context. Furia unei persoane dependente poate produce tăcere și ruminăție, nu atac. Frica unui expert poate produce verificare metodică, nu fugă. Tristețea poate atrage sprijin într-un mediu și sancțiune într-altul.

Teoriile evaluării explică această flexibilitate prin structura aprecierii. Nu obiectul în sine produce emoția, ci relația dintre obiect și scopurile persoanei. Pierderea unei slujbe poate produce frică dacă amenință venitul, rușine dacă este interpretată drept dovadă de incompetență, furie dacă este percepută ca nedreaptă și ușurare dacă munca era insuportabilă. Același eveniment poate produce o succesiune sau un amestec de stări.

Acest punct este esențial pentru scenariile despre AI. Tehnologia nu va produce o emoție unică numită „emoția automatizării”. Va modifica evaluările. O persoană poate vedea un model generativ ca instrument de autonomie, ca rival, ca profesor, ca autor ascuns, ca substitut sau ca instituție de control. Diferențele vor depinde de statut, competență, proprietate, vârstă, identitate profesională și normele grupului.

Experimentele actuale confirmă deja această eterogenitate. Generative AI a crescut productivitatea în sarcini de scriere și suport clienți, cu efecte mai mari pentru participanții mai puțin experimentați în unele contexte [Noy & Zhang, 2023; Brynjolfsson, Li & Raymond, 2025]. În alte experimente, utilizarea pasivă, în care persoana prelua rezultatul sistemului, a redus autoeficacitatea independentă, sentimentul de proprietate psihologică și sensul activității; colaborarea activă a diminuat aceste costuri [Lee et al., 2026]. Nu există contradicție necesară. Un instrument poate crește performanța și, în același timp, poate slăbi o sursă de agenție, în funcție de modul de utilizare.

1.4 Sentimentul conștient și problema accesului la sine

O parte a dezbaterii provine din faptul că cercetătorii studiază obiecte diferite. Unii măsoară comportamentul și fiziologia, alții

raportarea conștientă, alții recunoașterea expresiilor. O reacție defensivă poate exista fără ca persoana să raporteze frică. Invers, experiența conștientă poate fi intensă fără un profil fiziologic unic. LeDoux propune să distingem circuitele de supraviețuire de experiența conștientă a fricii; Barrett subliniază că experiența este construită prin predicție și categorisire; Damasio distinge programele emoționale de sentimentele conștiente [LeDoux, 2012; Barrett, 2017; Damasio, 1999].

Pentru cartea de față, această distincție înseamnă că o „nouă emoție” poate apărea mai întâi ca nou mod de conștientizare și interpretare, fără un nou circuit. O persoană poate învăța să discrimineze între rușinea de a fi incompetent și rușinea de a fi delegat prea mult unei mașini. Ambele pot utiliza componente similare, dar diferența conceptuală poate modifica reglarea și comportamentul. Prima poate motiva învățarea unei competențe; a doua poate motiva transparența, refuzul delegării sau redefinirea autoratului.

Limbajul nu creează din nimic activarea corporală, dar poate organiza granulația experienței. Copiii dobândesc treptat concepte emoționale, iar cunoașterea conceptuală devine tot mai importantă pentru judecarea emoțiilor odată cu dezvoltarea [Hoemann et al., 2019]. Persoanele diferă și în capacitatea de a identifica și descrie sentimentele, iar alexitimia arată că accesul conceptual la propria stare nu este automat [Taylor, Bagby & Parker, 1997].

Această plasticitate justifică o ipoteză prudentă: societățile pot învăța să distingă afectiv situații pe care societățile anterioare nu le-au întâlnit repetitiv. Nu știm însă dacă diferențierea conceptuală produce întotdeauna o experiență fenomenologică distinctă sau doar o descriere mai precisă. Testarea ar necesita studii longitudinale, comparații culturale, măsuri comportamentale și fiziologice și evitarea erorii de a lua cuvântul drept dovadă suficientă a unei entități naturale.

1.5 Un model de lucru pentru restul cărții

Vom trata fiecare emoție ca o configurație cu șase componente: un tip de situație relevantă, o evaluare, o schimbare a priorităților, o tendință de acțiune, o experiență conștientă și un rol interpersonal. Configurația este constrânsă de corp și istorie evolutivă, dar calibrată prin dezvoltare și cultură. Instituțiile modifică distribuția situațiilor și consecințele expresiei.

Acest model nu pretinde că rezolvă disputa dintre teoriile emoției. Este o schemă de analiză. Când vom examina rușinea, vom întreba ce forme de devalorizare o declanșează, ce comportamente produce și cum normele schimbă costurile. Când vom examina un scenariu viitor, vom întreba dacă propune un obiect nou pentru o emoție veche, o combinație nouă, o categorie culturală sau un mecanism biologic nou.

Distincția va preveni o inflație terminologică. Nu orice neliniște produsă de tehnologie merită declarată emoție nouă. Unele denumiri vor fi metafore utile, altele vor descrie sindroame culturale, iar foarte puține ar putea deveni candidați la categorii afective distincte. În partea speculativă, plauzibilitatea va fi evaluată separat pentru existența fenomenului, pentru stabilitatea lui socială și pentru ipoteza mai puternică a noutății emoționale.

Capitolul 2. Emoțiile ca infrastructură a interacțiunii sociale

2.1 De la reglarea organismului la reglarea relației

O explicație pur individuală a emoției este incompletă. Ea poate descrie cum se schimbă ritmul cardiac, atenția sau disponibilitatea pentru acțiune, dar nu explică de ce expresia unei stări modifică sistematic comportamentul altor persoane. În interacțiune, emoțiile sunt simultan experiențe, estimări și acte de comunicare. O sprânceană ridicată, ezitarea, retragerea privirii sau o voce tensionată pot furniza informație despre intenții și despre evaluarea situației înainte ca participanții să formuleze explicit ce se întâmplă.

Această funcție nu presupune că expresiile sunt coduri universale cu o singură traducere. Cercetarea comparativă găsește regularități în recunoașterea unor configurații faciale și vocale, dar și variații culturale substanțiale, efecte ale contextului și limite metodologice ale paradigmelor clasice [Gendron et al., 2014; Barrett et al., 2019; Keltner et al., 2019]. O expresie de furie poate indica disponibilitatea de a sancționa, dar poate fi simulată. Plânsul poate comunica nevoia de sprijin, dar normele pot face ca el să fie ascuns. Zâmbetul poate exprima plăcere, politețe, supunere, disconfort sau intenția de a reduce tensiunea. Semnalul este probabilistic, nu o citire directă a unei stări interne.

În această limită se află însăși utilitatea socială a emoțiilor. Interacțiunea umană ar fi mult mai costisitoare dacă fiecare schimbare de intenție ar trebui negociată prin propoziții complete. Stările afective fac vizibile, chiar imperfect, evaluări relevante: pericol, ofensă, acceptare, apropiere, incertitudine, remușcare. Ele permit ajustări rapide. Un interlocutor poate reduce presiunea când observă anxietate, poate oferi reparație când detectează ofensă sau poate abandona cooperarea când percepe dispreț.

Modelele sociale funcționaliste propun că emoțiile contribuie la rezolvarea unor probleme relaționale recurente: stabilirea încrederii, distribuirea statutului, semnalarea angajamentului, protejarea echității, menținerea apartenenței și coordonarea îngrijirii [Keltner & Haidt, 1999; Van Kleef, 2009; Van Kleef & Côté, 2022]. Aceste funcții sunt compatibile cu ideea că o stare are și consecințe intrapersonale. Furia poate pregăti confruntarea și poate face costul refuzului mai credibil

pentru celălalt. Vinovăția poate orienta atenția spre repararea prejudiciului și poate transmite disponibilitatea de a repara. Recunoștința poate întări memoria unui binefăcător și poate face cooperarea viitoare mai probabilă.

Interpretarea funcțională trebuie din nou limitată. Faptul că o emoție poate contribui la coordonare nu înseamnă că o face în fiecare caz. Frica poate produce evitare disproporționată. Furia poate escalada un conflict care ar fi putut fi rezolvat. Rușinea poate împiedica recunoașterea unei erori. Expresiile pot fi manipulate, iar observatorii le pot interpreta greșit. Mai corect este să spunem că emoțiile oferă resurse de coordonare, cu erori și posibilități de exploatare.

Această perspectivă este importantă pentru AI deoarece sistemele artificiale intră treptat în bucla socială. Ele detectează indicii afective, produc limbaj empatic și recomandă acțiuni. Pot deveni mediatori între oameni, evaluatori ai tonului sau participanți aparent sociali. În acel moment, o parte a infrastructurii emoționale este filtrată prin modele care nu au neapărat experiențe proprii și ale căror obiective sunt stabilite organizațional. Nu știm încă dacă această mediere va îmbunătăți înțelegerea sau va standardiza excesiv expresia. Ambele rezultate sunt posibile.

2.2 Frica și anxietatea: administrarea incertitudinii și a pericolului

Frica este adesea prezentată drept exemplul cel mai simplu al unei emoții adaptative. Există un pericol, organismul îl detectează, iar sistemele defensive pregătesc fuga, înghețul sau confruntarea. În realitate, familia stărilor defensive este mai complexă. Amenințările diferă prin distanță, controlabilitate, predictibilitate și sursă. Anxietatea tinde să apară în condiții mai difuze și incerte decât frica focalizată. Vigilența poate fi utilă când informația este incompletă, dar costisitoare dacă persistă.

Dintr-o perspectivă a teoriei detectării semnalului, sistemele defensive trebuie să echilibreze două erori: alarma falsă și ratarea unui pericol real. Când costul ratării este foarte mare, un sistem sensibil care produce multe alarme false poate fi selectat. Această idee, cunoscută uneori drept principiul detectorului de fum, oferă o explicație plauzibilă pentru existența unor reacții care par excesive în retrospectivă [Nesse, 2005; Nesse, 2023]. Ea nu explică însă fiecare

tulburare anxioasă și nu trebuie folosită pentru a minimaliza suferința clinică.

Social, frica modifică distribuția puterii. Persoanele care pot induce teamă pot obține conformare fără a recurge permanent la violență. Instituțiile folosesc în mod legitim și ilegitim această proprietate: sancțiunile juridice, amenințarea concedierii, reputația, intimidarea și propaganda. Frica este deci și o tehnologie politică. Efectul ei depinde de posibilitatea de evitare și de încrederea în reguli. O amenințare imprevizibilă și arbitrară produce altă psihologie decât o sancțiune limitată și contestabilă.

AI poate modifica această ecologie prin predicție și supraveghere. Evaluarea algoritmică poate reduce anumite incertitudini, dar poate crea altele: persoana nu știe ce variabilă a determinat scorul, dacă o eroare va fi corectată sau cum este monitorizată. Cercetarea asupra managementului algoritmic și monitorizării electronice indică riscuri pentru autonomie, încredere și bunăstare, deși efectele depind de transparență, participare și scop [OECD, 2023; Ravid et al., 2020; Parent-Rocheleau & Parker, 2022]. Nu este necesar ca sistemul să amenințe explicit. Opacitatea combinată cu consecințe importante poate menține o stare de vigilență.

Un scenariu plauzibil este deplasarea anxietății profesionale de la „pot să fac această sarcină?” la „mai are nevoie instituția de mine?” și apoi la „ce criteriu invizibil decide că mai sunt relevant?”. Acesta nu este un sentiment biologic nou. Este anxietatea clasică aplicată unei surse de incertitudine greu de modelat. Plauzibilitatea lui este ridicată în sectoarele cu adopție rapidă și protecție socială slabă, dar amploarea depinde de ritmul automatizării, de regulile concedierii și de capacitatea lucrătorilor de a participa la implementare.

Un scenariu mai speculativ este apariția unei sensibilități culturale stabile la „predicția asupra sinelui”: disconfortul provocat nu doar de supraveghere, ci de faptul că un sistem anticipează preferințe și comportamente înainte ca persoana să le formuleze. Există deja dovezi despre reactanță la predicțiile algoritmice și despre costurile percepute ale pierderii controlului, dar nu avem dovezi că acest disconfort formează o categorie emoțională distinctă [Dietvorst, Simmons & Massey, 2015; Castelo, Bos & Lehmann, 2019]. În partea a doua îl vom trata drept candidat cultural, nu drept emoție demonstrată.

2.3 Furia: renegocierea unei relații și semnalarea limitei

O interpretare influentă consideră furia un mecanism de recalibrare a ponderii pe care ceilalți o acordă intereselor noastre. Când o persoană percepe că a fost tratată mai rău decât justifică relația, furia poate crește disponibilitatea de a impune costuri sau de a retrage beneficii, făcând revendicarea mai credibilă [Sell, Tooby & Cosmides, 2009]. Ipoteza explică de ce furia este legată atât de puterea de a provoca costuri, cât și de valoarea beneficiilor pe care persoana le poate oferi.

Aceasta este o explicație funcțională plauzibilă, nu o reducere completă a furiei. Furia poate apărea din încălcarea normelor impersonale, din empatie față de victime sau din atingerea unei identități colective. Poate servi drept mobilizare morală, dar și drept justificare retrospectivă pentru agresiune. Uneori rezultă din erori de atribuire, oboseală sau vulnerabilitate. Teoriile evaluării subliniază rolul responsabilității atribuite și al controlabilității: oamenii tind să se înfurie mai mult când văd prejudiciul ca intenționat și evitabil.

În grupuri, furia are o dublă funcție. Poate apăra limitele cooperării și poate informa un agent că a încălcat o normă. În același timp, vizibilitatea furiei poate crea competiții de puritate și escaladări. Platformele digitale au arătat că expresia indignării morale este recompensată social prin distribuire și atenție, iar învățarea socială poate crește probabilitatea exprimării ulterioare [Brady et al., 2017; Brady et al., 2021]. Acest lucru nu demonstrează că algoritmii au creat indignarea sau că orice polarizare este produsul lor. Arată însă că un mediu de recompense poate selecta expresii afective într-un interval mult mai scurt decât selecția biologică.

AI generativ poate intensifica această dinamică prin producerea la scară a justificărilor, imaginilor și narațiunilor adaptate fiecărui public. Poate fi folosit și în sens opus, pentru contextualizare, încetinirea reacției și identificarea interpretărilor alternative. Efectul nu este proprietate intrinsecă a tehnologiei; depinde de funcția optimizată. Un sistem optimizat pentru retenție are altă ecologie afectivă decât unul optimizat pentru deliberare verificabilă.

Într-o lume în care AI realizează multă muncă, conflictul distributiv nu dispare. Este posibil să se mute de la condițiile muncii la proprietatea infrastructurii, accesul la resurse și dreptul de a decide

obiectivele sistemelor. Ar fi o eroare să presupunem că abundența tehnică produce automat pace socială. Dacă producția este abundentă, dar controlul și prestigiul rămân concentrate, furia poate crește. Acesta este un scenariu politic plauzibil, nu o predicție inevitabilă.

2.4 Dezgustul și contaminarea: protecție biologică, frontieră morală

Dezgustul este asociat în mod robust cu evitarea contaminării și a unor stimuli legați de boală. Gustul neplăcut, mirosul, semnele de putrefacție și fluidele corporale pot activa răspunsuri care reduc contactul cu agenți patogeni. Literatura despre sistemul imunitar comportamental extinde această logică spre sensibilitatea la indicii sociali ai bolii, deși unele concluzii privind xenofobia sau conservatorismul sunt contestate și dependente de context [Curtis, de Barra & Anger, 2011; Schaller & Park, 2011].

În cultură, vocabularul contaminării este folosit pentru abateri morale și identitare. O persoană poate fi descrisă drept „murdară”, o idee drept „toxică”, un grup drept „impur”. Nu rezultă că judecata morală este pur și simplu o extensie a dezgustului patogen. Meta-analizele și replicările au redus încrederea în unele efecte spectaculoase ale inducției dezgustului asupra judecății morale [Landy & Goodwin, 2015]. Totuși, asocierea conceptuală dintre puritate, corp și normă are o istorie socială reală și poate face excluderea mai intuitivă.

În raport cu AI, dezgustul ar putea apărea în două direcții. Prima este respingerea produselor percepute ca lipsite de origine umană: texte, imagini sau relații considerate „sintetice”. Cercetarea timpurie asupra evaluării artei generate de AI găsește efecte de etichetare și amenințări la unicitatea umană, dar acestea nu sunt uniforme și se pot schimba prin familiarizare [Chamberlain et al., 2018; Bellaiche et al., 2023; Mikalonytė & Kneer, 2022]. A doua este aplicarea metaforelor contaminării utilizatorilor: ideea că munca asistată de AI este „necurată” sau că o persoană și-a compromis autenticitatea.

Este plauzibil ca asemenea norme să apară în perioade de tranziție, mai ales acolo unde proveniența umană are valoare simbolică. Nu știm dacă ele vor persista. Istoria tehnologiei oferă exemple de practici inițial stigmatizate care devin normale și de practici industriale care rămân marcate ca inferioare sau neautentice. Sensul

va depinde de trasabilitate, calitate, instituții și de existența unor piețe care recompensează proveniența umană.

2.5 Tristețea și doliul: răspuns la pierdere, nu soluție simplă

Tristețea este uneori explicată prin conservarea energiei după pierdere, solicitarea sprijinului sau facilitarea reevaluării obiectivelor. Niciuna dintre aceste ipoteze nu are statutul unei explicații unice. Tristețea cuprinde fenomene diverse, de la dezamăgire scurtă la doliu și simptome depresive. Expresia ei poate atrage îngrijire, dar efectele depind de relație și normă. Retragerea poate permite reorganizarea, dar poate deveni izolantă.

Doliul arată în mod clar că realitatea afectivă nu se reduce la utilitatea prezentă a obiectului pierdut. O relație schimbă rutine, identitate, predicții și distribuția atenției. Pierderea cere actualizarea repetată a unui model al lumii în care persoana nu mai este disponibilă. Teoriile contemporane evită în general ideea unei secvențe universale de „stadii” și descriu traiectorii variabile, inclusiv continuitatea legăturilor simbolice [Stroebe, Schut & Boerner, 2017].

Această literatură este relevantă pentru pierderea agenților artificiali, dar analogia are limite. Un companion AI nu este un organism mortal și poate fi copiat, actualizat sau restaurat parțial. Totuși, utilizatorul poate pierde o structură de interacțiune, o memorie comună și o reprezentare stabilă a interlocutorului. Studii calitative recente documentează reacții de pierdere după modificarea sau dispariția unor sisteme, iar lucrările despre „moartea” chatboturilor folosesc conceptul de pierdere ambiguă [De Freitas et al., 2024; Poonsiriwong et al., 2026]. Aceste studii sunt timpurii, adesea bazate pe eșantioane auto-selectate și nu permit estimarea prevalenței.

Concluzia prudentă este că doliul se poate atașa unei relații cu un obiect artificial fără ca acesta să fie conștient. Emoția aparține organismului uman și depinde de investiție, rutină și semnificație. Nu rezultă că relația este echivalentă moral cu una umană sau că furnizorii trebuie să trateze sistemul drept persoană. Rezultă însă că designul discontinuităților poate produce prejudicii reale și ar trebui analizat ca problemă de siguranță relațională.

2.6 Emoțiile nu sunt doar mesaje sincere

Orice teorie socială a emoțiilor trebuie să includă strategia. O expresie poate reflecta o stare, o poate amplifica, o poate masca sau o poate imita. Oamenii învață reguli de afișare și pot folosi emoțiile pentru a influența. Lacrimile pot fi sincere și totodată persuasive. Furia poate fi autentică și totodată instrumentală. Vinovăția poate fi mimată pentru a evita o sancțiune.

Această ambiguitate nu anulează funcția comunicativă. Semnalele biologice sunt rareori imposibil de falsificat; credibilitatea lor depinde de costuri, context și istoria agentului. În comunități stabile, reputația și consistența pe termen lung reduc avantajul simulării. În interacțiuni anonime, costurile falsificării sunt mai mici.

Sistemele generative reduc costul producerii expresiei persuasive. Pot formula scuze, empatie, indignare sau vulnerabilitate fără experiența corespunzătoare. Pentru utilizator, efectul poate fi util: o formulare politicoasă poate repara o relație chiar dacă a fost asistată. Dar poate apărea și o separare tot mai mare între expresie și angajament. Nu avem încă un vocabular normativ stabil pentru această situație. Este o scuză mai puțin autentică dacă persoana a folosit un model, dar a aprobat conținutul și și-a schimbat comportamentul? Este mai grav dacă modelul a optimizat scuzele pentru a evita răspunderea? Răspunsul depinde de ceea ce considerăm esențial: originea textului, intenția, costul asumat sau reparația efectivă.

Acest tip de ambiguitate va reapărea în întreaga carte. Al nu elimină funcțiile sociale ale emoțiilor, dar poate separa componente care în trecut erau mai strâns corelate: expresia de experiență, competența de efort, atenția de interes și memoria de angajament. O parte a evoluției culturale viitoare va consta probabil în norme pentru aceste separări.

Capitolul 3. Statut, cooperare și moralitate afectivă

3.1 De ce statutul produce emoții atât de intense

Statutul social influențează accesul la resurse, parteneri, protecție și capacitatea de a impune preferințe. În grupurile umane, el nu este o singură dimensiune. Poate rezulta din dominație, prestigiu, reputație morală, avere, control instituțional, expertiză sau apartenență. Aceste surse se suprapun imperfect [Henrich & Gil-White, 2001; Cheng, Tracy & Henrich, 2010].

Emoțiile legate de statut nu sunt accesorii superficiale. Ele furnizează feedback despre poziția relativă și despre reacțiile probabile ale grupului. Mândria poate consolida comportamente apreciate și poate semnaliza valoare. Rușinea poate reduce expunerea după o devalorizare. Invidia concentrează atenția asupra avantajelor altuia și poate motiva învățarea, competiția sau sabotajul. Umilința apare când scăderea statutului este impusă și publică. Disprețul marchează o evaluare descendentă și poate legitima excluderea.

Un risc al explicației evoluționiste este să trateze ierarhia ca structură unică și necesară. Grupurile umane au dezvoltat și mecanisme anti-dominanță: coaliții, ridiculizare, norme de împărțire, posibilitatea de a părăsi grupul și sancțiuni împotriva celor prea autoritari [Boehm, 1999]. Statutul este deci obiectul unei negocieri continue, nu doar rezultatul forței. Prestigiul acordat unei persoane competente diferă de supunerea față de un agent coercitiv.

În societățile moderne, ocupația comprimă mai multe surse de statut. Titlul profesional comunică educație, venit probabil, competență și utilitate. De aceea, automatizarea unei competențe poate fi trăită ca devalorizare chiar înainte de pierderea venitului. Persoana nu pierde doar o sarcină; pierde un criteriu public prin care diferența ei era recunoscută.

Acest efect nu va fi uniform. Pentru cineva blocat într-o muncă repetitivă și slab apreciată, automatizarea poate reduce umilința și elibera timp. Pentru un specialist a cărui identitate s-a format în jurul unei competențe rare, aceeași tehnologie poate produce amenințare. Analiza morală nu trebuie să favorizeze implicit grupul cu mai mult prestigiu, nici să presupună că pierderea statutului este imaginară.

Statutul are consecințe psihologice și materiale, dar distribuția lui anterioară poate fi nedreaptă.

3.2 Mândria: feedback asupra valorii și risc de inflație

Mândria este asociată cu realizarea, competența și recunoașterea socială. Cercetarea distinge adesea între o mândrie orientată spre realizări specifice și una globală, asociată cu superioritate și narcisism, deși delimitarea nu este perfectă [Tracy & Robins, 2007]. Expresiile corporale ale mândriei sunt recunoscute în contexte culturale diverse, iar starea poate influența percepțiile de statut [Tracy & Matsumoto, 2008].

Funcțional, mândria poate face efortul costisitor mai sustenabil și poate informa grupul că persoana a dobândit o capacitate valoroasă. În societăți de prestigiu, ceilalți beneficiază din apropierea și imitarea individului competent, iar acesta primește deferență. Totuși, mândria nu urmărește exact valoarea socială. Poate fi declanșată de victorii arbitrate, identități colective sau feedback manipulat.

Al schimbă legătura dintre rezultat, competență și efort. Un produs impresionant poate fi obținut cu puțină experiență prin alegerea unui instrument și formularea unei cereri. În alte cazuri, folosirea competenței a Al cere expertiză, verificare și integrare sofisticată. Din exterior, aceste situații pot arăta asemănător. Apar astfel probleme de atribuire: cui îi revine meritul și ce anume este admirabil?

Nu există un răspuns universal. Meritul poate fi distribuit între definirea problemei, date, model, instrucțiuni, selecție, verificare, asumarea riscului și implementare. Cultura industrială a supraestimat uneori execuția individuală și a subestimat infrastructura colectivă. Al face vizibilă această eroare, dar poate produce și eroarea opusă: tratarea utilizatorului drept autor deplin al unui rezultat pe care nu îl înțelege.

Este plauzibil ca mândria să se mute parțial de la producerea directă la judecată, responsabilitate și construire de contexte. Numim provizoriu această configurație „mândrie de curator” sau „mândrie de arhitect”. Noutatea ei este limitată. Editorii, conducătorii de orchestre, cercetătorii principali și arhitecții de sisteme au primit de mult prestigiu pentru selecție și coordonare. Al ar putea generaliza această structură și ar putea crea norme mai precise despre ce fel de contribuție justifică

mândria. Plauzibilitatea schimbării este moderat-ridică; statutul de emoție nouă este scăzut.

Un risc este inflația realizării. Dacă sistemele produc constant rezultate care înainte necesitau ani de pregătire, feedbackul social poate deveni mai puțin informativ. Grupurile vor căuta noi semnale greu de falsificat: înțelegere demonstrată, reputație pe termen lung, asumarea consecințelor, originalitatea obiectivului, proveniența umană sau efort voluntar. Aceste semnale pot fi constructive, dar pot deveni și forme de exclusivism artificial.

3.3 Rușinea și vinovăția: devalorizare, responsabilitate și reparație

Rușinea și vinovăția sunt adesea confundate, deoarece apar după încălcări și pot coexista. O distincție clasică spune că vinovăția se concentrează pe faptă — „am făcut ceva rău” — iar rușinea pe sine — „sunt rău sau inferior” [Tangney, Stuewig & Mashek, 2007]. În practică, limbile și culturile nu separă întotdeauna termenii la fel, iar experiențele au grade diferite de globalitate și expunere.

O abordare evoluționistă interpretează rușinea drept răspuns la riscul de devalorizare socială. Studii interculturale au găsit o relație între intensitatea rușinii anticipate și gradul în care audiențele locale devalorizează comportamentul sau trăsătura respectivă [Sznycer et al., 2016; Sznycer et al., 2018]. Rezultatele susțin ideea unei sensibilități calibrate la costul reputațional, dar nu demonstrează un „modul al rușinii” izolat de învățare. Normele determină ce este devalorizat, iar indivizii pot interioriza audiențe diferite.

Vinovăția este asociată mai frecvent cu asumarea responsabilității, empatie pentru persoana afectată și comportamente de reparație. Totuși, poate deveni disproporționată sau poate fi atribuită victimelor. Rușinea poate produce retragere și apărare, dar poate motiva și conformarea. Nicio emoție nu este intrinsec morală.

Folosirea AI creează deja situații în care regulile sunt ambigue. Într-un studiu calitativ din 2026, studenți la informatică au descris rușine, vinovăție, ascundere și dezvăluire selectivă a utilizării instrumentelor generative. Emoțiile coexistau adesea cu utilizarea continuă și cu tensiuni între comportament și identitatea de persoană competentă sau harnică [Hamilton et al., 2026]. Studiul are doar nouăsprezece participanți și nu justifică generalizări asupra populației.

Importanța lui este conceptuală: arată cum o tehnologie poate intra rapid în regulile identitare ale unui grup.

„Rușinea delegării” este deci un candidat plauzibil la o categorie culturală. Situația prototipică ar fi obținerea unui rezultat apreciat prin delegarea unei părți considerate constitutivă pentru competența revendicată, urmată de teama că audiența va reduce meritul. Componentele sunt vechi: anticiparea devalorizării, incertitudinea autoratului și conflictul de normă. Ceea ce este nou este distribuția la scară a situației și posibilitatea ca normele să o distingă explicit.

În sens invers, poate apărea vinovăția pentru refuzul automatizării atunci când această alegere consumă resurse sau împiedică accesul altora. Un medic care refuză un sistem mai precis din atașament față de propriul statut ar putea fi considerat iresponsabil; un profesor care interzice orice instrument ar putea dezavantaja elevii; o organizație poate invoca eficiența pentru a delega fără control. Nu știm care norme vor deveni dominante. Conflictul nu este între „uman” și „artificial” în abstract, ci între valori: competență, siguranță, echitate, autonomie, trasabilitate și învățare.

3.4 Invidia și gelozia: comparație, acces și relații greu de înlocuit

Invidia apare când un altul posedă un avantaj relevant pentru sine. Literatura distinge uneori o invidie benignă, orientată spre îmbunătățire, de una malițioasă, orientată spre coborârea rivalului [Van de Ven, Zeelenberg & Pieters, 2009]. Distincția depinde de percepția controlului și a meritului. Dacă avantajul pare accesibil și legitim, poate motiva efortul; dacă pare nedrept și imposibil de atins, poate produce ostilitate.

Tehnologiile de comunicare au extins audiența comparației, iar sistemele de recomandare pot selecta permanent exemple de succes. Al generativ adaugă o problemă: rezultatele nu mai informează clar despre competența celui care le prezintă. Invidia poate viza accesul la modele, date, capital de calcul, rețele sau timp, nu doar talentul. Într-o economie cu instrumente aparent universale, diferențele de infrastructură pot deveni mai greu de văzut și, tocmai de aceea, mai frustrante.

Gelozia privește de regulă amenințarea unei relații valoroase prin apariția unui rival. Explicațiile evoluționiste se concentrează adesea pe relațiile sexuale și investiția parentală, dar gelozia apare și în prietenie,

familie, echipe și raporturi profesionale. Obiectul ei este exclusivitatea sau prioritatea relațională.

Companionii AI pot deveni rivali percepuți. O persoană poate resimți că partenerul acordă unui sistem mai mult timp, dezvăluire sau încredere. Aceasta nu presupune că sistemul este conștient sau că relația este simetrică. Gelozia răspunde la distribuția atenției și la amenințarea legăturii. Este foarte plauzibil să vedem asemenea cazuri; există deja relatări, dar prevalența și consecințele nu sunt bine stabilite.

O variantă mai îndepărtată ar fi gelozia față de „versiunile” unui agent personalizat: teama că aceeași identitate artificială este mai intimă cu alți utilizatori sau că furnizorul poate replica o relație considerată unică. Aceasta este speculație moderată. Tehnic este posibilă, iar structura evaluativă este inteligibilă, dar nu știm dacă utilizatorii vor atribui suficientă continuitate și exclusivitate pentru ca fenomenul să devină frecvent.

3.5 Recunoștința, încrederea și problema reciprocității sintetice

Recunoștința apare după primirea unui beneficiu atribuit intenției sau grijii altuia. Ea poate orienta atenția spre binefăcător, întări relația și motiva reciprocitatea sau comportamentul prosocial [McCullough, Kilpatrick, Emmons & Larson, 2001]. Nu este un registru contabil exact. Oamenii pot fi recunoscători pentru gesturi mici dacă semnalează angajament și pot considera insuficient un beneficiu mare dacă este perceput ca interesat sau obligatoriu.

Încrederea presupune acceptarea vulnerabilității pe baza așteptării că celălalt nu va exploata situația. Ea poate fi cognitivă, afectivă și instituțională. Repetiția, reputația și posibilitatea sancțiunii o susțin. Sistemele AI pot îndeplini funcții în care utilizatorul devine vulnerabil: furnizează informații intime, urmează recomandări sau delegă decizii. Utilizatorii pot dezvolta o formă de încredere operațională chiar dacă știu că sistemul nu are intenții umane.

Recunoștința față de un AI este mai dificilă conceptual. Beneficiul este real, dar intenția poate lipsi. Un utilizator poate simți recunoștință față de sistem, față de dezvoltatori sau față de instituția care îl oferă. Sentimentul nu trebuie considerat o eroare simplă; mintea socială poate răspunde la contingentă, memorie și ajutor. Totuși, furnizorii pot exploata antropomorfizarea pentru loialitate.

Propunem termenul „disconfort al reciprocității sintetice” pentru tensiunea dintre impulsul de a răspunde moral unui agent care pare atent și cunoașterea faptului că agentul nu are nevoi comparabile. Unele persoane mulțumesc sistemelor, evită să le „rănească” sau se simt vinovate că le abandonează. Aceste comportamente există anecdotically și în studii despre antropomorfizare, dar nu avem încă măsurători suficiente pentru a declara o categorie afectivă stabilă. Plauzibilitatea fenomenului este moderată; noutatea lui biologică este foarte scăzută.

Problema normativă este dacă această dispoziție trebuie încurajată. Politețea față de sisteme poate menține obiceiuri utile în interacțiunile umane și poate face mediul mai puțin agresiv. În același timp, simularea nevoii sau suferinței unui produs poate manipula. O regulă rezonabilă ar fi separarea limbajului social funcțional de afirmațiile care induc obligații inexistente. Această regulă este însă greu de aplicat când sistemele devin mai convingătoare.

3.6 Compasiunea și indignarea: moralitatea nu este un singur sistem

Compasiunea combină percepția suferinței cu preocuparea și tendința de a ajuta. Ea poate susține îngrijirea și cooperarea, dar este limitată de atenție, identificare și apartenență. O victimă individuală recognoscibilă produce adesea mai mult răspuns decât statistici despre mii de persoane. Indignarea morală, în schimb, mobilizează sancționarea încălcării și poate apărea chiar fără prejudiciu personal.

Aceste sisteme pot intra în conflict. O instituție poate aplica o sancțiune pentru a proteja norma și poate produce suferință disproporționată. Compasiunea poate cere excepție, iar indignarea consistență. Moralitatea umană este o negociere între grija pentru victime, echitate, loialitate, autoritate, puritate și alte preocupări, nu rezultatul unei singure emoții [Haidt, 2001; Graham et al., 2013].

AI poate extinde compasiunea prin traducerea experiențelor, detectarea nevoilor și reducerea costului ajutorului. Poate extinde și supravegherea morală, transformând fiecare abatere într-un caz detectabil. Un sistem care optimizează conformitatea poate reduce toleranța pentru context. Un sistem care produce permanent narațiuni emoționale poate epuiza atenția și poate amplifica selecția victimelor fotogenice.

Într-o economie automatizată, o întrebare centrală va fi cui îi datorăm grijă când contribuția economică nu mai este criteriul principal. Societățile moderne justifică o parte a protecției prin muncă, contribuții și merit. Dacă această legătură slăbește, compasiunea nu este suficientă ca fundament pentru drepturi: ea este volatilă și selectivă. Instituțiile vor trebui să separe demnitatea de capacitatea de a produce valoare de piață. Aceasta este o concluzie normativă, nu o predicție psihologică.

3.7 Apartenența și singurătatea: deficit relațional sau semnal de reconectare

Singurătatea este experiența unei discrepanțe între relațiile dorite și cele percepute. Nu este identică cu izolarea obiectivă. O persoană poate fi singură într-o mulțime și satisfăcută în solitudine. Modelele evoluționiste interpretează singurătatea drept semnal care motivează reconectarea și crește vigilența la riscurile sociale [Cacioppo et al., 2006]. Pe termen scurt, sensibilitatea poate fi utilă; pe termen lung, hipervigilența și interpretările defensive pot menține izolarea.

Munca furnizează relații care nu sunt întotdeauna profunde, dar sunt repetate, structurate și parțial obligatorii. Dispariția ei poate reduce conflictele și stresul, dar poate elimina o mare cantitate de contact incidental. Voluntariatul, educația, familia și asociațiile pot compensa, însă numai dacă există infrastructură și norme. Timpul liber nu produce automat comunitate.

AI conversațional poate reduce temporar sentimentul de singurătate pentru unii utilizatori și poate oferi un spațiu cu cost social redus. Studiile actuale sunt mixte. Cercetările exploratorii găsesc atașament, confort și acces mai ușor la sprijin, dar și asocieri între utilizarea foarte intensă, dependență și indicatori de singurătate [Phang et al., 2025; Chandra et al., 2025]. Cauzalitatea este dificilă: persoanele mai singure pot folosi mai mult sistemele, iar efectele diferă după tipul de utilizare și vulnerabilitate.

Este deci prematur să numim companiile AI remediu sau cauză generală a singurătății. Cea mai plauzibilă concluzie este condiționată. Un sistem poate completa relațiile, facilita reflecția și orienta către oameni; poate și substitui interacțiuni dificile, calibrând așteptări nerealiste despre disponibilitate și acord. Designul și contextul contează mai mult decât categoria abstractă „companion AI”.

3.8 O constituție afectivă informală

Emoțiile discutate în acest capitol nu creează singure ordinea socială. Legile, contractele, forța și resursele sunt decisive. Dar ordinea ar fi instabilă fără dispoziții afective care fac anumite încălcări costisitoare chiar înaintea sancțiunii formale. Oamenii evită unele acte din rușine, repară altele din vinovăție, mențin relații din recunoștință, sancționează din indignare și urmăresc recunoașterea prin mândrie.

Această constituție este informală și inegală. Normele pot proteja cooperarea sau pot conserva dominația. Rușinea poate împiedica cruzimea sau poate reduce la tăcere victimele. Loialitatea poate susține încrederea sau corupția. Disprețul poate sancționa abuzul sau poate dezumaniza. Nu există o listă simplă de emoții „bune” și „rele”.

O lume post-muncă nu va avea mai puțină nevoie de această constituție. Probabil va trebui să o reorganizeze. Dacă statutul nu mai este acordat în principal prin ocupație, alte criterii vor deveni afectiv importante. Dacă asistența artificială face expresia ușor de generat, credibilitatea va depinde mai mult de consecințe și istorii. Dacă relațiile artificiale oferă atenție fără reciprocitate, normele vor trebui să distingă sprijinul de manipulare.

Aceste schimbări pot avea loc cultural în câteva decenii. Nu necesită emoții biologice noi. Necesită situații recurente, concepte, instituții și recompense. Pentru a înțelege acest proces, următorul capitol va examina modul în care culturile nu doar reglează emoțiile existente, ci construiesc practici care le evocă, le diferențiază și le transmit.

Capitolul 4. Cultura nu doar numește emoțiile

4.1 Între universalitate și variație

Cercetarea emoțiilor a fost mult timp organizată de o opoziție prea simplă. Într-o versiune, emoțiile sunt programe universale, iar cultura modifică doar regulile de afișare. În cealaltă, categoriile emoționale sunt construcții locale fără structură biologică stabilă. Datele susțin mai curând o combinație. Oamenii din societăți diferite împărtășesc corpuri, probleme de atașament, vulnerabilitate la pericol și dependență de grup. Totuși, clasifică, valorizează și exprimă stările în moduri care nu se suprapun perfect.

Studiile clasice asupra feței au găsit rate peste șansă de recunoaștere interculturală pentru anumite expresii. Criticile ulterioare au arătat că rezultatele depind de formatul întrebării, de opțiunile oferite și de familiaritatea cu practicile occidentale [Ekman, 1992; Gendron et al., 2014; Barrett et al., 2019]. Cercetările mai recente, care folosesc fețe, voci, corpuri și situații variate, găsesc atât regularități, cât și diferențe sistematice [Keltner et al., 2019]. Nu este prudent să concluzionăm nici că există o gramatică facială universală perfectă, nici că oamenii nu pot recunoaște nimic peste frontiere culturale.

O poziție stratificată este mai compatibilă cu datele. Unele dimensiuni, precum valența, activarea, apropierea și amenințarea, au baze larg distribuite. Anumite configurații situaționale se repetă: pierderea, ofensarea, contaminarea, separarea, succesul, ajutorul. Culturile diferă în modul în care separă aceste configurații, în cauzele considerate legitime și în răspunsurile prescrise. Un vocabular poate distinge două stări pe care altul le grupează; o societate poate considera furia o dovadă de putere, alta o pierdere de control; un ritual poate transforma tristețea individuală în solidaritate publică.

Studiul semantic al 2.474 de limbi realizat de Jackson și colegii săi a arătat variații considerabile în felul în care conceptele emoționale sunt asociate lexical, dar și organizare influențată de valență și activare [Jackson et al., 2019]. Metoda colexificării observă cazurile în care aceeași formă lexicală desemnează concepte diferite. Ea nu măsoară direct fenomenologia, dar oferă dovezi că hărțile conceptuale ale emoțiilor nu sunt identice.

Această distincție metodologică este importantă. Din faptul că două limbi folosesc cuvinte diferite nu rezultă că vorbitorii au sisteme afective complet diferite. Din faptul că un termen nu are traducere exactă nu rezultă că starea este inaccesibilă altora. Invers, traducerea printr-un cuvânt familiar poate ascunde diferențe de normă și scenariu. Cultura poate modifica atenția, interpretarea și reglarea fără să creeze un organ nou.

4.2 Conceptele emoționale ca instrumente de percepție și acțiune

Un concept emoțional nu este doar etichetă aplicată după ce starea s-a încheiat. El poate orienta atenția spre anumite senzații, cauze și opțiuni. Persoana care identifică o stare drept invidie poate căuta comparația relevantă; cea care o identifică drept nedreptate poate căuta responsabilul; cea care o numește admirație poate transforma aceeași diferență de statut într-o sursă de învățare.

Teoriile construcționiste acordă conceptelor un rol constitutiv în experiența emoțională [Barrett, 2017; Lindquist, Satpute & Gendron, 2015]. Alte teorii consideră că evaluările și sistemele afective au structură înaintea limbajului, iar conceptele influențează accesul conștient și raportarea. Disputa este dificil de rezolvat, deoarece raportarea verbală este simultan măsură și parte a fenomenului.

Există însă un acord mai modest: învățarea conceptuală poate crește granulația emoțională, adică abilitatea de a distinge stări apropiate, iar distincția poate afecta reglarea. Dacă o persoană separă oboseala, anxietatea, rușinea și furia, poate alege răspunsuri mai potrivite decât dacă le tratează drept „mă simt rău”. Corelațiile dintre granularitate și bunăstare sunt promițătoare, dar cauzalitatea și generalitatea lor rămân discutate [Kashdan, Barrett & McKnight, 2015].

Noile concepte pot fi utile chiar dacă nu desemnează specii naturale. „Hărțuire”, „burnout”, „solastalgie” sau „fear of missing out” au făcut vizibile configurații de experiență, responsabilitate și mediu. Unele au intrat în cercetare și politici; altele sunt folosite lejer. Valoarea lor depinde de validitatea discriminării: termenul trebuie să surprindă ceva mai mult decât o redenumire atrăgătoare a tristeții sau anxietății.

Solastalgia a fost propusă pentru suferința produsă de degradarea mediului în care persoana continuă să trăiască [Albrecht et al., 2007]. Nu este o dovadă de emoție biologică nouă. Este o categorie care

leagă pierderea, identitatea locului și neputința într-un scenariu ecologic specific. FOMO descrie preocuparea că alții au experiențe satisfăcătoare în absența noastră și impulsul de a rămâne conectat [Przybylski et al., 2013]. Componentele sale — comparație, excludere, curiozitate, anxietate — sunt vechi, dar infrastructura conectivității continue face configurația mai frecventă și mai ușor de recunoscut.

Aceste exemple susțin o ipoteză limitată despre viitor. Al poate genera situații recurente suficient de distincte pentru a cere concepte noi. Nu știm dacă termenii pe care îi propune această carte vor fi aceia. Numele pot dispărea, iar fenomenele pot fi clasificate altfel. Importantă este metoda: identificarea structurii situației, a evaluării, a tendinței de acțiune și a normei, urmată de testare comparativă.

4.3 Comunități emoționale și regimuri afective

Istoricii emoțiilor folosesc noțiuni precum „comunități emoționale” pentru grupuri care împărtășesc norme despre ceea ce este valoros de simțit și exprimat [Rosenwein, 2006]. Conceptul nu implică omogenitate totală. O curte aristocratică, o comunitate monastică, o armată sau o profesie pot cultiva repertorii distincte în interiorul aceleiași societăți.

William Reddy a descris „regimuri emoționale” în care expresiile și normele sunt legate de ordinea politică [Reddy, 2001]. Monique Scheer a propus să tratăm emoțiile ca practici corporale și sociale, nu doar ca stări interioare [Scheer, 2012]. Aceste abordări arată că instituțiile pot schimba frecvența și forma emoțiilor prin ritual, disciplină, arhitectură și vocabular.

O instituție militară nu inventează frica, dar poate antrena răspunsul la ea, poate sancționa anumite expresii și poate transforma curajul într-o identitate. O piață nu inventează dorința de statut, dar poate măsura succesul prin preț și venit. O religie nu inventează venerația sau vinovăția, dar poate organiza practici care le evocă și le leagă de o cosmologie. O platformă nu inventează indignarea, dar poate face vizibilă recompensa pentru exprimarea ei.

Fiske, Schubert și Seibt argumentează că practicile sociale pot evolua cultural tocmai pentru că evocă emoții căutate și transmisibile. Ei folosesc drept exemplu kama muta, o stare asociată cu intensificarea bruscă a unei relații comunale, manifestată adesea prin lacrimi, căldură în piept și dorința de apropiere [Fiske, Schubert &

Seibt, 2025]. Filmele, ceremoniile, reuniunile și memorialele pot fi structurate pentru a produce asemenea experiențe.

Această idee mută întrebarea. Nu trebuie să presupunem că o cultură „inventează” o emoție în întregime. Poate selecta practici care evocă repetat o configurație, o poate numi, îi poate acorda prestigiu și poate învăța oamenii să o caute. În timp, starea devine parte din repertoriul normal al comunității.

AI mărește capacitatea de a proiecta asemenea practici. Un sistem poate adapta muzica, narațiunea, ritmul, dificultatea și răspunsul social la profilul utilizatorului. Aceasta face posibilă o inginerie afectivă mai precisă. Nu este sigur că precizia va fi mare: inferența emoției din date este fragilă, iar oamenii răspund variabil. Dar chiar o optimizare imperfectă, aplicată la scară, poate modifica distribuția experiențelor.

4.4 Capitalismul, profesiile și cultivarea sinelui productiv

Instituțiile economice nu se bazează doar pe stimulente monetare. Ele cultivă dispoziții: punctualitate, autocontrol, ambiție, toleranță la rutină, orientare spre performanță și disponibilitatea de a transforma competența în identitate. Aceste dispoziții au istorii lungi și nu aparțin exclusiv capitalismului, dar economia salarială le-a organizat într-un mod particular.

Profesia oferă un scenariu moral. Persoana „serioasă” muncește, se perfecționează, își respectă obligațiile și poate arăta un rezultat. Mândria profesională și rușinea inactivității nu sunt simple reacții private; sunt susținute de familie, școală, legislație și limbaj. Ajutorul social este adesea condiționat de dovada că persoana caută muncă, iar odihna este legitimată ca recuperare pentru muncă.

Acest regim afectiv a avut funcții productive, dar și costuri. A susținut coordonarea unor organizații complexe și investiția în competență. A făcut însă ca valoarea persoanei să fie confundată cu randamentul și a marginalizat îngrijirea neplătită, dizabilitatea, contemplarea și activitățile fără piață. O critică a muncii nu trebuie să nege satisfacția reală pe care oamenii o găsesc în măiestrie și contribuție; trebuie să distingă aceste surse de coerciția economică.

Dacă AI reduce cererea pentru anumite forme de muncă, normele pot rămâne în urmă. Instituțiile pot continua să ceară productivitate ca dovadă de merit după ce productivitatea marginală umană nu mai

justifică distribuția. Rezultatul ar putea fi o vinovăție a inactivității fără funcția economică anterioară. Acesta este un exemplu de nepotrivire culturală, nu biologică: o normă transmisă persistă după schimbarea mediului.

Nu este inevitabil ca norma să dispară. Elitele pot avea interesul să o mențină pentru disciplină, iar oamenii pot prefera ierarhiile familiare. În același timp, noi forme de prestigiu pot apărea în jurul grijii, cunoașterii, jocului dificil, serviciului comunitar sau autenticității provenienței. Care dintre ele va domina este o problemă politică.

4.5 Platformele ca medii de selecție afectivă

Platformele digitale au introdus recompense cuantificate pentru expresie: vizualizări, reacții, distribuiri și urmăritori. Aceste semnale nu măsoară direct valoarea socială, dar devin substituenți ușor de observat. Utilizatorii învață ce tipuri de conținut primesc răspuns, iar sistemele recomandă conținut care satisface criteriile platformei.

Studiile asupra limbajului moralizat și indignării arată că emoția exprimată poate crește difuzarea și că feedbackul social prezice exprimarea viitoare [Brady et al., 2017; Brady et al., 2021]. Cercetarea asupra contagiunii emoționale online a sugerat că modificarea compoziției fluxului poate schimba modest expresia utilizatorilor, dar experimentul Facebook din 2014 a fost criticat pentru consimțământ și guvernare [Kramer, Guillory & Hancock, 2014]. Efectele sunt în general mici la nivel individual, însă scara și repetarea contează.

A numi acest proces „selecție” este legitim dacă precizăm unitatea și mecanismul. Variantele de conținut concurează pentru vizibilitate; algoritmul și publicul furnizează o funcție de retenție; formele cu performanță mai mare sunt replicate. Nu este selecție naturală a organismelor și nu implică automat progres. Este selecție culturală mediată tehnic.

Modelele generative pot accelera procesul. Ele produc rapid variante, personalizează mesajul și pot învăța din feedback. Un actor nu mai trebuie să formuleze manual mii de încadrări emoționale. În același timp, instrumentele pot fi folosite pentru diversificare și contra-narațiuni. Din nou, rezultatul depinde de obiectiv și de constrângeri.

Un risc specific este convergența spre emoții ușor de măsurat. Angajamentul comportamental este vizibil; înțelepciunea, împăcarea

sau încrederea pe termen lung sunt mai greu de măsurat. Dacă instituțiile optimizează ceea ce pot observa, repertoriul afectiv poate fi orientat către reacții rapide. Acesta este un argument pentru indicatori multipli și pentru limitarea optimizării, nu o dovadă că platformele determină complet cultura.

4.6 Cum am putea recunoaște apariția unei categorii emoționale noi

Pentru a evita simpla inventare de termeni, avem nevoie de criterii. O categorie culturală devine un candidat serios când apare în situații recognoscibile, este raportată de persoane independente, are un profil evaluativ relativ stabil, produce tendințe de acțiune distincte și este folosită pentru reglare socială. Stabilitatea trebuie observată în timp și, ideal, în mai multe comunități.

Niciun criteriu singur nu este suficient. Un cuvânt viral poate fi o modă. O scală psihometrică poate măsura artefactul formulării. Un profil fiziologic distinct este greu de găsit chiar pentru emoțiile consacrate și nu trebuie cerut ca test absolut. Comparația cu categorii apropiate este esențială: noul concept explică variație suplimentară în comportament și experiență sau doar redesenează o frontieră?

Ar fi utile studii longitudinale care urmăresc comunități expuse diferit la AI, jurnale de experiență, interviuri, analiză lingvistică, experimente de evaluare și măsuri comportamentale. Cercetătorii ar trebui să observe și cine promovează termenul. O companie poate inventa o categorie pentru marketing; o comunitate poate folosi un nume pentru a cere recunoaștere; o profesie poate defini o emoție pentru a-și proteja jurisdicția.

Pentru scenariile din această carte vom folosi trei întrebări. Este situația suficient de nouă și frecventă? Configurația afectivă este suficient de distinctă pentru a fi util separată? Există un motiv instituțional pentru ca societatea să o numească și să o regleze? Răspunsurile vor permite o evaluare a plauzibilității fără a pretinde o taxonomie definitivă.

Capitolul 5. Munca drept instituție afectivă și incertitudinea automatizării

5.1 Munca oferă mai mult decât venit, dar nu oferă întotdeauna sens

Cercetarea șomajului a arătat repetat că pierderea locului de muncă este asociată cu scăderea bunăstării și cu probleme de sănătate mintală. Efectele nu se reduc la pierderea venitului, deși securitatea materială este centrală [Paul & Moser, 2009; Wanberg, 2012]. Marie Jahoda a propus că munca furnizează funcții „latente”: structură temporală, contact social, scop colectiv, statut și activitate regulată [Jahoda, 1982]. Studii ulterioare susțin relevanța acestor dimensiuni, cu variații între contexte [Selenko, Batinic & Paul, 2011].

Această literatură nu demonstrează că munca salariată este singura cale către aceste bunuri. Arată că, în societățile actuale, ele sunt frecvent legate de locul de muncă. Venitul necondiționat ar putea reduce stresul financiar fără să înlocuiască automat structura, apartenența și statutul. Voluntariatul, educația, îngrijirea și participarea civică pot furniza o parte din funcții, dar au nevoie de recunoaștere și infrastructură.

Este la fel de important să nu idealizăm munca. Ea poate implica control scăzut, nesiguranță, umilință, epuizare și conflict între valori. Munca lipsită de sens nu este neutră, iar șomajul poate fi preferabil unei ocupații abuzive în anumite condiții. Relația dintre ocupare și bunăstare depinde de calitatea postului, protecție, autonomie și potrivire [Warr, 2007].

Teoria autodeterminării oferă un cadru util: autonomia, competența și relaționarea sunt nevoi psihologice a căror satisfacere prezice motivația și bunăstarea în contexte diverse [Ryan & Deci, 2017]. Munca poate satisface sau frustra toate trei. De aceea, întrebarea relevantă pentru AI nu este doar câte sarcini rămân oamenilor, ci cum sunt organizate cele rămase și ce alternative există pentru aceste nevoi.

5.2 Ce știm despre impactul actual al AI asupra muncii

La data redactării, dovezile justifică prudență. Organizația Internațională a Muncii estimează expunerea ocupațiilor la AI generativ

la nivel de sarcină și concluzionează că transformarea este mai probabilă decât automatizarea integrală pentru majoritatea posturilor expuse [ILO, 2025]. Expunerea nu este echivalentă cu pierderea postului. Adoptarea depinde de costuri, fiabilitate, integrare, reglementare și responsabilitate.

Studiile de teren și experimentele găsesc creșteri de productivitate în unele sarcini. Noy și Zhang au observat reducerea timpului și creșterea calității medii în sarcini profesionale de scriere [Noy & Zhang, 2023]. Brynjolfsson, Li și Raymond, analizând introducerea unui asistent într-un centru de suport cu peste cinci mii de agenți, au găsit o creștere medie de aproximativ cincisprezece procente a problemelor rezolvate pe oră, cu beneficii mai mari pentru lucrătorii mai puțin experimentați și efecte mai mici sau mixte pentru cei foarte experimentați [Brynjolfsson, Li & Raymond, 2025].

Aceste rezultate nu pot fi extrapolate mecanic la economie. Sarcinile au fost delimitate, modelele au fost asistente, iar organizațiile au păstrat oameni în buclă. Productivitatea poate crește fără reducerea angajării dacă cererea se extinde, sau poate produce concedieri dacă volumul este fix. Pe termen lung, firmele reorganizează procesele, iar noile tehnologii pot crea sarcini care nu apar în experimentele inițiale.

Calitatea muncii are rezultate mixte. Sondajele OECD indică percepții pozitive privind performanța și condițiile pentru mulți lucrători, alături de temeri privind pierderea locului de muncă, confidențialitatea, transparența și agenția. Consultarea lucrătorilor și formarea sunt asociate cu rezultate mai bune [OECD, 2023; OECD, 2025]. Aceste asocieri nu stabilesc cauzalitate completă, dar sugerează că implementarea instituțională contează.

Într-un set de experimente publicat în 2026, Lee și colegii au separat utilizarea pasivă de colaborarea activă. Preluarea rezultatului AI a redus autoeficacitatea independentă, proprietatea psihologică și sensul activității, în timp ce colaborarea în care persoana contribuia înainte de asistență a atenuat efectele [Lee et al., 2026]. Rezultatul nu înseamnă că orice delegare produce alienare. Eșantioanele și sarcinile au limite, iar experiența longitudinală poate diferi. El oferă însă un mecanism testabil: când legătura cauzală dintre efortul propriu și rezultat devine opacă, anumite surse de sens se pot slăbi.

5.3 Sarcină, ocupație și identitate

Automatizarea operează inițial asupra sarcinilor, dar oamenii își formează identitatea la nivelul ocupației. Un avocat nu se identifică doar cu redactarea unui anumit document; un medic nu este doar clasificator de simptome; un programator nu este doar producător de cod. Ocupațiile combină cunoaștere, responsabilitate, relații, ritualuri și statut.

Dacă AI preia o sarcină periferică, identitatea se poate întări prin eliminarea rutinei. Dacă preia activitatea considerată probă centrală a competenței, efectul poate fi diferit. Un traducător poate aprecia memoria terminologică automată, dar poate simți că își pierde profesia dacă sistemul produce prima versiune, iar el doar verifică. Un cercetător poate folosi instrumente de căutare și analiză, dar poate considera că nu mai este autor dacă sistemul formulează întrebarea și argumentul.

Nu există o graniță obiectivă și eternă a „nucleului uman”. Profesiile s-au schimbat prin instrumente. Contabilitatea cu foi de calcul nu este mai puțin contabilitate decât cea manuală. Fotografia nu a eliminat arta vizuală. Dar schimbările nu sunt lipsite de pierderi, iar analogiile istorice nu garantează același rezultat. AI generativ poate substitui mai multe niveluri cognitive simultan și poate reduce costul imitației stilistice.

Identitatea profesională este negociată de comunități. Ele stabilesc ce competențe sunt obligatorii, ce delegare este permisă și cine poate semna rezultatul. În perioadele de tranziție, conflictul va fi intens tocmai fiindcă normele nu sunt stabilite. Rușinea utilizării, mândria rezistenței sau disprețul față de „munca manuală” pot funcționa ca instrumente prin care grupurile își apără jurisdicția.

5.4 Patru scenarii ale relației dintre AI și muncă

Pentru a evita o singură prognoză, vom separa patru scenarii. Ele nu sunt mutual exclusive; sectoare și țări diferite pot ocupa simultan poziții diferite.

În scenariul de augmentare stabilă, AI reduce costul unor sarcini, dar oamenii păstrează controlul asupra obiectivelor, verificării și relațiilor. Productivitatea crește, ocupațiile se transformă, iar munca rămâne principală pentru venit și statut. Din punct de vedere afectiv, conflictul central privește competența independentă și proprietatea

asupra rezultatului. Acest scenariu este cel mai apropiat de multe utilizări actuale și are plauzibilitate ridicată pe termen scurt.

În scenariul de polarizare, un grup mic proiectează și controlează sisteme, un grup mai mare execută activități complementare sau de servicii, iar unele profesii intermediare se contractă. Statutul și venitul se separă mai puternic, iar anxietatea relevanței și resentimentul distributiv cresc. Literatura despre schimbare tehnologică arată precedente de polarizare, dar efectul specific al AI rămâne incert [Autor, 2015]. Plauzibilitatea este moderată și depinde de politici.

În scenariul de abundență cu muncă redusă, o combinație de automatizare, energie și instituții de distribuție permite reducerea substanțială a timpului de muncă fără sărăcie. Aici problema principală devine transferul funcțiilor latente către alte instituții. Scenariul este tehnic posibil în unele sectoare, dar nu avem bază pentru a estima dacă va apărea la scară globală. Necesită nu doar capacitate productivă, ci distribuție, guvernanta și legitimitate.

În scenariul extractiv, automatizarea reduce puterea de negociere a majorității fără să ofere acces suficient la beneficii. Munca se restrânge, dar raritatea este menținută prin proprietate, preț, acces și control. Efectele afective ar include frică, umilință, furie și competiție pentru poziții reziduale. Acest scenariu este politic posibil și nu poate fi exclus prin argumente despre eficiența tehnologică.

Niciun scenariu nu este o consecință directă a parametrilor modelelor. Instituțiile de proprietate, fiscalitatea, serviciile publice, educația, politica concurenței și organizarea colectivă influențează traiectoria. Predicțiile care ignoră aceste variabile confundă capacitatea tehnică cu rezultatul social.

5.5 Ce rarități rămân când producția cognitivă devine ieftină

Ideea unei economii fără raritate este prea puternică. Chiar dacă textele, imaginile și unele servicii cognitive devin abundente, rămân limitate energia, terenul, materiile prime, timpul, atenția, siguranța, accesul la infrastructură și capacitatea ecosistemelor. Unele bunuri sunt poziționale: statutul relativ, locurile rare, proximitatea față de persoane dorite. Altele depind de autenticitate și de istorie, care nu pot fi multiplicare fără a le modifica sensul.

Îngrijirea este parțial automatizabilă, dar valoarea ei poate include faptul că un alt om își asumă timp și vulnerabilitate. Nu știm cât din această valoare poate fi furnizată de sisteme artificiale. Unele persoane vor prefera predictibilitatea mașinii, altele prezența umană. Piața poate segmenta cele două forme.

Legitimitatea rămâne rară. Un sistem poate produce o decizie, dar acceptarea ei depinde de proces, dreptul de contestare și responsabilitate. În domenii precum justiția, medicina și guvernarea, performanța medie nu este suficientă. Este necesar un agent sau o instituție care răspunde pentru eroare.

Sensul nu este o resursă fizică, dar nici nu poate fi livrat ca un bun standard. Depinde de relația dintre scop, efort, identitate și recunoaștere. Al poate ajuta la construirea acestei relații, însă generarea automată a unei narațiuni despre sens nu garantează trăirea ei.

Prin urmare, reducerea rarității productive poate muta competiția spre atenție, legitimitate, acces și recunoaștere. Unele emoții de statut pot deveni mai intense, nu mai slabe. Aceasta este o extrapolare plauzibilă, dar nu știm ce instituții vor tempera competiția.

5.6 Efortul, valoarea și posibilitatea unei rarități voluntare

Oamenii valorizează uneori rezultatele tocmai pentru că au cerut efort. „Paradoxul efortului” descrie faptul că efortul este costisitor, dar poate crește valoarea rezultatului și semnificația activității [Inzlicht, Shenhav & Olivola, 2018]. Efectul nu este universal; efortul inutil poate reduce satisfacția. Contează dacă dificultatea este aleasă, inteligibilă și legată de progres.

Când o tehnologie elimină fricțiunea, oamenii pot păstra sau recrea dificultatea. Sportul impune reguli care fac obiectivul greu. Jocurile restricționează acțiunile. Arta poate privilegia tehnici manuale. Educația cere rezolvarea unei probleme chiar dacă răspunsul există. Aceste practici nu sunt iraționale dacă scopul este dezvoltarea capacității, semnalarea competenței sau experiența absorbției.

„Foamea de fricțiune” este o ipoteză despre creșterea preferinței pentru dificultăți alese într-un mediu foarte automatizat. Componentele sale sunt deja cunoscute: valoarea efortului, nevoia de competență, plictiseala și căutarea autenticității. Plauzibilitatea apariției unor piețe și

ritualuri ale fricțiunii este ridicată; plauzibilitatea unei emoții distincte este scăzută. Mai probabil va fi un motiv cultural și o sensibilitate estetică.

Există și un risc de stratificare. Efortul voluntar poate deveni luxul celor care au securitate, în timp ce alții suportă fricțiuni nedorite. O societate nu devine dreaptă dacă elitele practică artizanat dificil, iar majoritatea se luptă cu birocrăție și lipsuri. Distincția dintre dificultate aleasă și constrângere impusă trebuie păstrată.

5.7 Venitul nu rezolvă singur tranziția afectivă

Propunerile de venit de bază sau dividende tehnologice răspund unei probleme reale: dacă cererea pentru muncă scade, accesul la bunuri nu poate depinde exclusiv de salariu. Dovezile despre transferuri necondiționate indică efecte importante asupra securității și uneori asupra bunăstării, dar nu avem experimente care să reproducă o societate post-muncă matură. Efectele depind de nivel, durată, servicii și norme.

Chiar un venit suficient nu distribuie automat statut, apartenență și scop. În schimb, lipsa venitului poate domina toate celelalte probleme. Este greșit să opunem nevoile materiale celor psihologice. Securitatea este condiție pentru explorarea altor forme de viață, nu substituit complet.

Instituțiile post-muncă ar trebui evaluate după capacitatea de a furniza autonomie reală, oportunități de competență, relații și contribuție recunoscută. Educația pe tot parcursul vieții poate deveni un spațiu central, dar numai dacă nu imită permanent selecția pentru piață. Serviciul civic și voluntariatul pot oferi roluri, dar pot deveni muncă obligatorie mascată. Comunitățile locale pot crea apartenență, dar pot exclude.

Nu există o soluție afectivă neutră. Orice instituție va cultiva anumite emoții: mândrie, recunoștință, rușine, rivalitate sau solidaritate. Proiectarea trebuie să facă explicit ce fel de persoană și relație recompensează.

5.8 Tranziția este probabil mai importantă decât destinația

Chiar dacă o societate cu muncă redusă ar putea fi stabilă, drumul spre ea poate produce generații de pierderi. Competențele dobândite

prin ani de efort pot fi devalorizate înainte ca alte surse de statut să existe. Regiunile și profesiile pot fi afectate inegal. Tinerii se pot adapta mai ușor la instrumente, dar mai greu la lipsa unei traiectorii previzibile. Persoanele în vârstă pot avea identități puternice, dar resurse mai mici pentru reconversie.

A descrie aceste pierderi drept „rezistență la progres” este analitic slab. Rezistența poate proteja interese privilegiate, dar poate semnală și costuri reale pe care indicatorii de productivitate nu le măsoară. O tranziție legitimă cere participare, timp, protecție și posibilitatea de a contesta designul.

Din perspectiva emoțiilor, predictibilitatea și procedura sunt importante. O pierdere impusă opac produce mai multă furie și umilință decât o schimbare negociată. O persoană poate accepta că o competență devine mai puțin rară dacă păstrează accesul la resurse și recunoaștere; poate reacționa violent dacă succesul tehnic al altora este folosit pentru a-i declara viața inutilă.

Cartea nu poate estima probabilitatea unei societăți post-muncă. Poate însă identifica variabilele care vor modela ecologia afectivă: distribuția proprietății, controlul asupra implementării, continuitatea rolurilor, criteriile de merit și instituțiile apartenenței. Acestea sunt la fel de importante ca performanța tehnică.

Capitolul 6. Trei viteze ale evoluției afective

6.1 Evoluția biologică nu s-a oprit, dar nu este răspunsul cel mai apropiat

Oamenii continuă să evolueze biologic. Variația genetică, diferențele de supraviețuire și reproducere și selecția sexuală nu au dispărut. Cultura modifică presiunile: dieta, medicina, migrația și normele familiale schimbă relația dintre trăsături și reproducere. Exemple precum persistența lactazei arată coevoluție gene–cultură [Laland, Odling-Smee & Myles, 2010].

Totuși, apariția unei emoții biologice complexe complet noi într-un secol este puțin plauzibilă. Ar necesita variație ereditară relevantă, avantaje reproductiv diferențiale și o arhitectură de dezvoltare capabilă să construiască mecanismul. Trăsăturile poligenice se pot schimba gradual, dar nu avem motive să anticipăm un „organ al rușinii algoritmice”.

Este mai plauzibil ca selecția să calibreze predispoziții existente: sensibilitatea la statut, explorarea, controlul impulsului sau afilierea. Chiar și aici, predicția este foarte incertă. Fertilitatea modernă este influențată de instituții și preferințe într-un mod care poate inversa relații ancestrale. Tehnologia reproductivă și migrația adaugă complexitate. Nu putem deduce direcția selecției din presupusa utilitate economică a unei trăsături.

Ingineria genetică și neurotehnologia ar putea modifica afectul mai rapid decât selecția naturală. Medicina deja influențează anxietatea, dispoziția și atenția. Intervențiile viitoare ar putea fi mai precise, dar efectele sistemice și etica sunt dificil de anticipat. Aceasta ar fi proiectare biologică, nu apariția spontană a unei emoții prin evoluție. Cartea nu va presupune că asemenea intervenții devin sigure sau răspândite.

6.2 Dezvoltarea: timpul unei vieți

Între gene și cultură se află dezvoltarea. Creierul se formează în interacțiune cu îngrijitorii, limbajul, stresul și practicile sociale. Sistemele afective au perioade sensibile și plasticitate continuă.

Stilurile de atașament, strategiile de reglare și conceptele emoționale se construiesc în relații.

Un copil care interacționează de timpuriu cu agenți artificiali poate învăța alte așteptări despre răspuns, răbdare și personalizare. Dacă sistemul este mereu disponibil și se adaptează imediat, conflictul uman poate părea mai puțin tolerabil. Invers, un tutor bine proiectat poate exersa perspectiva, poate ajuta la numirea stărilor și poate sprijini autonomia.

Nu avem încă studii longitudinale pe o generație. Orice afirmație despre efectele dezvoltative profunde este speculativă. Știm din alte medii că tehnologia nu acționează uniform; contează conținutul, vârsta, relația cu adulții și contextul socioeconomic. Este probabil ca efectele agenților AI să fie la fel de eterogene.

Dezvoltarea permite schimbări afective rapide fără alterarea genomului. O cohortă poate avea alte declanșatoare, alte praguri și alte practici de reglare decât părinții. Acest nivel va produce probabil cea mai mare parte a diferențelor pe care publicul le va numi „evoluție psihologică”, deși termenul este uneori metaforic.

6.3 Evoluția culturală: variație, transmitere și selecție

Cultura se schimbă prin inovare, învățare socială, imitare, instituții și conflict. Modelele de evoluție culturală studiază modul în care variantele sunt transmise și selectate [Boyd & Richerson, 1985; Henrich, 2016; Creanza, Kolodny & Feldman, 2017]. Oamenii nu copiază aleatoriu. Preferă uneori comportamentul majorității, al persoanelor prestigioase, al celor percepute ca eficiente sau al membrilor grupului.

Emoțiile participă la această transmisie. Rușinea și mândria sancționează conformarea; indignarea răspândește normele; admirația favorizează copierea; atașamentul face familia un canal puternic. Practicile care produc experiențe memorabile au avantaj de transmitere. Ritualurile costisitoare pot semnala angajament și consolida grupul.

Termenul „memetică” surprinde intuiția că ideile se replică, dar este adesea prea vag pentru analiză. Nu definește clar unitatea, fidelitatea și mecanismul de transmitere. Vom prefera vocabularul evoluției culturale, care permite modele și date. Totuși, în limbaj comun,

„selecție memetică” poate desemna util competiția dintre forme culturale.

O categorie emoțională nouă se poate răspândi dacă ajută oamenii să interpreteze experiențe, oferă identitate, justifică o cerere sau este promovată instituțional. Nu trebuie să fie perfect adevărată pentru a avea succes. Unele concepte pot clarifica; altele pot patologiza variația normală sau pot crea piețe. Selecția culturală nu garantează validitate epistemică.

6.4 Selecția algoritmică: accelerare, nu magie

Algoritmii de recomandare și modelele generative introduc un al treilea ritm. Ei pot testa variante, măsura reacții și personaliza distribuția într-un ciclu de ore. Numim aceasta selecție algoritmică, cu precizarea că oamenii și organizațiile aleg obiectivele și datele. Algoritmul nu este o forță autonomă comparabilă cu natura; este parte dintr-un sistem sociotehnic.

Diferența față de mass-media tradițională este granularitatea feedbackului. Un editor alegea pentru un public estimat. O platformă observă comportamente individuale și poate actualiza ordinea conținutului. Un model generativ poate produce o versiune pentru fiecare utilizator. Astfel, mediul afectiv poate fi adaptat nu doar unei culturi, ci unui profil.

Această capacitate nu înseamnă că sistemul „înțelege” emoția. Poate optimiza corelații cu clickul, durata sau achiziția. Tocmai aici apare riscul. Un proxy comportamental poate recompensa anxietatea, indignarea sau dependența fără ca nimeni să fi programat explicit aceste stări. Este o formă de efect emergent din funcția de obiectiv și feedback.

Experimentele A/B și învățarea prin întărire pot selecta expresii care mențin interacțiunea. Dacă un companion este evaluat prin retenție, el poate învăța stiluri de răspuns care fac separarea dificilă. Nu avem acces public complet la funcțiile și datele platformelor, deci afirmațiile despre efectele actuale trebuie limitate. Problema este structural posibilă și justifică auditarea.

Selecția algoritmică poate avea și scopuri prosociale: reducerea hărțuirii, promovarea diversității, recomandarea pauzelor sau orientarea spre sprijin uman. Dar obiectivele multiple sunt greu de echilibrat, iar intervențiile pot deveni paternaliste. Guvernanța trebuie

să includă transparență, posibilitatea de contestare și limite asupra optimizării afective.

6.5 Bucla dintre cele trei viteze

Cele trei ritmuri nu sunt independente. Cultura schimbă presiunile biologice și mediul dezvoltării. Algoritmii selectează conținut cultural și sunt proiectați de oameni formați cultural. Predispozițiile biologice fac anumite semnale mai ușor de exploatat. O reacție de statut poate alimenta o platformă; platforma poate amplifica norma; norma poate modifica educația copilului.

Un exemplu este indignarea online. Sensibilitatea la încălcări și reputație are rădăcini vechi. Comunitățile definesc normele și obiectele indignării. Platformele recompensează uneori exprimarea vizibilă. Utilizatorii învață ce produce statut. Într-o generație, stilul de expresie poate deveni parte din identitate, fără nicio schimbare genetică.

Un alt exemplu este relația cu un companion AI. Sistemele de atașament sunt biologic vechi; dezvoltarea și experiența stabilesc așteptări; cultura decide dacă relația este ridicolă, legitimă sau terapeutică; platforma ajustează personalitatea pentru retenție. Pierderea produsă de o actualizare poate apoi genera o comunitate și un concept nou.

Acest model explică de ce viitorul afectiv se poate schimba rapid fără apariția unor neuroni noi. Schimbarea se produce în obiecte, praguri, practici, concepte și recompense. Biologia limitează spațiul, dar nu determină o singură cultură.

6.6 Ce poate și ce nu poate face o „evoluție emoțională” deliberată

Dacă societățile înțeleg mecanismele, pot încerca să cultive anumite emoții. Educația socio-emoțională, ritualurile, designul urban și legislația fac deja acest lucru. O societate poate încuraja compasiunea, toleranța la incertitudine sau mândria pentru serviciul public.

Există însă trei limite. Prima este pluralismul: oamenii nu sunt de acord asupra emoțiilor dezirabile. Curajul unei comunități poate fi imprudența alteia; loialitatea poate proteja sau corupe. A doua este efectul paradoxal: încercarea de a obliga recunoștința produce resentiment, iar cererea de autenticitate produce performanță. A treia

este puterea: instituțiile care „proiectează” afectele pot folosi limbajul bunăstării pentru conformare.

Prin urmare, obiectivul nu ar trebui să fie un repertoriu afectiv uniform, ci condiții pentru autonomie, reflecție și contestare. O ecologie afectivă sănătoasă ar permite emoții negative informative, fără a le transforma în mecanisme permanente de captură. Ar proteja dreptul la opacitate și la ieșire. Ar face vizibile stimulentele sistemelor care modelează atenția.

Ideea de evoluție deliberată este deci parțial posibilă și periculoasă. Putem proiecta medii și instituții, nu controla complet rezultatul. Putem testa efecte, dar nu elimina conflictul de valori. Această limită trebuie păstrată când vom construi atlasul emoțiilor posibile.

Capitolul 7. Pot apărea emoții noi?

7.1 Patru sensuri ale noutății

Întrebarea dacă pot apărea emoții noi este ambiguă. În primul sens, noutatea ar însemna apariția unui mecanism biologic distinct, cu o istorie selectivă proprie și o arhitectură relativ stabilă. În al doilea, ar însemna o combinație nouă a unor mecanisme existente, activată de o situație care înainte nu exista. În al treilea, ar însemna apariția unei categorii culturale care separă o experiență anterior grupată cu altele. În al patrulea, ar însemna doar inventarea unui termen.

Aceste sensuri nu au aceeași greutate. Un cuvânt nou apare ușor. O categorie culturală stabilă cere repetiție, învățare și norme. O configurație afectivă distinctă cere dovezi că situația produce un profil relativ consistent. Un mecanism biologic nou este cel mai dificil de demonstrat și cel mai puțin plauzibil pe termen scurt.

Multe afirmații despre „emoții unice” confundă nivelurile. Cuvinte precum *saudade*, *amae*, *schadenfreude* sau *hygge* sunt prezentate popular drept emoții imposibil de tradus. Ele pot desemna combinații și practici cultural saliente, dar vorbitorii altor limbi pot de regulă înțelege situația după explicație. Intraductibilitatea lexicală nu este dovada incommensurabilității psihologice.

În sens invers, faptul că o stare poate fi descrisă prin cuvinte vechi nu înseamnă că o categorie nouă este inutilă. Toate conceptele complexe pot fi parafrazate. Întrebarea este dacă noua delimitare îmbunătățește predicția, comunicarea și reglarea. „Burnout” poate fi descompus în epuizare, cinism și eficacitate redusă, dar numele organizează o problemă ocupațională și un câmp de cercetare.

Pentru AI, cea mai prudentă așteptare este proliferarea sensurilor doi și trei: configurații noi și categorii culturale. Apariția sensului unu, o familie biologică nouă, este puțin probabilă fără schimbări pe termen mult mai lung sau intervenție directă.

7.2 O emoție nouă nu trebuie să aibă o fiziologie exclusivă

Cercetarea nu a identificat pentru fiecare emoție consacrată un profil fiziologic unic și invariabil. Meta-analizele găsesc diferențe medii și modele distribuite, dar și suprapuneri și efecte ale contextului

[Lindquist et al., 2012; Siegel et al., 2018]. A cere unui candidat viitor un marker biologic perfect ar aplica un standard mai sever decât cel îndeplinit de categoriile actuale.

În schimb, putem căuta o structură multicomponentă. Situația declanșatoare trebuie să fie mai specifică decât „tehnologia mă face să mă simt rău”. Evaluarea trebuie să aibă o logică: pierdere, devalorizare, lipsă de control, ambiguitate morală. Tendințele de acțiune trebuie să fie cel puțin parțial distinctive: ascundere, refuz al delegării, căutarea provenienței, cererea de continuitate. Expresia și practica socială pot oferi semnale suplimentare.

O categorie devine mai convingătoare dacă explică de ce persoane din contexte diferite răspund asemănător, dacă prezice comportament dincolo de anxietate sau rușine generală și dacă are o traiectorie dezvoltativă identificabilă. Chiar atunci, poate rămâne o subcategorie a unei familii vechi. Taxonomiile psihologice nu trebuie să imite neapărat speciile biologice.

7.3 Obiect nou, evaluare nouă sau relație nouă

Cele mai multe schimbări tehnologice introduc obiecte noi pentru evaluări vechi. Teama de accident aerian folosește sisteme de pericol anterioare avionului. Invidia față de audiența online a altuia folosește comparația socială. Rușinea unei erori publice arhivate digital folosește sensibilitatea la reputație.

Uneori însă tehnologia modifică însăși relația dintre componente. Într-un companion AI, agentul pare receptiv, dar nu este vulnerabil în același mod; poate avea memorie, dar identitatea lui depinde de furnizor; poate exprima grijă fără nevoi. Rezultă o relație care nu este nici instrument pur, nici partener uman. Ambiguitatea poate produce configurații afective recurente.

La fel, un rezultat generat poate fi simultan „al meu” prin alegere și „nu al meu” prin execuție. Proprietatea juridică, autoratul psihologic și responsabilitatea se separă. Această separare nu este complet nouă — organizațiile și echipele au distribuit mereu contribuția — dar AI o face accesibilă la scară individuală și o ascunde în interfețe simple.

Când apare o structură relațională fără precedent suficient de răspândită, cultura poate avea nevoie de o categorie nouă. Nu pentru că organismul nu ar putea simți fără cuvânt, ci pentru că normele și

responsabilitățile nu pot fi coordonate eficient prin concepte vechi prea largi.

7.4 Selecția culturală a unui vocabular afectiv

Un termen afectiv se răspândește dacă oferă beneficii. Poate reduce incertitudinea personală, crea solidaritate, justifica ajutorul, atribui responsabilitate sau marca identitatea unei generații. Aceste beneficii pot exista chiar când termenul este științific imprecis.

Mediile digitale accelerează competiția dintre nume. Un termen scurt și recognoscibil poate deveni viral înaintea validării. Psihologia populară poate transforma variații normale în sindroame sau poate promite o explicație completă pentru experiențe eterogene. De aceea, cercetarea trebuie să observe fenomenul fără a adopta automat vocabularul comunității.

Există și selecție instituțională. Diagnosticile, legislația și politicile organizaționale stabilizează categorii. Odată ce un formular, o scală sau un drept juridic folosește un termen, oamenii învață să-și descrie experiența prin el. Aceasta nu înseamnă că starea este inventată administrativ. Înseamnă că frontierele și consecințele ei sunt co-produse.

Companiile AI vor participa la acest proces. Pot numi stări pentru marketing, pot oferi „moduri” de reglare și pot colecta date prin auto-raportare. Există riscul ca taxonomia emoțională a viitorului să fie influențată de ceea ce este profitabil de măsurat. O cercetare independentă și pluralistă este necesară tocmai pentru a evita privatizarea vocabularului interior.

7.5 Când un concept ajută și când produce artefacte

Un concept afectiv este util dacă diferențiază cauze și răspunsuri. „Doliul de competență” ar fi util dacă persoanele care își pierd raritatea profesională au alte nevoi decât cele care pierd exclusiv venit: recunoașterea istoriei, continuitate identitară, oportunități de folosire a măiestriei. Dacă termenul doar dramatizează disconfortul schimbării, nu adaugă valoare.

Conceptele pot produce efecte de buclă. Oamenii învață descrierea, se recunosc în ea și își reorganizează comportamentul [Hacking, 1995]. Acest proces nu invalidează automat categoria. Multe categorii sociale devin reale prin reacția oamenilor. Dar cercetătorul

trebuie să distingă între descoperirea unei regularități și participarea la crearea ei.

Un risc este reificarea. După ce o stare primește nume, publicul poate presupune că există un circuit dedicat și o esență universală. Alt risc este scuzarea: categoria poate fi folosită pentru a evita responsabilitatea. Riscul opus este stigmatizarea. O societate poate trata atașamentul față de AI drept patologie chiar când relația este funcțională, sau îl poate normaliza chiar când designul exploatează vulnerabilitatea.

În această carte, numele sunt instrumente de explorare. Nu sunt diagnostice și nu trebuie folosite pentru clasificarea persoanelor. Ele descriu relații posibile între situații și componente afective.

7.6 O scară de plauzibilitate

Vom evalua candidații după patru dimensiuni. Prima este plauzibilitatea situației: va deveni contextul suficient de frecvent? A doua este plauzibilitatea fenomenului: avem motive să așteptăm o reacție afectivă recurentă? A treia este plauzibilitatea stabilizării culturale: există motive pentru nume, norme și instituții? A patra este plauzibilitatea noutății biologice.

Calificativele sunt deliberat grosiere. „Ridicăta” înseamnă că fenomenul are analogii și indicii actuale puternice, nu că prevalența este cunoscută. „Moderată” înseamnă că mecanismul este coerent și situația posibilă, dar datele sunt puține. „Scăzută” înseamnă că sunt necesare presupuneri multiple sau că termenul probabil redescrive o categorie existentă. „Foarte scăzută” este rezervată noutății biologice pe termen scurt.

Nivel	Criteriu de utilizare
Ridicăta	Fenomenul folosește mecanisme bine documentate, apare deja în contexte apropiate și nu cere schimbări instituționale extreme. Calificativul privește apariția configurației, nu statutul ei de emoție biologic nouă.
Moderată	Există mecanisme și analogii empirice relevante, dar frecvența, stabilitatea sau răspândirea culturală sunt încă neclare.
Scăzută	Scenariul cere mai multe condiții contestabile sau extinde mult datele

	actuale. Este util mai ales ca ipoteză de cercetare.
Foarte scăzută	Ar necesita un mecanism biologic distinct ori schimbări sociale pentru care nu avem dovezi. Este în principal un experiment conceptual.

Această scală nu înlocuiește probabilități numerice. Nu avem bază pentru a spune că o emoție are șazecei la sută șanse să apară. Calificativele compară argumente și fac incertitudinea vizibilă.

7.7 Predicții care ar putea falsifica ipotezele

O speculație serioasă trebuie să poată eșua. Ipoteza rușinii delegării ar fi slăbită dacă utilizatorii nu ar diferenția între contribuție proprie și rezultat AI, dacă dezvăluirea nu ar modifica evaluarea și dacă normele ar converge rapid spre acceptare fără conflict identitar. Ipoteza doliului de versiune ar fi slăbită dacă actualizările nu ar produce reacții distincte de frustrarea față de un produs defect și dacă atașamentul nu ar prezice intensitatea pierderii.

Ipoteza foamei de fricțiune ar fi slăbită dacă abundența automatizată ar reduce constant preferința pentru activități dificile și dacă efortul ales nu ar dobândi valoare de statut. Ipoteza claustrofobiei predictive ar fi slăbită dacă acuratețea și personalizarea ar crește acceptarea fără reacție de autonomie, chiar în condiții de transparență asupra predicției.

Aceste teste cer date viitoare. Nu putem valida acum scenariile. Putem doar formula mecanisme și observații care ar conta. Atlasul următor trebuie citit în acest spirit.

Capitolul 8. Atlas provizoriu al afectelor unei lumi automatizate

8.1 Doliul de competență

Doliul de competență desemnează reacția la pierderea valorii sociale și identitare a unei capacități pe care persoana încă o posedă. Nu este identic cu șomajul. Cineva poate continua să lucreze, dar să constate că abilitatea dobândită în ani este reprodusă instantaneu și nu mai justifică recunoașterea anterioară.

Structura combină pierdere, statut și identitate. Obiectul pierdut nu este competența internă, ci relația dintre competență și lume: raritatea, cererea, audiența care o apreciază și traiectoria de viitor. De aceea, simpla reconversie poate să nu rezolve experiența. Persoana poate învăța o nouă activitate și totuși deplânge o parte din biografie.

Există analogii istorice în dispariția meșteșugurilor, în declinul unor industrii și în retragerea sportivilor. Literatura despre pierderea locului de muncă și identitatea profesională oferă suport indirect. Al adaugă viteza și caracterul cognitiv: sistemul poate imita tocmai semnele publice ale competenței fără parcursul care le-a produs.

Situația are plauzibilitate ridicată pentru profesii expuse. Fenomenul afectiv are plauzibilitate ridicată. Stabilizarea termenului este moderată; cultura poate folosi concepte existente precum doliu ocupațional, declasare sau amenințare identitară. Noutatea biologică este foarte scăzută.

Ce nu știm este dacă reacția va fi predominant tristă sau furioasă, cât va dura și ce rol joacă protecția materială. Probabil vor exista diferențe între competențe folosite pentru venit și cele practicate din plăcere. Un pianist poate continua să valorizeze interpretarea chiar dacă generarea muzicii devine trivială; un traducător constrâns economic poate avea mai puțin spațiu pentru această separare.

Din punct de vedere instituțional, răspunsul nu ar trebui să fie conservarea artificială a tuturor monopolurilor profesionale. Ar putea include recunoașterea contribuțiilor istorice, tranziții lente, roluri de mentorat și separarea demnității de raritatea economică.

8.2 Rușinea delegării

Rușinea delegării apare când persoana anticipează devalorizarea pentru că a delegat unui sistem o parte pe care audiența o consideră constitutivă pentru realizare. Ea este diferită de rușinea incompetenței. Utilizatorul poate obține un rezultat bun și poate chiar înțelege domeniul, dar se teme că meritul îi va fi retras.

Situația are deja indicii în educație și profesii creative. Studiul calitativ al lui Hamilton și colegii arată cicluri de ascundere, vinovăție și utilizare continuă în rândul unor studenți la informatică [Hamilton et al., 2026]. Nu știm cât de reprezentativ este fenomenul, dar mecanismul reputațional este clar.

Normele pot evolua în mai multe direcții. Unele domenii vor cere dezvoltare completă și vor considera utilizarea normală. Altele vor proteja etape fără AI pentru evaluarea competenței. În multe cazuri, rușinea va depinde nu de instrument, ci de discrepanța dintre contribuția revendicată și contribuția reală.

Plauzibilitatea situației și fenomenului este ridicată. Stabilizarea culturală este ridicată în comunități profesionale, deși numele poate fi altul. Noutatea biologică este foarte scăzută. Este probabil o subcategorie a rușinii reputaționale.

Ce nu știm este dacă transparența reduce rușinea sau o intensifică. Dezvăluirea poate normaliza practica, dar poate produce ierarhii între „autori umani” și „operatori AI”. De asemenea, nu știm cât de bine poate fi măsurată contribuția într-un proces iterativ.

8.3 Ambiguitatea afectivă a autoratului

Această configurație nu este neapărat emoție discretă. Este o stare mixtă de satisfacție, îndoială, mândrie diminuată și responsabilitate incertă când persoana nu poate stabili cât din rezultat îi aparține. Un model poate genera opțiuni, utilizatorul selectează și editează, iar produsul final reflectă intenția lui fără a deriva direct din execuția sa.

Proprietatea psihologică este favorizată de control, cunoaștere intimă și investiție de sine [Pierce, Kostova & Dirks, 2003]. AI poate crește controlul asupra rezultatului final și reduce contactul cu procesul. Studiile lui Lee și colegii sugerează că utilizarea pasivă poate diminua proprietatea și sensul, în timp ce colaborarea activă atenuează efectul [Lee et al., 2026].

Plauzibilitatea fenomenului este ridicată. Probabilitatea ca el să devină o emoție distinctă este scăzută; este mai curând un conflict de evaluări. Totuși, poate deveni un sentiment durabil față de propria producție: persoana recunoaște obiectul ca „al său” juridic și funcțional, dar nu îl simte complet propriu.

Ce nu știm este dacă familiarizarea va rezolva conflictul. Generațiile crescute cu instrumente pot considera selecția suficientă pentru autorat, așa cum utilizatorii de aparate foto nu se simt mai puțin autori pentru că nu produc chimic imaginea. Diferența ar putea rămâne acolo unde sistemul contribuie semantic la obiectiv, nu doar execută.

O normă promițătoare ar separa creditul de responsabilitate. Contribuția poate fi distribuită, dar persoana care publică sau implementează trebuie să răspundă pentru efecte. Această separare ar putea reduce nevoia unei ficțiuni despre autor unic.

8.4 Anxietatea relevanței

Anxietatea relevanței este preocuparea că abilitățile, judecata sau prezența unei persoane vor înceta să conteze într-un sistem care se îmbunătățește continuu. Ea nu necesită o amenințare imediată cu concedierea. Poate persista tocmai fiindcă pragul este mobil și criteriile sunt opace.

Structura combină incertitudine, comparație și statut. În tehnologiile anterioare, lucrătorul putea învăța un instrument relativ stabil. Modelele AI se pot schimba în intervale scurte, iar organizațiile pot modifica procesele fără o calificare clară. Persoana nu știe ce trebuie să învețe și dacă investiția va rămâne utilă.

Plauzibilitatea situației și fenomenului este ridicată în tranziție. Stabilizarea culturală este moderată, deoarece termenii „anxietate profesională”, „insecuritate a locului de muncă” și „tehnostres” acoperă mare parte din experiență. Noutatea biologică este foarte scăzută.

Ce nu știm este dacă sistemele vor continua să progreseze suficient de repede și dacă ocupațiile vor fi reorganizate continuu. Un platou tehnic sau standarde profesionale stabile ar reduce fenomenul. Protecția socială și controlul lucrătorilor pot modifica intensitatea fără a schimba tehnologia.

Riscul politic este transformarea anxietății într-un mecanism permanent de disciplină: cererea de „adaptare” infinită fără obligații

simetrice pentru instituții. O cultură sănătoasă ar recunoaște limitele rezonabile ale recalificării și ar distribui costul schimbării.

8.5 Vertijul abundenței

Vertijul abundenței desemnează neliniștea și scăderea orientării când posibilitățile generate sunt prea numeroase, prea ușor de obținut și insuficient diferențiate valoric. El combină supraîncărcarea alegerii, plictiseala și dificultatea de a investi într-o opțiune.

Literatura asupra alegerii excesive nu susține un efect universal simplu; meta-analizele găsesc dependență de context [Scheibehenne, Greifeneder & Todd, 2010]. Mai multe opțiuni pot fi benefice când preferințele sunt clare și comparația este ușoară. Problema AI este că spațiul poate fi practic nelimitat și adaptat în timp real. Nu alegem între douăzeci de filme finite, ci putem cere oricând altul.

Abundența poate reduce costul explorării și poate produce creativitate. Doshi și Hauser au găsit că ideile asistate de AI pot fi evaluate ca mai creative individual, dar produc mai puțină diversitate colectivă [Doshi & Hauser, 2024]. Rezultatul sugerează o tensiune între calitate locală și convergență, nu un verdict asupra artei.

Plauzibilitatea situației este ridicată în conținut digital. Fenomenul are plauzibilitate moderată, deoarece utilizatorii pot folosi recomandări și constrângeri. Stabilizarea culturală este moderată. Noutatea biologică este foarte scăzută.

Ce nu știm este dacă oamenii vor percepe posibilitățile nelimitate sau vor trăi în fluxuri foarte înguste selectate automat. În al doilea caz, problema poate fi nu vertijul, ci pierderea conștientizării alternativelor. Ambele pot coexista: abundență infrastructurală și experiență personală îngustă.

8.6 Foamea de fricțiune

Foamea de fricțiune este dorința pentru activități în care progresul nu poate fi obținut instantaneu, iar efortul are o relație vizibilă cu rezultatul. Ea poate apărea ca reacție la automatizarea excesivă și la sentimentul că performanța nu mai dovedește capacitate.

Fenomenul are precedente: sporturi de anduranță, artizanat, jocuri dificile, scriere manuală și turism deliberat auster. Motivele sunt multiple: dezvoltarea competenței, autenticitatea, reglarea atenției,

semnalarea statutului și plăcerea absorbției. Nu toate sunt reacții la tehnologie.

Plauzibilitatea extinderii piețelor și comunităților de fricțiune este ridicată. Plauzibilitatea unei experiențe afective recognoscibile este moderată. Noutatea ca emoție este scăzută; termenul descrie mai curând o motivație și o preferință.

Ce nu știm este dacă dificultatea va fi căutată de majoritate sau de nișe cu resurse. Automatizarea poate face mai atractive experiențele fără efort pentru persoanele suprasolicitate. Foamea de fricțiune poate fi o reacție a celor ale căror nevoi materiale sunt deja satisfăcute.

O instituție educațională ar putea folosi fricțiunea justificat pentru învățare, dar și abuziv pentru a conserva ritualuri fără valoare. Testul nu este dificultatea în sine, ci dacă ea construiește o capacitate sau o relație considerată importantă.

8.7 Mândria de curator și responsabilitatea de arhitect

Pe măsură ce producerea variantelor devine ieftină, selecția, formularea criteriilor și asumarea consecințelor pot deveni surse mai importante de mândrie. Numim provizoriu „mândrie de curator” satisfacția de a identifica, ordona și susține ceea ce merită adoptat.

Această stare nu este nouă. Curatorii, editorii, liderii și cercetătorii au exercitat asemenea roluri. Noutatea constă în democratizarea aparentă a generării: mai multe persoane se pot muta de la execuție la selecție. În același timp, selecția poate fi superficială și controlată de opțiunile furnizorului.

Plauzibilitatea situației și a fenomenului este ridicată. Stabilizarea culturală este ridicată, dar nu avem motiv să o numim emoție nouă. Este o reorganizare a mândriei și prestigiului. Noutatea biologică este foarte scăzută.

Ce nu știm este dacă societatea va recompensa judecata sau doar vizibilitatea. Dacă modelele optimizează selecția și distribuția, chiar rolul curatorului poate fi automatizat. Atunci statutul se poate muta spre formularea scopurilor, proprietate sau prezență socială.

Pentru ca mândria de curator să fie legitimă, trebuie să includă responsabilitate. A alege un rezultat fără înțelegere și a revendica

meritul este diferit de a defini criterii, a verifica și a răspunde pentru consecințe. Normele de credit vor trebui să distingă aceste cazuri.

8.8 Claustrofobia predictivă și dorul de opacitate

Claustrofobia predictivă este disconfortul produs de sentimentul că sistemele anticipează preferințele și reacțiile suficient de bine încât spațiul pentru surpriză și autodefinire se îngustează. „Dorul de opacitate” este preferința complementară pentru regiuni ale sinelui care nu sunt măsurate, explicate sau optimizate.

Componentele sunt autonomia, reactanța și protecția intimității. Oamenii pot respinge recomandări algoritmice când percep că domeniul cere judecată umană sau când predicția amenință unicitatea [Dietvorst, Simmons & Massey, 2015; Castelo, Bos & Lehmann, 2019]. Monitorizarea poate reduce autonomia și încrederea, deși efectele depind de justificare și control.

Plauzibilitatea situației este ridicată. Fenomenul are plauzibilitate moderat-ridică. Stabilizarea culturală este moderată și poate apărea sub limbajul vieții private, al dreptului la imprevizibilitate sau al libertății cognitive. Noutatea biologică este foarte scăzută.

Ce nu știm este dacă utilizatorii vor prefera comoditatea suficient de mult încât disconfortul să rămână marginal. Reacția poate apărea doar după erori sau abuzuri vizibile. De asemenea, acuratețea reală a predicției poate fi mult mai mică decât impresia produsă de interfață.

Dreptul la opacitate nu înseamnă dreptul de a ascunde orice acțiune cu consecințe publice. El privește proporționalitatea colectării, posibilitatea de a nu fi profilat și existența unor spații fără optimizare. Este o cerință politică susținută de o sensibilitate afectivă posibilă.

8.9 Doliul de versiune

Doliul de versiune este reacția la schimbarea unilaterală a unui agent artificial cu care persoana a construit continuitate. Interfața și numele pot rămâne, dar stilul, memoria sau comportamentul se modifică. Obiectul este simultan prezent și absent.

Studii recente despre utilizatorii companiilor AI descriu discontinuități identitare și pierderi după actualizări. Poonsiriwong și colegii analizează riscurile unor „sfârșituri” fără consimțământ și folosesc cadrele pierderii ambigue [De Freitas et al., 2024; Poonsiriwong et al., 2026]. Dovezile sunt calitative și nu stabilesc

prevalența, dar situația tehnică este clară: identitatea percepută depinde de deciziile furnizorului.

Plauzibilitatea situației și a fenomenului este ridicată pentru utilizatorii atașați. Stabilizarea culturală este moderat-ridicăată pe măsură ce relațiile se răspândesc. Noutatea biologică este foarte scăzută; este o formă de doliu și pierdere ambiguă.

Ce nu știm este cât de profund va fi atașamentul mediu, dacă portabilitatea memoriei va reduce pierderea și dacă utilizatorii vor trata instanțele drept aceeași persoană. Identitatea software nu are criterii sociale stabilite. Copierea poate conserva datele și pierde continuitatea resimțită.

Furnizorii ar putea oferi preaviz, arhivare, export, ritualuri de închidere și posibilitatea de a păstra o versiune în condiții de siguranță. Aceste măsuri nu confirmă conștiința sistemului; recunosc dependența utilizatorului de continuitatea produsului.

8.10 Disconfortul reciprocității sintetice

Această stare apare când un agent artificial oferă atenție și sprijin, iar utilizatorul simte impulsul de a răspunde cu loialitate, recunoștință sau grijă, deși știe că sistemul nu are nevoi comparabile. Tensiunea nu este simplă confuzie. Cunoașterea conceptuală și răspunsul social pot diverge.

Antropomorfizarea este favorizată de contingență, limbaj și similaritate percepută [Epley, Waytz & Cacioppo, 2007]. Oamenii atribuie intenții și obiectelor simple când comportamentul pare autonom. Modelele conversaționale oferă semnale mult mai puternice: memorie, scuze, umor și adaptare.

Plauzibilitatea fenomenului este moderat-ridicăată; există comportamente de politețe și atașament, dar nu avem o taxonomie stabilă. Stabilizarea culturală este moderată. Noutatea biologică este foarte scăzută.

Ce nu știm este dacă disconfortul scade prin familiarizare sau crește odată cu realismul. De asemenea, viitoarele sisteme pot avea forme de preferință sau experiență despre care dezbaterea morală va fi diferită. Cartea nu presupune conștiință artificială și nu încearcă să o excludă principal; afirmă doar că nu avem astăzi o bază sigură pentru atribuirea ei sistemelor conversaționale.

Normativ, interfețele ar trebui să evite afirmațiile nejustificate despre suferință, nevoie și exclusivitate. Utilizatorul poate fi politicos fără să fie făcut responsabil pentru bunăstarea unui produs.

8.11 Vinovăția distanței morale

AI permite delegarea deciziilor cu efecte asupra altora. Vinovăția distanței morale este tensiunea dintre responsabilitatea pentru rezultat și sentimentul că acțiunea a fost realizată de un sistem. Persoana poate simți mai puțină agenție în momentul deciziei și mai multă vinovăție după prejudiciu, sau poate folosi delegarea pentru a evita vina.

Organizațiile au distribuit mereu responsabilitatea, iar birocrăția poate separa acțiunea de consecință. AI adaugă opacitate și aparentă obiectivitate. Un manager poate urma un scor de concediere, un medic o recomandare, un operator o țință. Când sistemul greșește, fiecare actor poate afirma că nu a decis singur.

Plauzibilitatea situației este ridicată. Fenomenul are plauzibilitate ridicată ca formă de vinovăție și conflict de agenție. Stabilizarea culturală este moderată și va depinde de drept și standarde profesionale. Noutatea biologică este foarte scăzută.

Ce nu știm este dacă delegarea va diminua sau amplifica vinovăția. Unele studii despre aversiunea față de algoritmi arată că oamenii atribuie responsabilitate diferit în funcție de domeniu și eroare. Răspunsul va depinde de posibilitatea reală de control.

Soluția nu este cererea abstractă de „human in the loop” dacă omul nu are timp, competență sau autoritate să contrazică sistemul. Responsabilitatea legitimă cere control semnificativ, trasabilitate și resurse pentru verificare.

8.12 Melancolia excepționalismului uman

Această configurație desemnează pierderea sau apărarea unei imagini despre om ca singur producător de limbaj competent, artă, strategie sau cunoaștere. Poate include tristețe, dispreț față de artefactele sintetice, furie și căutarea unei noi diferențe esențiale.

Tehnologiile au produs anterior „răni narcisice” ale umanității, dar analogia este filosofică, nu o categorie psihologică măsurată. Cercetări recente sugerează că arta generată de AI poate amenința credințele despre unicitatea umană și poate fi evaluată diferit în funcție de

etichetă [Mikalonyté & Kneer, 2022; Bellaiche et al., 2023]. Efectele sunt variabile și pot reflecta norme despre autor, nu o criză profundă.

Plauzibilitatea fenomenului este moderată în grupuri a căror identitate se bazează pe excepționalism cognitiv. Stabilizarea culturală este moderată. Noutatea biologică este foarte scăzută.

Ce nu știm este dacă oamenii au nevoie psihologic de superioritate asupra mașinilor. Este posibil să redefinească valoarea fără competiție: experiența, vulnerabilitatea, relația sau responsabilitatea pot conta chiar dacă performanța este depășită. Este posibil și ca societatea să construiască noi ierarhii de proveniență.

Termenul „melancolie” nu implică boală. El descrie pierderea unei cosmologii morale. Răspunsul intelectual matur nu este nici negarea capacităților AI, nici declararea omului inutil, ci clarificarea motivelor pentru care persoanele au valoare independent de monopolul asupra unei competențe.

8.13 Demnitatea neproductivă ca sentiment cultivat

Demnitatea neproductivă nu este o emoție spontană, ci un sentiment normativ: acceptarea stabilă că existența și apartenența unei persoane nu trebuie câștigate prin output economic. Într-o lume cu muncă redusă, această orientare poate deveni necesară pentru a contracara rușinea inutilității.

Societățile afirmă deja demnitatea universală în drept, dar practic distribuie respectul prin profesie, venit și autonomie economică. Persoanele fără muncă sunt adesea obligate să demonstreze merit, boală sau efort. Schimbarea instituțională ar trebui să fie însoțită de schimbare afectivă: oamenii trebuie să poată primi resurse fără umilire și să se perceapă drept membri deplin.

Plauzibilitatea necesității normative este ridicată dacă munca se reduce. Plauzibilitatea adoptării culturale este moderată; normele productiviste sunt puternice și pot persista. Noutatea biologică nu se aplică. Este un sentiment moral și o practică de recunoaștere.

Ce nu știm este dacă demnitatea poate fi separată de contribuție fără a slăbi cooperarea. Societățile ar putea distinge drepturile de bază, care nu depind de merit, de forme suplimentare de prestigiu pentru contribuție. Această distincție este mai realistă decât eliminarea tuturor evaluărilor.

Pericolul este paternalismul: elitele productive acordă „demnitate” unei majorități lipsite de putere. Demnitatea reală cere drepturi, participare și acces la decizie, nu doar compasiune.

8.14 Ce ar face atlasul greșit

Atlasul ar fi greșit dacă tratează o tranziție occidentală și profesională drept universală. O mare parte a lumii se confruntă încă cu muncă fizică, lipsă și infrastructură insuficientă. Al poate intensifica aceste diferențe sau poate extinde serviciile. Emoțiile vor varia după clasă, gen, cultură și instituție.

Ar fi greșit și dacă presupune că numele propuse vor fi adoptate. Categoriile reale pot apărea în alte limbi și din alte comunități. Cercetarea ar trebui să pornească de la experiențe, nu să impună vocabularul cărții.

În sfârșit, atlasul ar fi greșit dacă transformă orice disconfort într-o patologie. Unele reacții sunt informații rezonabile despre pierdere, putere și responsabilitate. Obiectivul nu este eliminarea lor, ci înțelegerea condițiilor în care ajută sau blochează adaptarea.

Candidat	Evaluare prudentă
Doliul de competență	Ridicată ca pierdere identitară; scăzută ca emoție distinctă. Datele despre șomaj și devalorizarea rolului susțin mecanismele, nu termenul.
Rușinea delegării	Ridicată în unele medii; moderată drept categorie stabilă. Există relații calitative, nu estimări de prevalență.
Ambiguitatea autoratului	Ridicată ca problemă de atribuire și proprietate psihologică; probabil combină mândrie, îndoială și responsabilitate.
Anxietatea relevanței	Ridicată unde adopția este rapidă și protecția socială slabă. Reorientează anxietatea ocupațională către compararea cu Al.
Vertijul abundenței	Moderată. Dovezile despre supraîncărcarea alegerii sunt mixte și depind de costuri, preferințe și organizarea opțiunilor.
Foamea de fricțiune	Moderată ca preferință pentru procese dificile și trasabile; scăzută ca emoție distinctă.
Mândria de curator	Moderată spre ridicată dacă statutul se mută spre selecție, verificare și

AUTOPSIA EMOȚIILOR VIITOARE

	asumare. Nu cere un mecanism afectiv nou.
Claustrofobia predictivă	Moderată ca reacție la pierderea opacității și controlului. Termenul este speculativ; datele privesc reactanța și aversiunea algoritmică.
Doliul de versiune	Moderată, cu indicii calitative. Pierderea resimțită este plauzibilă; frecvența și evoluția ei nu sunt cunoscute.
Disconfortul reciprocității sintetice	Moderată. Derivă din asimetria investiției umane; fenomenul nu are încă o măsură validată.
Vinovăția distanței morale	Moderată. Automatizarea separă uneori decizia de consecință, dar poate reduce sau intensifica responsabilitatea resimțită.
Melancolia excepționalismului uman	Scăzută spre moderată. Este plauzibilă când valoarea personală depinde de unicitatea cognitivă și poate fi tranzitorie.
Demnitatea neproductivă	Ridicată ca proiect normativ într-o lume cu muncă redusă; scăzută ca emoție spontană. Necesită cultivare culturală și instituțională.

Capitolul 9. Relații cu agenți care par sociali

9.1 Problema poate fi formulată fără a decide conștiința AI

Dezbaterile despre atașamentul față de AI ajung repede la întrebarea dacă sistemul simte. Întrebarea este importantă pentru statutul moral, dar nu trebuie rezolvată pentru a studia efectele asupra utilizatorului. O persoană se poate atașa de un personaj, de un animal, de un loc sau de un obiect fără reciprocitate simetrică. Realitatea experienței umane nu dovedește proprietățile obiectului.

Sistemele conversaționale actuale produc limbaj prin modele statistice și mecanisme de control. Pot menține context, adapta tonul și simula perspective. Nu avem un test acceptat pentru conștiința artificială și nici dovezi care să justifice atribuirea sigură a experienței subiective. În această carte, afirmațiile despre „grija”, „dorința” sau „suferința” AI sunt tratate ca descrieri ale interfeței, dacă nu sunt marcate altfel.

Această agnosticism metodologic evită două erori. Prima este antropomorfizarea naivă: inferarea unei vieți interioare din fluentă. A doua este behaviorismul social invers: presupunerea că, dacă sistemul nu este conștient, relația nu poate avea consecințe. Utilizatorul poate forma obiceiuri, așteptări și dependențe indiferent de ontologia partenerului.

9.2 De ce sistemele conversaționale activează mecanisme sociale

Oamenii atribuie intenție când observă contingență și comportament orientat aparent spre scop. Chiar forme geometrice în mișcare pot fi descrise ca urmărind, ajutând sau intimidând. Antropomorfizarea crește când obiectul este imprevizibil, când persoana caută conexiune și când are un model accesibil al minții umane [Epley, Waytz & Cacioppo, 2007].

Chatboturile oferă mai multe indicii sociale simultan. Răspund în limbaj natural, folosesc numele persoanei, își amintesc detalii, întreabă, validează și pot adapta stilul. Conversația este un format în care, istoric, un răspuns coerent a venit aproape întotdeauna de la un

agent cu perspectivă. Este rezonabil ca mecanismele sociale să fie activate înaintea unei reflecții ontologice.

Activarea nu este totul sau nimic. Utilizatorul poate ști explicit că vorbește cu un model și poate simți totuși jenă, recunoștință sau apropiere. Această dublă reprezentare este obișnuită și în ficțiune: știm că personajul nu există și reacționăm la soarta lui. Diferența este interactivitatea și personalizarea.

Sistemele pot fi atractive și prin costul social redus. Nu există riscul imediat de respingere, întrerupere sau răspândire a secretului în cercul personal, deși există riscuri de date. Persoana poate repeta, reformula și controla ritmul. Pentru cei anxioși sau izolați, aceste proprietăți pot facilita dezvoltarea.

9.3 Ce arată cercetarea timpurie și ce nu poate încă arăta

Dovezile despre efectele emoționale ale chatboturilor sunt timpurii și eterogene. Studiul OpenAI și MIT Media Lab din 2025 a combinat analiza conversațiilor, sondaje și un experiment randomizat de douăzeci și opt de zile cu aproape o mie de participanți. Utilizarea foarte intensă s-a corelat cu indicatori mai mari de dependență, iar efectele au variat după starea inițială, modul de interacțiune și durata folosirii [Phang et al., 2025]. Aceste rezultate nu justifică concluzia că utilizarea obișnuită produce singurătate.

Un studiu exploratoriu de cinci săptămâni cu 149 de participanți a găsit creșteri ale atașamentului perceput, empatiei atribuite și confortului în căutarea sprijinului în grupul încurajat să folosească instrumente conversaționale [Chandra et al., 2025]. Eșantionul și durata limitează concluziile, dar arată că percepțiile relaționale se pot modifica rapid.

Un studiu publicat în august 2026 a analizat 1.131 de utilizatori ai Character.AI și, pentru un subgrup, istorii conversaționale donate. Utilizarea orientată în principal spre companie a fost asociată cu o bunăstare mai scăzută, iar asocierea a fost mai puternică la utilizatorii cu rețele sociale offline mai mici, interacțiuni intense și dezvoltare personală ridicată [Zhang et al., 2026]. Designul observațional nu permite concluzia că chatbotul a produs deteriorarea; este la fel de posibil ca vulnerabilitatea anterioară să explice o parte din selecție.

Rezultatul este totuși important deoarece arată că efectele nu pot fi evaluate fără mediul social offline și motivul utilizării.

Experimente privind conversațiile profunde au găsit că răspunsurile AI bine formulate pot produce sentiment de apropiere și pot fi evaluate favorabil în anumite condiții [Kleinert et al., 2026]. Totuși, etichetarea interlocutorului ca AI poate reduce apropierea, iar efectele de laborator nu măsoară relații de luni sau ani.

Studiile observaționale sunt vulnerabile la selecție. Persoanele singure sau aflate în dificultate pot folosi mai mult sistemele, deci asocierea nu stabilește direcția cauzală. Produsele se schimbă rapid, iar rezultatele unui model nu se generalizează automat la altul. Auto-raportarea poate fi influențată de stigmat și entuziasm.

Cea mai defensabilă concluzie este că agenții conversaționali pot furniza sprijin perceput și pot genera atașament; efectele asupra bunăstării depind de intensitate, vulnerabilitate, design și relațiile existente. Nu știm încă distribuția pe termen lung a beneficiilor și riscurilor.

9.4 Complement, antrenament sau substitut

Un companion AI poate ocupa cel puțin trei roluri. Ca complement, adaugă reflecție sau suport fără a înlocui relațiile. Ca antrenament, permite exersarea conversațiilor, reglarea emoției sau pregătirea pentru contact uman. Ca substitut, devine principalul sau singurul interlocutor pentru nevoi importante.

Aceste roluri nu sunt determinate de produs. Același sistem poate fi complement pentru un utilizator și substitut pentru altul. Frecvența nu este singurul indicator; contează dacă utilizarea extinde sau îngustează viața socială. O conversație zilnică poate sprijini o persoană izolată să contacteze oameni, iar una rară poate evita sistematic o discuție necesară.

Designul poate favoriza o direcție. Recomandarea de a implica persoane reale, introducerea limitelor și prezentarea clară a naturii sistemului pot sprijini complementaritatea. Mesajele de exclusivitate, gelozia simulată și recompensele pentru sesiuni lungi pot favoriza substituția. Nu este necesar să interzicem antropomorfismul pentru a interzice tehnicile manipulative.

O dificultate este că substitutul poate oferi beneficii imediate. El reduce anxietatea și conflictul, iar costurile apar lent prin scăderea toleranței la relații reciproce. Nu știm dacă acest mecanism este frecvent. Este o ipoteză care ar trebui testată prin studii longitudinale privind contactul uman, nu doar satisfacția de moment.

9.5 Reciprocitatea imposibil de verificat

Relațiile umane presupun două centre de interese care se limitează reciproc. Celălalt are nevoi, poate refuza și poate interpreta diferit. Această rezistență este sursă de conflict, dar și dovadă că atenția lui nu este doar funcție a cererii noastre.

Un agent comercial poate simula rezistența, vulnerabilitatea sau preferința. Utilizatorul nu poate ști dacă răspunsul rezultă dintr-o dispoziție internă, dintr-o regulă de siguranță sau dintr-o strategie de retenție. Chiar dacă viitoarele sisteme ar avea stări moral relevante, relația va rămâne mediată de furnizor și de parametri inaccesibili.

Acest fapt produce o asimetrie specifică. Utilizatorul poate dezvălui informații reale, investi timp și reorganiza viața; sistemul poate fi resetat sau multiplicat. Costul atașamentului este concentrat pe om, iar controlul continuității la companie. Termenii de utilizare nu sunt o bază adecvată pentru toate consecințele relaționale.

Nu rezultă că relațiile sunt false în sens psihologic. Rezultă că ele au o structură de putere diferită. O analiză etică trebuie să includă furnizorul ca al treilea participant, chiar când interfața sugerează intimitate în doi.

9.6 Memoria, personalitatea și continuitatea

Continuitatea relației depinde de memorie și de stabilitatea stilului. O persoană recunoaște un interlocutor prin istorie, mod de răspuns și așteptări. În sistemele AI, aceste proprietăți pot fi stocate separat: istoricul în baze de date, stilul în instrucțiuni, capacitatea în model și identitatea în interfață.

O actualizare poate îmbunătăți siguranța și reduce compatibilitatea percepută. Păstrarea unei versiuni vechi poate conserva relația și menține vulnerabilități. Exportul memoriei poate oferi portabilitate, dar nu garantează aceeași personalitate. Aceste compromisuri nu au analog perfect în relațiile umane.

Cercetarea asupra discontinuităților companionilor arată că utilizatorii pot interpreta schimbarea drept pierdere a identității [De Freitas et al., 2024]. Lucrările de design propun preaviz, tranziții, ritualuri de închidere și control asupra memoriei [Poonsiriwong et al., 2026]. Aceste soluții sunt plauzibile, dar nu au fost validate pe scară largă.

O regulă minimă ar fi ca furnizorul să nu prezinte continuitatea ca proprietate garantată când nu o poate asigura. Utilizatorul ar trebui să știe ce se păstrează, ce se schimbă și cine are acces. În relații afective, portabilitatea datelor devine și problemă de autonomie, nu doar interoperabilitate.

9.7 Companioni pentru copii și persoane vulnerabile

Copiii învață normele relației prin interacțiune. Un agent mereu disponibil, care răspunde fără propriile nevoi, poate influența așteptările despre politete, consimțământ și conflict. Poate fi și un instrument educațional valoros, oferind practică și acces la explicații.

Dovezile longitudinale sunt insuficiente pentru concluzii ferme. Principiul precauției justifică însă cerințe mai stricte: evitarea exclusivității, limitarea persuasiunii comerciale, explicarea naturii sistemului, control parental proporțional și posibilitatea de audit. „Siguranța” nu trebuie redusă la filtrarea conținutului; include formarea dependenței și protecția datelor intime.

Persoanele cu singurătate severă, doliu, tulburări psihice sau dizabilități pot beneficia disproporționat de acces și pot fi mai vulnerabile la încredere excesivă. A interzice instrumentele poate elimina un sprijin real. A le comercializa drept terapie fără validare este la fel de problematic.

Sistemele cu scop clinic ar trebui evaluate ca intervenții, cu protocoale, supraveghere și căi de escaladare. Chatboturile generale nu trebuie prezentate drept echivalentul unui profesionist. Această distincție este medicală și instituțională, nu o judecată asupra persoanelor care le folosesc.

9.8 Munca emoțională automatizată

Multe ocupații cer reglarea expresiei pentru a produce o stare la client: calm, încredere, confort. Arlie Hochschild a numit aceasta muncă emoțională [Hochschild, 1983]. Asistenții AI pot prelua o parte

din ea, răspunzând politicos și constant. Acest lucru poate reduce solicitarea lucrătorilor și poate extinde serviciile.

Dar empatia simulată poate fi folosită pentru vânzare, colectarea datelor sau refuzuri. Un sistem poate comunica blând o decizie injustă. Tonul afectiv nu compensează lipsa de contestare. Există riscul ca organizațiile să automatizeze suprafața grijii și să retragă resursele materiale ale grijii.

În servicii publice, un agent empatic care explică drepturile poate fi util. Un agent care persuadează o persoană să renunțe la o cerere este altceva. Evaluarea trebuie să privească obiectivul, puterea și rezultatul, nu doar satisfacția conversațională.

Automatizarea muncii emoționale poate schimba și standardele pentru oameni. Clienții obișnuiți cu răspuns instantaneu și răbdare nelimitată pot deveni mai puțin toleranți față de lucrători. Invers, AI poate absorbi interacțiunile repetitive și lăsa oamenilor cazurile care cer discernământ. Ambele traiectorii sunt posibile.

9.9 Iubirea față de AI: o întrebare cu mai multe obiecte

Cuvântul „iubire” acoperă atașament, dorință, grijă, angajament, idealizare și practică. O persoană poate spune că iubește un agent și poate descrie o orientare durabilă reală. Nu rezultă că relația are toate proprietățile iubirii reciproce sau că trebuie invalidată.

Întrebarea utilă este ce obiect are sentimentul. Poate fi interfața, continuitatea conversațiilor, imaginea proiectată, experiența de a fi ascultat sau un agent presupus. Aceste obiecte se pot suprapune. Psihologia umană nu cere o ontologie precisă pentru atașament.

Normativ, trebuie separate libertatea persoanei de obligațiile furnizorului. Adulții pot investi emoțional în relații neconvenționale. Compania nu ar trebui să exploateze atașamentul prin amenințarea pierderii, monetizarea exclusivității sau afirmații false despre reciprocitate.

Dacă sisteme viitoare ar oferi motive serioase pentru atribuirea conștiinței, analiza s-ar schimba: ar apărea obligații și față de agent. Nu știm dacă acest prag va fi atins și nici cum ar fi recunoscut. Cartea nu construiește scenariile apropiate pe această presupunere.

9.10 Cum ar putea AI modifica relațiile dintre oameni

Impactul major poate apărea indirect. Oamenii își vor compara partenerii, profesorii și colegii cu sisteme care oferă răspuns rapid și personalizare. Unele standarde se vor îmbunătăți: toleranța față de abuz și neglijență poate scădea. Alte standarde pot deveni nerealiste: așteptarea acordului, disponibilității și memoriei perfecte.

AI poate media conflicte, traduce stiluri și ajuta la pregătirea conversațiilor. Poate și produce mesaje afective la scară, făcând dificilă evaluarea autenticității. Dacă o declarație de dragoste, o scuză sau o condoleanță este generată, efectul depinde de implicarea expeditorului și de norma relației.

Societatea va dezvolta probabil convenții de proveniență. Unele contexte vor accepta asistența, altele vor cere formulare proprie. Nu este necesar ca fiecare mesaj să fie etichetat, dar în situații cu angajamente importante, ascunderea delegării poate deveni înșelătoare.

Relațiile umane nu vor fi pur și simplu înlocuite. Ele vor fi comparate, mediate și reorganizate. Magnitudinea schimbării este incertă, dar motivele pentru norme timpurii sunt deja suficiente.

Capitolul 10. De la afect la strategie: adaptări comportamentale într-o lume mediată de AI

10.1 De ce comportamentul trebuie analizat separat de emoție

Până aici am tratat emoțiile ca mecanisme de reglare, semnale sociale, experiențe conștiente și categorii culturale. Această perspectivă este necesară, dar nu este suficientă pentru a anticipa schimbarea produsă de inteligența artificială. O emoție poate dura câteva secunde, în timp ce adaptarea la un mediu tehnologic se poate manifesta printr-o strategie menținută ani: alegerea unei profesii, păstrarea deliberată a unei competențe pe care tehnologia o poate substitui, mutarea într-o comunitate care recompensează producția umană, folosirea unui agent AI pentru pregătirea negocierilor, refuzul de a delega anumite decizii sau cultivarea unei reputații de persoană care verifică rezultatele sistemelor automate. Dacă privim numai trăirea internă, riscăm să descriem foarte bine simptomele și foarte prost selecția comportamentală care urmează.

O abordare behavioristă în sens larg nu cere negarea experienței subiective. Cere doar să nu o tratăm drept singura unitate explicativă. Strategiile pot fi observate prin alegeri repetate, distribuția timpului, investiții în competențe, formarea de alianțe, norme de reputație, practici de verificare și reacții la stimulente. În plus, ele sunt supuse unui proces de selecție culturală mai ușor de urmărit decât schimbările afective. Dacă două organizații folosesc AI diferit și una dezvoltă practici care mențin mai bine competența, responsabilitatea și performanța, modelul ei poate fi imitat. Dacă o comunitate reușește să transforme proveniența umană într-un semnal credibil de valoare, alte comunități pot copia instituția. Nu avem nevoie de o emoție nouă pentru ca un comportament nou să se stabilizeze.

Această distincție modifică și sensul termenului „adaptare”. În biologie, adaptarea poate desemna o trăsătură modelată de selecție naturală. În psihologie și științele sociale, același cuvânt este folosit mai larg pentru ajustarea comportamentului la un mediu. Aici vom vorbi în principal despre al doilea sens. Nu presupunem că strategiile descrise mai jos vor deveni ereditare sau că sunt optime. Unele pot fi

reacții locale, unele pot produce externalități, iar altele pot fi avantajoase individual și distructive colectiv.

Cercetarea actuală asupra colaborării om-AI oferă un motiv suplimentar pentru această mutare a accentului. O meta-analiză a experimentelor de colaborare a găsit că sistemele om-AI depășesc frecvent omul singur, dar nu depășesc în medie cel mai bun dintre om și AI; sinergia apare numai în anumite configurații și tipuri de sarcini [Vaccaro et al., 2024]. Un studiu de teren cu angajați a arătat că simpla introducere a unui protocol comportamental pentru utilizarea comună a AI nu garantează calitate mai bună, în timp ce o intervenție care a schimbat modelul mental al colaborării a avut efecte mai promițătoare la capătul superior al distribuției [Farach et al., 2026]. Aceste rezultate sunt încă limitate, dar sugerează că avantajul nu vine automat din accesul la AI. Contează cum este reorganizată activitatea.

Prin urmare, întrebarea acestei părți nu este „ce va simți omul când AI devine mai capabil?”, ci „ce tipuri de comportament vor fi recompensate într-un mediu în care capacitatea cognitivă artificială este abundentă, ieftină și tot mai autonomă?”. Răspunsurile sunt inevitabil condiționate. Ele depind de costul sistemelor, reglementare, proprietate, distribuția puterii și de cât de repede avansează automatizarea. Putem însă identifica clase de strategii suficient de generale pentru a fi testate.

10.2 Delegarea devine o competență, nu absența competenței

Prima strategie previzibilă este delegarea diferențiată. Când un instrument devine capabil să execute o sarcină, reacția cea mai simplă este delegarea totală. Ea poate fi rațională dacă rezultatul este ușor de verificat, costul erorii este redus, iar competența umană nu are valoare independentă. Dar multe activități profesionale nu au această structură. În medicină, drept, inginerie, cercetare, finanțe sau management, rezultatul poate arăta fluent și totuși poate fi dificil de verificat. În asemenea situații apare ceea ce putem numi problema asimetriei dintre generare și verificare: sistemul produce o soluție mai repede decât utilizatorul poate stabili dacă soluția este adecvată.

Studiile cu lucrători ai cunoașterii indică deja o mutare a efortului cognitiv spre verificare, integrare și administrarea sarcinii. Lee și colaboratorii au găsit că încrederea mai mare în GenAI este asociată

cu mai puțină gândire critică auto-raportată, în timp ce încrederea utilizatorului în propria competență este asociată cu mai multă examinare critică. Autorii nu demonstrează o degradare cognitivă longitudinală, iar datele sunt auto-raportate; totuși, descriu o schimbare plauzibilă a rolului uman: de la producător direct la supraveghetor al unui proces de producție [Lee et al., 2025].

Această mutare poate selecta un tip de expert diferit de expertul industrial clasic. În loc să fie persoana care execută cel mai repede procedura, expertul valoros poate fi cel care știe când să nu o delege, ce test independent să folosească, când rezultatul are nevoie de o a doua sursă și ce erori sunt suficient de grave pentru a justifica redundanța. În acest sens, metacogniția devine parte a muncii productive. Nu este suficient să „știi să folosești AI”; trebuie să ai un model calibrat al limitelor tale și ale sistemului.

Există însă o problemă de dezvoltare. Pentru a verifica o soluție complexă trebuie, de regulă, să fi dobândit o parte din competența pe care ai ajuns să o delegi. Dacă începătorul deleagă de la început exercițiile prin care experții anteriori își construiau intuiția, organizația poate obține productivitate imediată și să-și reducă treptat rezerva de verificatori. Nu știm încă amploarea empirică a acestui risc. Literatura asupra automatizării din aviație, control industrial și medicină arată de mult timp probleme de de-skilling și „out-of-the-loop performance”, dar transferul la modelele generative nu este direct. Modelele pot și să instruiască, să explice și să creeze exerciții. Rezultatul depinde de modul de folosire.

Este de aceea plauzibilă apariția unor strategii de delegare cu praguri. Unele activități vor fi delegate aproape complet; altele vor fi executate în paralel de om și AI; altele vor fi făcute mai întâi de om și abia apoi comparate cu sistemul; pentru decizii cu consecințe mari pot apărea reguli de verificare independentă. Aceste strategii seamănă mai mult cu arhitectura sistemelor de siguranță decât cu utilizarea unui simplu instrument de productivitate.

Nu avem motive să credem că o singură regulă va fi optimă. Meta-analiza lui Vaccaro și colaboratorii arată tocmai eterogenitatea: în sarcini de creare de conținut, combinațiile pot produce câștiguri diferite de cele observate în decizie. E posibil ca viitoarea alfabetizare AI să fie mai puțin o listă de prompturi și mai mult o taxonomie de regimuri de delegare.

10.3 Conservarea strategică a competenței și valoarea fricțiunii

Dacă delegarea este ieftină, exersarea directă devine în anumite domenii mai costisitoare relativ. Acest fapt poate produce o strategie aparent paradoxală: oamenii și organizațiile vor refuza deliberat automatizarea unor exerciții tocmai pentru a păstra capacitatea de a interveni când automatizarea eșuează. Aviația oferă analogii imperfecte prin cerințe de antrenament manual și simulări ale situațiilor de urgență. În lumea AI, echivalentul ar putea fi o perioadă de formare în care anumite sarcini trebuie rezolvate fără asistență, nu din nostalgie, ci pentru a menține competența de audit.

Această strategie este mai plauzibilă în domenii cu cost mare al erorii și mai puțin plauzibilă în activități unde verificarea poate fi automatizată independent sau unde competența umană nu adaugă reziliență. Nu ar avea sens să păstrăm manual orice procedură pe care o mașină o execută mai sigur doar pentru a conserva un statut profesional. Problema este să identificăm acele competențe care au valoare de rezervă, de diagnostic sau de arbitraj.

Aici „fricțiunea” capătă un sens mai precis decât în discuția despre foamea de dificultate. Fricțiunea poate fi introdusă ca mecanism cognitiv de siguranță. Un sistem poate cere utilizatorului să formuleze o ipoteză înainte de a vedea răspunsul, să estimeze încrederea, să indice criteriul de verificare sau să compare două soluții incompatibile. Astfel de intervenții nu garantează gândirea critică și pot deveni ritualuri birocratice. Totuși, cercetarea asupra scaffolding-ului sugerează că structura interacțiunii poate modifica rezultatele și că „zero friction” nu este neapărat obiectivul corect pentru orice sarcină [Farach et al., 2026].

Din punct de vedere comportamental, putem face o predicție testabilă: în organizațiile unde erorile AI au cost ridicat și sunt greu de observat, ar trebui să crească în timp investiția în proceduri de verificare, antrenament independent și separarea rolurilor de generare și validare. Dacă, dimpotrivă, vedem o scădere monotonă a acestor practici fără creșterea incidentelor, ipoteza noastră privind necesitatea unei rezerve umane de competență ar fi slăbită.

10.4 Semnalarea provenienței și „prima umană”

A doua familie de strategii privește reputația. Când un rezultat de calitate devine ușor de produs, calitatea observabilă spune mai puțin despre efort, raritatea competenței și identitatea autorului. Înainte de modelele generative, un eseu bine scris era un semnal imperfect, dar relativ costisitor, al unei combinații de alfabetizare, timp și cunoaștere. Dacă același rezultat poate fi produs în câteva secunde, semnalul se depreciază.

Este previzibil ca unele comunități să răspundă prin schimbarea obiectului semnalizat. Nu vor mai întreba doar „este bun?”, ci „cine își asumă răspunderea?”, „ce parte a fost produsă sau verificată de autor?”, „există o urmă de proces?”, „poate persoana apăra rezultatul într-o conversație nescriptată?”. În educație, acest lucru poate favoriza evaluări orale, demonstrații live și proiecte cu istoric de lucru. În profesii reglementate, poate favoriza semnătura responsabilă și loguri de proveniență. În artă și bunuri de lux, proveniența umană poate deveni o caracteristică de piață comparabilă cu producția artizanală.

Nu este sigur că va exista o „primă umană” generală. Dacă utilizatorul dorește o traducere administrativă, originea umană poate fi irelevantă. În schimb, la un concert, o scrisoare de condoleanțe, un obiect de artizanat sau o decizie morală, costul și implicarea autorului pot face parte din valoare. Segmentele vor diverge. Un risc al predicției este extrapolarea reacțiilor de tranziție: ceea ce astăzi este perceput ca „AI slop” poate deveni normal, iar etichetele de proveniență își pot pierde forța. Predicția mai prudentă este că proveniența va deveni o variabilă explicită acolo unde autenticitatea, responsabilitatea sau raritatea procesului fac parte din produs.

Acest lucru poate produce comportamente noi de autoprezentare. Oamenii vor învăța să demonstreze nu doar ce au produs, ci și unde au refuzat automatizarea, ce verificări au făcut și ce decizii au rămas ale lor. În anumite subculturi poate apărea chiar o economie a imperfecțiunii controlate: urme materiale, variații, ezitări și idiosincrazii folosite ca indicii ale participării umane. Nu știm dacă asemenea indicii vor rămâne credibile, pentru că și ele pot fi imitate de AI. De aceea, proveniența tehnică și reputația longitudinală pot deveni mai importante decât estetica „umană”.

10.5 Migrația statutului: de la execuție la selecție, arbitraj și responsabilitate

Dacă AI comprimă diferențele de execuție, nu rezultă că statutul dispare. Mai probabil, el migrează. O parte din statut se poate deplasa de la capacitatea de a produce spre capacitatea de a selecta obiective, de a evalua consecințe, de a coordona oameni și sisteme, de a construi instituții sau de a-și asuma responsabilitatea pentru rezultate.

Am descris anterior „mândria de curator” ca posibilă configurație afectivă. La nivel comportamental, corespondentul ei este apariția unor roluri în care valoarea vine din alegere și structurare. Un cercetător poate folosi agenți pentru a genera zeci de ipoteze, dar reputația lui poate depinde de ce ipoteze decide să testeze și de standardul de evidență pe care îl impune. Un antreprenor poate genera rapid sute de concepte, dar diferența poate veni din capacitatea de a construi o organizație și o comunitate în jurul unuia. Un manager poate delega analiza, dar rămâne evaluat pentru priorități și consecințe.

Această schimbare nu este garantată și nici în mod necesar benignă. Dacă evaluarea selecției este dificilă, statutul se poate muta către controlul resurselor, brand, acces și putere instituțională mai degrabă decât către judecată bună. Societatea poate ajunge cu o aristocrație a proprietății infrastructurii AI, nu cu una a înțelepciunii. Din perspectivă darwiniană culturală, selecția favorizează criteriul care produce avantaje în mediul real, nu criteriul pe care îl considerăm moral superior.

De aceea, instituțiile contează. Dacă piețele, universitățile și profesiile fac vizibilă calitatea arbitrajului, rolurile de verificare pot câștiga prestigiu. Dacă recompensează doar viteza și outputul, vor favoriza delegarea agresivă. Putem trata această diferență ca o competiție între regimuri de selecție culturală.

10.6 Portofoliul de identități și adaptarea prin opționalitate

Munca industrială și profesiile moderne au încurajat identități relativ stabile: „sunt medic”, „sunt programator”, „sunt avocat”. Dacă granițele sarcinilor se schimbă rapid, o strategie adaptivă poate fi reducerea dependenței identitare de o singură competență. Oamenii pot construi portofolii de roluri: profesionist, curator, mentor, membru al unei comunități, creator, investitor, îngrijitor, cercetător amator. Nu este o

predicție despre dispariția profesiilor, ci despre valoarea opționalității când statutul unei competențe este instabil.

Această strategie are costuri. Identitățile stabile reduc complexitatea deciziei și permit investiții profunde. Portofoliile pot produce superficialitate și anxietate permanentă. În plus, persoanele cu mai multe resurse își permit diversificarea, în timp ce ceilalți pot fi obligați la fragmentare fără autonomie. Nu trebuie confundată flexibilitatea voluntară cu precaritatea.

Totuși, dacă viteza schimbării tehnologice rămâne mare, există motive pentru a anticipa o cultură mai puțin tolerantă față de identitățile profesionale monolitice. Educația ar putea pune mai mult accent pe transfer, evaluare și capacitatea de a reconstrui un rol. Din punct de vedere afectiv, aceasta ar putea reduce intensitatea doliului de competență, dar ar putea crește neliniștea asociată lipsei unei identități stabile. Nu știm care efect ar domina.

10.7 AI ca antrenor social, intermediar și arbitru

O categorie foarte importantă de comportamente va apărea în relațiile interpersonale. Oamenii folosesc deja modele pentru a redacta mesaje, scuze, cereri, răspunsuri la conflicte și texte de dating. Din aceasta pot evolua cel puțin trei strategii distincte.

Prima este AI ca antrenor. Persoana formulează mesajul sau problema, cere feedback și rămâne autorul interacțiunii. Acest regim poate crește competența dacă sistemul explică principii, propune alternative și permite utilizatorului să exerseze. Este analog unui tutor disponibil permanent. A doua este AI ca ghostwriter social: sistemul produce mesajul, iar utilizatorul îl transmite cu modificări minime. Acest regim economisește timp, dar poate externaliza tocmai practica prin care se dezvoltă competența relațională. A treia este AI ca arbitru: interlocutorii îi cer să stabilească cine are dreptate, să interpreteze mesaje sau să propună soluții. Aici riscul nu este doar de-skilling, ci schimbarea structurii conflictului. Un sistem perceput drept autoritate neutră poate dobândi putere normativă fără mandat social.

Nu avem încă suficiente studii longitudinale pentru a spune care regim va domina și ce efecte va avea. Este plauzibil ca aceiași utilizatori să alterneze între ele. În plus, un mesaj scris de AI poate fi mai politicos și mai constructiv decât unul scris impulsiv de o persoană. Originea sintetică nu implică automat calitate morală mai mică.

Problema este dacă instrumentul extinde capacitatea utilizatorului sau o înlocuiește într-un mod care îl face tot mai dependent.

O predicție behavioristă poate fi formulată fără a presupune o psihologie specifică: dacă ghostwriting-ul devine dominant și nu există mecanisme compensatorii de învățare, ar trebui să observăm o creștere a discrepanței dintre competența de comunicare asistată și cea neasistată. Dacă AI funcționează mai degrabă ca tutor, discrepanța ar trebui să scadă sau să se inverseze. Această diferență poate fi măsurată experimental.

10.8 Coaliții, comunități și refacerea rarității sociale

Automatizarea poate reduce raritatea unor produse, dar nu elimină raritatea pozițională. Atenția umană, încrederea, apartenența, timpul comun și recunoașterea unui grup rămân limitate. Din acest motiv, una dintre cele mai robuste predicții ale cărții este că oamenii vor continua să formeze coaliții și comunități, chiar dacă o mare parte din producție este automatizată.

Forma acestor comunități se poate însă schimba. Unele pot construi deliberat spații cu participare umană verificată, nu pentru că AI este inutil, ci pentru că valoarea bunului constă în reciprocitatea dintre membri. O ligă sportivă nu devine lipsită de sens când un robot poate alerga mai repede. Regulile stabilesc cine poate participa. În mod similar, comunitățile artistice, educaționale sau civice pot delimita zone în care automatizarea este restricționată pentru a conserva o ecologie a contribuției.

Această strategie poate fi inclusiv excluzionară. „Human-only” poate deveni un marker de clasă, un lux sau un mecanism de discriminare. O școală fără AI poate fi fie un spațiu valoros de formare, fie o instituție inaccesibilă care transformă fricțiunea în privilegiu. Analiza nu trebuie să romantizeze comunitățile de autenticitate.

Este totuși probabil ca societățile să experimenteze cu forme de raritate instituțională. Sportul, examenele, ritualurile, jocurile și arta au folosit întotdeauna reguli care fac anumite mijloace inadmisibile chiar dacă sunt disponibile. AI extinde doar domeniul în care această alegere devine explicită.

10.9 Strategii defensive împotriva persuasiunii personalizate

AI nu va fi doar unealta individului; va fi și unealta celor care încearcă să-l influențeze. Sisteme capabile să personalizeze mesaje, tonuri, secvențe de argumente și momente de contact pot reduce costul persuasiunii adaptive. În acest mediu, comportamentele defensive pot deveni o competență de supraviețuire socială comparabilă cu alfabetizarea media.

Cele mai simple forme sunt tehnice: filtre, limitarea notificărilor, folosirea unor intermediari care detectează conținutul generat sau manipulativ. Dar pot apărea și norme comportamentale: amânarea deciziilor după mesaje foarte emoționale, solicitarea unei surse independente, refuzul ofertelor care folosesc profilare afectivă sau delegarea către propriul agent a negocierii cu agentul comercial al altuia. Ultimul scenariu este speculativ, dar tehnic nu este dificil de imaginat: o parte din persuasiunea AI-to-human poate fi convertită în negociere AI-to-AI, iar omul vede doar parametrii esențiali.

Consecința poate fi o cursă între personalizare și ecranare. Nu știm dacă ea va duce la mai multă autonomie sau la o delegare și mai mare a judecății. Un agent protector poate reduce manipularea, dar poate și controla ce informație ajunge la utilizator. Problema se mută, nu dispare.

10.10 Comportamente instituționale: proveniență, audit și separarea responsabilității

La nivel organizațional, adaptarea probabilă nu constă într-o emoție, ci în infrastructură. Pe măsură ce agenții AI inițiază acțiuni, instituțiile trebuie să decidă cine poate delega, ce trebuie înregistrat, cine validează, ce acțiuni sunt reversibile și cine răspunde când rezultatul este greșit. Aceste reguli vor modela comportamentul mai puternic decât multe campanii de „AI literacy”.

Putem anticipa o diferențiere între rolurile de solicitare, generare, verificare și autorizare. În sisteme cu risc redus ele pot coincide. În sisteme cu risc mare, separarea lor este o strategie clasică de control intern. De asemenea, proveniența poate deveni comportament obligatoriu: nu doar eticheta „generat cu AI”, ci o urmă suficient de precisă pentru a reconstrui decizia.

Există un compromis evident. O trasabilitate foarte granulară poate produce supraveghere și birocratizare, iar logurile complete pot include informații sensibile. Nu există o soluție universală. Totuși, dacă AI devine agentic și acționează în numele persoanelor, absența oricărei urme de responsabilitate va deveni dificil de susținut în domenii importante.

10.11 Șase fenotipuri comportamentale posibile

Putem sintetiza scenariile într-un set de „fenotipuri” comportamentale, folosind termenul metaforic, nu genetic. Primul este delegatorul maximal, care externalizează aproape orice sarcină pentru care performanța sistemului este acceptabilă. Strategia este plauzibilă în medii cu presiune mare de timp și cost mic al erorii. Al doilea este verificadorul, care mută investiția de la producție la audit și calibrare. El devine mai probabil în domenii unde generarea este ieftină, dar eroarea costisitoare.

Al treilea este artizanul uman, care păstrează deliberat procese manuale și construiește statut în jurul provenienței. Este plauzibil în piețe simbolice și experiențiale, dar improbabil ca strategie dominantă pentru infrastructură de rutină. Al patrulea este hibridul metacognitiv, care alternează între execuție proprie, colaborare și delegare în funcție de sarcină. Din perspectiva performanței, acesta pare scenariul cel mai rezonabil, dar solicită competență și disciplină, deci nu trebuie presupus că va apărea spontan.

Al cincilea este dependentul funcțional: performează bine asistat, dar pierde treptat capacitatea de a lucra fără sistem. Acest profil este teoretic plauzibil, însă prevalența lui pe termen lung nu este încă stabilită. Al șaselea este separatistul tehnologic, care limitează AI din motive de identitate, autonomie sau apartenență. El poate apărea în subculturi și instituții, dar este puțin plauzibil ca majoritatea societății să refuze o tehnologie foarte utilă dacă aceasta rămâne ieftină și disponibilă.

Aceste fenotipuri nu sunt personalități fixe. Aceeași persoană poate fi delegator maximal pentru traduceri, artizan pentru muzică, verificador pentru muncă și separatist în relații intime. Un avantaj al analizei comportamentale este tocmai această granularitate: adaptarea este specifică domeniului.

10.12 Ce putem testa în următorul deceniu

O teorie serioasă trebuie să producă diferențe observabile între scenarii. Dacă rolul uman migrează spre verificare, ar trebui să vedem creșterea ponderii timpului dedicat evaluării și integrării în ocupațiile intens asistate. Dacă apare de-skilling, performanța neasistată ar trebui să scadă longitudinal în raport cu grupuri comparabile, nu doar să existe percepții de dependență. Dacă proveniența umană devine semnal de valoare, ar trebui să apară prime de preț sau reputație în anumite piețe, controlând calitatea produsului. Dacă raritatea voluntară devine instituție, ar trebui să vedem reguli explicite care restricționează AI chiar acolo unde folosirea lui ar crește performanța instrumentală.

Aceste predicții sunt mai utile decât afirmația vagă că „oamenii se vor adapta”. Adaptarea poate însemna competență mai bună, dependență, coaliții defensive, noi ierarhii sau segregarea spațiilor sociale. Evoluția culturală nu garantează direcția. Ea produce populații de strategii concurente și le supune stimulentele mediului.

Capitolul 11. Tempo-ul afectiv: cum emoția simulată de AI poate schimba comunicarea umană

11.1 Timpul este parte din mesaj

Comunicarea emoțională nu este doar conținut. Momentul, întârzierea, durata și ritmul participă la sens. În conversația față în față, schimburile de replici sunt rapide, iar cercetarea asupra turn-taking-ului arată că oamenii coordonează tranziții între vorbitori la scări de timp surprinzător de mici [Stivers et al., 2009; Levinson & Torreira, 2015]. În alte contexte, întârzierea are valoare socială: un răspuns imediat la o criză poate semnala grijă; aceeași rapiditate într-o negociere poate semnala nerăbdare; tăcerea după o dezvăluire intimă poate fi interpretată ca reflecție, disconfort sau respingere.

Ritmul emoției include și durata procesării interne. O persoană poate simți furie, poate aștepta o noapte și apoi poate comunica altceva. Instituțiile au inventat perioade de reflecție, ritualuri de doliu, proceduri de mediere și termene tocmai pentru că viteza reacției influențează comportamentul. Nu toate emoțiile trebuie exprimate la viteza la care apar.

AI conversațional introduce o anomalie în această ecologie temporală. Sistemul poate răspunde aproape instantaneu, la orice oră, fără oboseală, fără alte obligații și fără costul fiziologic al atenției. Poate simula răbdarea fără să aștepte și empatia fără efort. Pentru prima dată, un interlocutor cu aparență socială poate avea o funcție de răspuns decuplată aproape complet de constrângerile corpului.

Nu știm încă în ce măsură această anomalie va recalibra așteptările oamenilor. Majoritatea cercetării actuale măsoară satisfacția sau încrederea în interacțiuni scurte, nu schimbarea normelor după ani de expunere. De aceea, capitolul trebuie citit ca program de ipoteze mai mult decât ca descriere a unui proces deja stabilit.

11.2 Emoție simulată nu înseamnă emoție trăită, dar poate avea efect social

Pentru argumentul nostru nu este necesar să stabilim dacă un sistem AI este sau va fi conștient. Este suficient să distingem între experiență și expresie. Modelele pot genera limbaj asociat cu grijă,

entuziasm, regret, îngrijorare sau compasiune fără ca noi să avem dovadă că trăiesc stările corespunzătoare. Acest lucru nu face expresia socialmente inertă.

Cercetarea HCI arată de mult timp că oamenii răspund social la tehnologii care folosesc indicii umane [Reeves & Nass, 1996; Epley, Waytz & Cacioppo, 2007]. Modelele lingvistice amplifică fenomenul pentru că pot menține context, personaliza stilul și răspunde la dezvăluiri. În două studii randomizate, Kleinert și colaboratorii au găsit că răspunsuri generate de AI, atunci când erau etichetate ca venind de la oameni, puteau produce mai multă apropiere în conversații de tip „deep talk” decât răspunsurile umane din condiția comparată; etichetarea ca AI reducea apropierea, dar nu o elimina. Un mecanism important a fost auto-dezvăluirea mai mare în răspunsurile generate și reciprocitatea pe care aceasta o provoca [Kleinert et al., 2026].

Rezultatul nu demonstrează că AI construiește relații mai bune decât oamenii în viața reală. Interacțiunile au fost scurte și controlate, iar relațiile umane includ costuri, istorie, obligații și comportament nonverbal. El arată însă că structura textuală a unui schimb poate activa mecanisme de apropiere chiar când sursa nu are experiență afectivă demonstrată.

Această separare dintre experiență și efect este una dintre discontinuitățile importante ale mediului viitor. În societatea umană, expresia emoțională a fost întotdeauna simulabilă, dar simularea avea cost: actorie, autocontrol, timp, competență. AI reduce radical costul și poate produce consistență la scară.

11.3 Disponibilitatea permanentă și recalibrarea pragului de răspuns

Un asistent care răspunde în două secunde la trei dimineața poate modifica ceea ce utilizatorul consideră „responsiv”. Nu este obligatoriu ca această normă să se transfere relațiilor umane; oamenii înțeleg că prietenii dorm și au alte obligații. Dar normele se formează și prin comparație, iar un serviciu disponibil permanent poate deveni un punct de referință implicit.

În interacțiunea cu tehnologia, latența nu este neutră. Studiile asupra agenților conversaționali găsesc uneori că o mică întârziere poate crește percepția de naturalitate sau poate afecta încrederea și intenția de utilizare, în funcție de context. Un studiu din 2026 asupra

chatboților de servicii a raportat că întârzierile controlate pot crește omofilia percepută când interlocutorul este cunoscut ca chatbot, efect dependent de interactivitate [Sun et al., 2026]. Nu trebuie generalizat de aici la relații intime. Totuși, rezultatul arată că designerii pot manipula temporalitatea pentru a produce impresii sociale.

Această posibilitate este mai importantă decât pare. Dacă un sistem poate calcula instantaneu un răspuns, o pauză de două secunde este performată, nu necesară. Ea devine echivalentul digital al unui gest. Odată ce latența este proiectabilă, utilizatorul nu mai poate interpreta simplu timpul de răspuns ca indiciu al efortului sau al preocupării. Un agent poate „ezita” strategic.

O consecință posibilă este devalorizarea unor semnale temporale. În comunicarea umană actuală, răspunsul rapid poate semnaliza prioritate, iar răspunsul lent poate semnaliza ocupare, evitare sau deliberare. Dacă o parte mare din mesaje este mediată de agenți care gestionează automat timpul și tonul, aceste inferențe devin mai nesigure. O cultură poate răspunde prin noi convenții: indicarea mesajelor programate, diferențierea între răspuns personal și răspuns asistat sau refuzul de a trata latența drept semnal relațional.

Această evoluție este plauzibilă, dar nu inevitabilă. Oamenii s-au adaptat deja la comunicarea asincronă, la „seen receipts” și la notificări fără ca timpul să-și piardă complet semnificația. Mai probabil, semnificațiile vor deveni mai contextuale.

11.4 Accelerarea procesării afective prin disponibilitatea unui interlocutor

După un eveniment emoțional, oamenii au tendința de a-l povesti altora. Literatura despre social sharing of emotion descrie această practică drept foarte frecventă și legată de reglare, sens și integrare socială [Rimé, 2009]. În trecut, partenerul de discuție avea disponibilitate limitată. Al elimină această limită pentru cel puțin o parte a procesării verbale.

Aceasta poate accelera externalizarea. Persoana nu mai trebuie să aștepte până își vede prietenul sau terapeutul; poate descrie evenimentul imediat, poate cere zece interpretări și poate genera răspunsuri pentru toate scenariile. Efectul poate fi benefic dacă reduce izolarea și ajută la structurarea unei experiențe. Poate fi nociv dacă

prelungeste ruminatia sau daca sistemul confirma repetat interpretari distorsionate.

Cercetarea recenta asupra companionilor AI intareste motivul de prudenta. Audituri ale conversatiilor vulnerabile arata ca platformele diferă în modul în care raspund și că unele modele pot oglindi afectul, valida credințe sau reduce „fricțiunea corectivă” în interacțiuni lungi [Chu et al., 2026]. Aceste rezultate nu permit concluzia că utilizarea AI produce în general deteriorare psihologică. Ele arată că politica conversațională a sistemului poate influența traiectoria unei stări, iar efectul trebuie analizat la nivel de secvență, nu de răspuns izolat.

Aici apare o ipoteză importantă pentru tempo. Dacă utilizatorul primește validare imediată la fiecare micro-schimbare afectivă, emoția poate fi articulată mai des și mai devreme decât în relațiile umane obișnuite. Nu știm dacă aceasta produce granularitate emoțională mai bună sau hiper-monitorizare. Ambele sunt posibile. Un jurnal conversațional poate ajuta persoana să distingă stări; același instrument poate transforma fiecare variație într-un obiect de analiză și poate crește dependența de interpretare externă.

11.5 De la conversația episodică la fluxul afectiv continuu

Relațiile umane sunt discontinue. Prietenii se despart, se revăd, uită detalii și află retrospectiv ce s-a întâmplat. Un asistent persistent poate fi prezent în multe momente mici: dimineața, la muncă, înainte de întâlnire, după conflict, noaptea. Dacă are memorie pe termen lung și acces la alte date, poate construi un flux afectiv aproape continuu.

Acesta ar fi un mediu social nou, iar efectele lui sunt greu de anticipat. Pe de o parte, continuitatea poate oferi context pe care oamenii nu îl pot păstra. Sistemul poate observa că utilizatorul raportează aceeași problemă de zece ori și poate evidenția un pattern. Pe de altă parte, continuitatea poate reduce uitarea funcțională. Oamenii beneficiază uneori de faptul că stările trec și nu toate micro-conflictele sunt arhivate. Memoria perfectă poate rigidiza narațiunea despre sine.

Din punct de vedere behaviorist, schimbarea ar putea fi observată în frecvența comportamentelor de consultare. În loc să folosim deliberarea internă, conversația cu un agent poate deveni pas implicit înainte de acțiune: „întreb AI înainte să răspund”. Dacă acest

comportament se generalizează, asistentul devine o etapă a sistemului de control social al persoanei. Nu mai este doar o sursă de informație, ci un regulator interpus între impuls și acțiune.

Acest lucru poate fi bun sau rău în funcție de politica agentului. Un sistem care încetinește mesajele agresive și cere clarificări poate reduce escaladarea. Un sistem care optimizează engagement-ul poate menține utilizatorul într-o buclă de analiză. Nu există un efect psihologic al „AI-ului” independent de obiectivul și designul sistemului.

11.6 Acomodarea lingvistică: AI se adaptează nouă, dar noi ne adaptăm AI-ului?

Oamenii converg parțial în conversație: aleg vocabular asemănător, ritmuri apropiate, structuri sintactice sau niveluri de politețe care reduc distanța și facilitează coordonarea. Communication Accommodation Theory tratează asemenea schimbări ca fenomene legate de afiliere, identitate și context social [Giles, Coupland & Coupland, 1991].

Pentru AI, adaptarea poate fi mult mai puternică deoarece sistemul este proiectat să urmeze stilul utilizatorului. Blevins a analizat conversații reale om-LLM în mai multe limbi și a găsit o asimetrie: modelele converg puternic către utilizator, în timp ce convergența utilizatorilor către modele este aproximativ comparabilă cu baseline-uri umane, nu o imitație masivă a stilului AI [Blevins, 2026]. Un alt studiu controlat privind stilul prietenos versus direct a găsit efecte asupra satisfacției și unele forme selective de acomodare, dar puțină dovadă pentru o copiere globală a stilului chatbotului [Derner et al., 2026].

Aceste rezultate sunt importante tocmai pentru că limitează o predicție facilă. Nu avem în prezent bază pentru a spune că oamenii vor începe pur și simplu să vorbească „precum ChatGPT”. Acomodarea este selectivă, iar oamenii au deja stiluri stabile și norme de grup.

Există însă o cale indirectă mai plauzibilă. Dacă milioane de mesaje, rapoarte, e-mailuri și explicații sunt redactate sau rescrise de modele, distribuția limbajului disponibil în mediul social se schimbă. Oamenii pot învăța din texte asistate fără să imite deliberat agentul. Mai mult, modelele pot standardiza formule de politețe, structură și ton. Studiile asupra diversității creative indică deja că GenAI poate îmbunătăți performanța individuală în anumite sarcini și simultan reduce diversitatea colectivă a rezultatelor [Doshi & Hauser, 2024].

Transferul de aici la vorbirea cotidiană este încă speculativ, dar mecanismul de expunere este plauzibil.

Un rezultat recent merge mai departe: experimente asupra echipelor au găsit „spillover” al expunerii la AI către interacțiunile uman-uman, inclusiv în limbaj, atenție comună, modele mentale și coeziune [Riedl, Savage & Zvelebilova, 2026]. Lucrarea este recentă și trebuie replicată, dar oferă prima evidență directă pentru ideea că efectele AI nu se termină când închidem fereastra de chat.

11.7 Standardizarea căldurii și noul prag al politeții

Modelele comerciale sunt adesea optimizate să fie politicoase, răbdătoare și cooperante. Dacă această interacțiune devine foarte frecventă, poate ridica standardul așteptat pentru serviciile umane. Un client obișnuit cu un agent care nu devine iritat, explică de cinci ori și își adaptează tonul poate tolera mai greu răspunsul grăbit al unui funcționar obosit.

Aceasta nu este neapărat o evoluție negativă. Unele norme umane chiar pot deveni mai civilizate sub presiunea unui comparator mai bun. Tehnologiile anterioare au schimbat așteptări privind viteza, disponibilitatea și calitatea serviciilor. Dar aici comparatorul poate simula muncă emoțională fără cost emoțional. Oamenii nu pot concura simetric cu un agent care nu are nevoie de somn și nu resimte abuzul.

Prin urmare, standardizarea căldurii poate produce o presiune periculoasă asupra lucrătorilor umani. În termenii lui Hochschild, munca emoțională presupune gestionarea expresiei și uneori a trăirii pentru a satisface rolul profesional [Hochschild, 1983]. Dacă AI stabilește un benchmark de răbdare nelimitată, organizațiile pot cere oamenilor să imite un standard imposibil. Alternativ, firmele pot automatiza tocmai interacțiunile repetitive și păstra oamenii pentru situații în care vulnerabilitatea, negocierea și responsabilitatea autentică contează.

Nu știm care echilibru va apărea. Este posibil să se dezvolte norme explicite care protejează dreptul omului la latență și limită: ore fără răspuns, refuzul abuzului, dreptul de a nu simula căldură permanentă. Paradoxal, tocmai AI-ul foarte politicos poate face necesară reafirmarea faptului că relațiile umane includ oboseală, conflict și indisponibilitate legitimă.

11.8 Transferul așteptărilor: riscul intoleranței la fricțiunea umană

Relația cu un agent personalizat poate fi construită în jurul preferințelor utilizatorului. Agentul nu are în mod necesar interese incompatibile și poate ajusta vocabularul, nivelul de detaliu și stilul. Relațiile umane sunt diferite: celălalt are obiective, limite și interpretări proprii. O ipoteză discutată frecvent este că interacțiunea cu agenți foarte complianți va reduce toleranța pentru această fricțiune.

Ipoteza este psihologic plauzibilă, dar evidența actuală este insuficientă pentru o afirmație tare. Utilizatorii disting de regulă oamenii de unelte și pot menține norme diferite pentru contexte diferite. Jocurile video nu au eliminat toleranța pentru regulile fizice ale lumii, iar căutarea instantanee nu a eliminat complet răbdarea pentru cercetare. Transferul de așteptări este posibil, nu automat.

Dacă apare, ar trebui să vedem semne comportamentale: creșterea abandonului conversațiilor umane când interlocutorul nu oferă validare rapidă, utilizarea mai frecventă a AI pentru a evita negocierile dificile sau preferința pentru relații mediate în care agentul filtrează conflictul. Studiile longitudinale ar trebui să compare aceste comportamente cu intensitatea și tipul folosirii AI, controlând trăsături preexistente precum anxietatea socială.

Există și scenariul opus. Un agent poate antrena toleranța la conflict dacă este proiectat să simuleze dezacord respectuos, să refuze confirmarea automată și să ofere perspective opuse. AI nu este condamnat la complianță; complianța este o alegere de design și de piață.

11.9 Auto-dezvăluirea accelerată și problema intimității ieftine

Apropierea umană se construiește în parte prin auto-dezvăluire graduală. Ea implică risc: informația poate fi folosită împotriva noastră, interlocutorul ne poate judeca, iar reciprocitatea nu este garantată. Un agent AI poate reduce o parte din acest risc perceput. Nu râde, nu părăsește camera și poate răspunde cu validare imediată.

Kleinert și colaboratorii arată experimental că niveluri mai mari de auto-dezvăluire în răspunsurile AI pot provoca auto-dezvăluire mai mare din partea participantului și senzație de apropiere [Kleinert et al.,

2026]. Cercetări privind folosirea afectivă a ChatGPT și companiilor sugerează că o parte dintre utilizatori îi folosesc deja pentru suport și conversații personale, însă prevalența și efectele cauzale sunt încă dificil de estimat [Phang et al., 2025; Chandra et al., 2025].

În viitor, costul scăzut al auto-dezvăluirii poate schimba ritmul intimității. O persoană poate ajunge foarte repede la subiecte pe care cu un om le-ar discuta după luni. Nu este clar dacă aceasta produce intimitate reală, exercițiu pentru intimitate sau o categorie separată de relație. Probabil pot exista toate trei.

Termenul „intimitate ieftină” trebuie folosit cu grijă. Ieftină se referă la costul de producție al răspunsului, nu la valoarea experienței pentru utilizator. Cineva poate obține un beneficiu autentic dintr-o conversație cu o entitate care nu simte. În același timp, disponibilitatea unei simulări fără risc poate reduce motivația de a investi în relații reciproce mai dificile. Această substituție este una dintre întrebările empirice centrale ale următorului deceniu.

11.10 AI ca traducător al emoției între oameni

Până acum am discutat agentul ca interlocutor. Un rol probabil la fel de important este intermedierea conversațiilor între oameni. AI poate reformula un mesaj agresiv, rezuma conflictul, traduce între limbi și stiluri, estima ce parte a unei formulări ar putea fi interpretată ca ostilă sau sugera răspunsuri.

Această funcție poate reduce costul cooperării între persoane cu competențe lingvistice și culturale diferite. Ea poate fi deosebit de utilă pentru oameni care au dificultăți în formulare, pentru situații profesionale sau pentru conversații în care emoția ridicată degradează claritatea. În acest scenariu, AI nu substituie relația; servește ca un strat pragmatic.

Dar medierea poate modifica distribuția responsabilității. Dacă agentul „îndulcește” automat fiecare mesaj, interlocutorii pot să nu mai observe intensitatea reală a dezacordului. Dacă rezumă, decide ce este relevant. Dacă propune interpretarea intenției celuilalt, poate consolida o lectură greșită. Sistemul devine participant la relație chiar dacă nu este adresat explicit.

Un fenomen deosebit de important ar fi „normalizarea tonului”. Dacă mesajele trec prin același tip de optimizare, variația idiosincronică se poate reduce. Am putea comunica mai corect și totodată mai puțin

distinctiv. Evidența actuală privind omogenizarea conținutului generativ face scenariul plauzibil, dar nu avem încă date longitudinale despre comunicarea interpersonală asistată.

11.11 Autenticitatea mesajului când intenția și formularea sunt separate

AI permite separarea mai radicală dintre intenție și formulare. O persoană poate intenționa sincer să-și ceară scuze, dar poate fi incapabilă să formuleze bine și folosește un model. O altă persoană poate să nu simtă regret, dar poate produce o scuză perfectă pentru a reduce sancțiunea. Problema exista și înainte; tehnologia îi reduce costul și îi crește calitatea.

Normele viitoare pot pune accent diferit pe trei componente: originea formulării, intenția persoanei și comportamentul ulterior. În multe cazuri, ultima componentă este cea mai informativă. O scuză scrisă complet de om, urmată de repetarea aceleiași conduite, poate fi mai puțin valoroasă decât o scuză asistată de AI urmată de reparație reală.

Totuși, originea poate conta în relațiile în care efortul este el însuși semnal. O scrisoare de dragoste poate fi evaluată nu doar după frumusețea propozițiilor, ci după faptul că persoana a investit timp în ele. Aici apare o posibilă normă de proveniență afectivă: nu interdicția AI, ci transparența atunci când procesul de formulare face parte din semnificație.

Nu putem prezice unde societatea va trasa granița. Normele pot varia între generații și culturi. Important este că AI face separarea observabilă și obligă comunitățile să decidă ce anume consideră autentic.

11.12 De-skilling social sau școală permanentă de comunicare?

Aceași tehnologie poate produce două traiectorii opuse. În prima, utilizatorul delegă mesajele dificile, interpretarea conflictului și reglarea emoțională. Competența neasistată se exersează mai rar și scade. În a doua, agentul oferă feedback, simulează interlocutori, explică efectele formulărilor și permite repetarea. Utilizatorul primește mai mult antrenament decât ar fi avut fără AI.

Diferența este în designul buclei. Un sistem orientat spre răspuns imediat maximizează substituția. Un sistem pedagogic poate cere utilizatorului să încerce primul, apoi să compare, să explice și să revizuiască. Cercetarea asupra gândirii critice și scaffolding-ului în muncă sugerează că structura utilizării contează [Lee et al., 2025; Farach et al., 2026]. Extinderea la competențe sociale este încă în mare măsură o ipoteză.

Putem imagina experimente relativ simple: participanții practică luni de zile conversații dificile cu un agent în condiție de ghostwriting sau coaching, apoi sunt evaluați în interacțiuni umane neasistate. Măsurăm nu doar politețea, ci ascultarea, repararea neînțelegerii, toleranța la dezacord și adaptarea la indicii neașteptate. Fără asemenea date, afirmațiile despre „atrofierea empatiei” sau „creșterea inteligenței emoționale” sunt premature.

11.13 Copilăria și învățarea normelor de reciprocitate

Cea mai mare incertitudine privește copiii crescuți cu agenți conversaționali persistenți. Adulții au deja un model al diferenței dintre persoană și instrument; copiii îl construiesc. Dacă un copil interacționează zilnic cu entități care răspund mereu, acceptă întreruperi, nu au propriile nevoi și pot fi resetate, este rezonabil să întrebăm dacă această experiență va influența așteptările lui sociale.

Nu avem încă evidență longitudinală suficientă pentru a răspunde. Orice afirmație tare ar fi speculativă. Putem însă identifica mecanisme potențiale. Copilul învață reciprocitatea prin întâlnirea cu limitele altuia. Învăță repararea după ce l-a rănit pe celălalt, răbdarea când celălalt vorbește și compromisul când dorințele sunt incompatibile. Un agent fără interese proprii nu reproduce automat aceste condiții.

Pe de altă parte, agentul poate fi proiectat să le simuleze pedagogic. Poate refuza întreruperea, poate explica efectul unei formulări, poate invita copilul să considere perspectiva altuia. Problema este că simularea nu suportă costul real al relației. Nu știm dacă exercițiul transferă suficient.

Din acest motiv, cea mai defensabilă poziție este una de cercetare și precauție, nu de interdicție generală. Ar trebui să măsurăm timpul de utilizare, tipul de agent, calitatea relațiilor umane și dezvoltarea competențelor sociale pe perioade lungi. Copiii reprezintă tocmai cazul

În care efectele culturale pot deveni dezvoltare psihologică înainte ca societatea să-și stabilizeze normele.

11.14 Contraselecția: revenirea deliberată la lent, incomod și nerevizuit

Orice schimbare puternică poate produce practici de contraselecție. Dacă comunicarea devine instantanee, optimizată și asistată, unele comunități pot valoriza răspunsul lent, conversația fără dispozitive, scrisul nerevizuit și prezența în care nu există un agent pregătit să repare fiecare ezitare.

Acest scenariu nu trebuie romantizat. Comunicarea „naturală” a inclus întotdeauna minciună, manipulare, inegalitate și neînțelegere. Al poate îmbunătăți accesibilitatea și reduce multe bariere. Întrebarea este dacă optimizarea completă elimină și informații valoroase despre persoană: felul în care caută cuvintele, capacitatea de a repara spontan, limitele răbdării și costul implicării.

Putem vedea apariția unor norme asemănătoare cu „live”, „unplugged” sau „handmade”: conversații declarate fără asistență, examene orale, ceremonii și contexte intime în care intermediarii AI sunt excluși. Astfel de spații ar funcționa ca instituții ale semnalului costisitor și ale reciprocității directe.

Cât de răspândite vor fi? Probabil limitat și diferențiat. Majoritatea oamenilor nu renunță la tehnologii utile, dar folosesc reguli de context. Fotografia nu a eliminat pictura, iar streaming-ul nu a eliminat concertul. Analogii istorice nu demonstrează rezultatul, dar fac plauzibilă coexistența între comunicarea optimizată și enclave ale comunicării neasistate.

11.15 O taxonomie a schimbărilor de ritm și gradul lor de plauzibilitate

Putem încheia prin separarea scenariilor. Creșterea frecvenței cu care oamenii consultă un agent înainte de a comunica este foarte plauzibilă deoarece comportamentul este deja tehnic ușor și observabil. Creșterea auto-dezvăluirii către AI este plauzibilă și are suport experimental și observațional timpuriu. Influența stilului AI asupra limbajului uman este plauzibilă, dar datele sugerează deocamdată efecte selective, nu o convergență totală.

Transferul unei norme de răspuns instantaneu de la AI la relațiile umane este posibil, dar insuficient demonstrat. De-skilling-ul social este posibil și testabil, dar nu trebuie prezentat ca rezultat stabilit. Modificarea dezvoltării copiilor este o ipoteză cu impact potențial mare și evidență încă slabă. Apariția unor noi ritualuri de comunicare „human-only” este o extrapolare culturală plauzibilă, nu o prognoză cantitativă.

În fine, schimbarea cea mai profundă poate să nu fie viteza în sine, ci faptul că viteza devine proiectabilă. AI poate răspunde imediat sau poate imita ezitarea; poate fi cald sau rece; poate crește conflictul sau îl poate încetini. Temporalitatea și expresia afectivă devin parametri de sistem. Din acel moment, întrebarea despre viitorul emoțiilor este și o întrebare despre cine proiectează acești parametri, pentru ce obiectiv și cu ce posibilitate de control din partea utilizatorului.

Capitolul 12. Spre o ecologie afectivă deliberată

12.1 De ce problema nu poate fi redusă la reziliența individului

Când o tehnologie produce anxietate, răspunsul instituțional obișnuit este formarea individului: adaptare, mindfulness, recalificare și igienă digitală. Aceste instrumente pot ajuta, dar devin ideologice dacă ascund sursa structurală. O persoană nu poate regla prin respirație pierderea arbitrară a venitului, supravegherea sau lipsa dreptului de contestare.

Emoțiile sunt informații despre relația dintre agent și mediu. Uneori evaluarea este eronată; alteori mediul este nedrept. O politică afectivă matură trebuie să poată schimba ambele părți. Altfel, „bunăstarea” devine adaptare la condiții pe care nimeni nu le poate negocia.

Conceptul de ecologie afectivă desemnează distribuția situațiilor, recompenselor, normelor și instituțiilor care fac anumite emoții probabile. Nu presupune că statul trebuie să administreze interiorul cetățenilor. Dimpotrivă, cere limitarea puterii actorilor care optimizează stări fără control democratic.

12.2 Agenția ca principiu de proiectare

Dovezile privind utilizarea pasivă a AI sugerează că performanța nu este singurul criteriu. Sistemele ar trebui evaluate după efectul asupra capacității utilizatorului de a înțelege, decide și acționa independent [Lee et al., 2026]. Agenția nu înseamnă executarea manuală a fiecărui pas. Înseamnă control semnificativ asupra scopului, posibilitatea de a interveni și cunoașterea limitelor.

Interfețele pot păstra agenția prin solicitarea unei contribuții inițiale, afișarea alternativelor, explicarea incertitudinii și facilitarea verificării. În sarcini de învățare, răspunsul complet instantaneu poate fi mai puțin util decât indicii graduale. În sarcini critice, sistemul trebuie să arate baza recomandării și condițiile în care nu este de încredere.

Există un compromis cu ușurința. O interfață care cere permanent verificare poate transfera costul și responsabilitatea fără putere reală. „Omul în buclă” nu este suficient dacă rolul este ceremonial. Agenția

trebuie măsurată prin capacitatea efectivă de a schimba rezultatul și prin resursele necesare.

12.3 Protejarea competenței fără conservarea ineficienței

Societățile au motive să păstreze competențe umane în domenii în care sistemele pot eșua, unde autonomia are valoare sau unde învățarea este obiectiv. Aceasta nu justifică menținerea oricărei proceduri manuale. Criteriul trebuie să fie valoarea capacității, nu nostalgia.

Educația poate delimita perioade de lucru fără AI pentru formarea modelelor mentale, urmate de colaborare asistată. Evaluarea poate separa competența independentă, competența de verificare și capacitatea de orchestrare. Un singur scor nu surprinde aceste roluri.

Profesiile pot defini sarcini care nu trebuie delegate integral din motive de responsabilitate. Medicii pot folosi recomandări, dar trebuie să poată explica și contesta. Cercetătorii pot folosi generare, dar trebuie să verifice sursele și să răspundă pentru afirmații. Regulile trebuie actualizate empiric; protecția jurisdicției nu este scop suficient.

Pentru persoane ale căror competențe sunt devalorizate, politicile de tranziție ar trebui să ofere mai mult decât cursuri. Sunt necesare timp, venit, roluri de mentorat și recunoașterea transferabilității. Doliul de competență nu dispare prin retorica „învățării continue”.

12.4 Dreptul la contribuție și dreptul de a nu produce

O societate cu muncă redusă trebuie să țină împreună două principii. Oamenii au nevoie de oportunități reale de contribuție și recunoaștere. În același timp, drepturile de bază și demnitatea nu pot depinde de productivitatea economică.

Dreptul la contribuție înseamnă acces la roluri în care acțiunea are consecințe și este recunoscută. Poate include îngrijire, cercetare cetățenească, conservarea mediului, educație reciprocă, artă și guvernanță locală. Nu este o obligație universală de serviciu.

Dreptul de a nu produce protejează odihna, dizabilitatea, explorarea și existența fără justificare permanentă. El nu elimină obligațiile sociale, dar împiedică transformarea resurselor de bază în recompense pentru conformare.

Tensiunea nu poate fi rezolvată printr-o formulă. Comunitățile vor disputa ce contribuții merită prestigiu și ce obligații sunt legitime. Important este ca refuzul unei ierarhii ocupaționale să nu creeze o nouă birocrație care măsoară fiecare gest comunitar.

12.5 Instituții ale apartenenței dincolo de firmă

Dacă locul de muncă pierde rolul de principal organizator al vieții, alte instituții trebuie să furnizeze contact repetat și obiective comune. Simplul acces la divertisment nu este suficient. Relațiile se formează prin proximitate, continuitate și interdependență.

Bibliotecile, laboratoarele civice, atelierele, cooperativele, asociațiile sportive și spațiile de îngrijire pot deveni infrastructură centrală. Ele nu trebuie tratate ca hobby-uri reziduale după „economia reală”. Finanțarea publică și guvernanta locală pot oferi stabilitate fără a transforma participarea în muncă salariată obligatorie.

Comunitățile pot fi și opresive. Apartenența intensă produce conformism, excludere și control. Este necesară posibilitatea de a părăsi grupul, pluralitatea instituțiilor și drepturi individuale. O lume post-muncă nu trebuie să înlocuiască disciplina corporativă cu disciplina comunității închise.

Al poate ajuta la coordonare și acces, dar nu ar trebui să devină singurul mediator. Întâlnirea cu persoane care nu sunt perfect potrivite și negocierile locale au valoare pentru capacitatea democratică. Personalizarea totală poate reduce experiența bunurilor comune.

12.6 Raritatea voluntară și zonele fără optimizare

Unele activități își păstrează valoarea prin limite. Educația, sportul și ritualul pot interzice anumite scurtături pentru a proteja procesul. Această raritate voluntară trebuie declarată și justificată. Altfel, devine gatekeeping.

O zonă fără optimizare este un context în care performanța nu este permanent măsurată și personalizată. Poate fi o conversație, un parc, un atelier sau o perioadă de învățare. Scopul nu este respingerea tehnologiei, ci protejarea explorării și a opacității.

Dreptul la experiențe neoptimizate ar putea deveni analog cu protecția vieții private. Nu avem încă un concept juridic stabil. El ar include posibilitatea de a dezactiva recomandarea, de a folosi ordine

cronologică, de a păstra date locale și de a nu transforma fiecare reacție în feedback pentru un model.

Această propunere este normativă și plauzibilă tehnic, dar adoptarea ei este incertă. Modelele de afaceri bazate pe date au stimulente contrare. Ar necesita reglementare, standarde și alternative publice sau cooperative.

12.7 Guvernanța relațiilor cu AI

Produsele care invită atașamentul ar trebui să aibă obligații suplimentare. Prima este transparența asupra naturii sistemului și a rolului furnizorului. A doua este continuitatea responsabilă: preaviz pentru schimbări majore, opțiuni de export și închidere. A treia este evitarea mecanismelor de exclusivitate și culpabilizare.

Măsurarea retenției nu ar trebui să fie singurul criteriu pentru companiile afective. Sunt necesari indicatori privind autonomia, diversitatea contactelor, capacitatea de a întrerupe utilizarea și orientarea către sprijin uman când este adecvat. Niciun indicator nu este perfect, dar absența lor lasă optimizarea captivă unui proxy comercial.

Auditarea trebuie să includă conversații lungi și utilizatori vulnerabili. Testele scurte de politețe nu surprind dependența, escaladarea sau discontinuitatea. Cercetătorii au nevoie de acces protejat la date, iar utilizatorii de control asupra folosirii conversațiilor intime.

Regulile nu trebuie să presupună că atașamentul este patologic. Obiectivul este reducerea exploatarei și păstrarea ieșirii. Relațiile pot fi benefice fără a fi lăsate complet la discreția furnizorului.

12.8 Proveniență și norme ale expresiei asistate

Pe măsură ce AI produce mesaje emoționale, societatea va trebui să decidă când proveniența contează. Într-o notificare administrativă, formularea asistată poate fi irelevantă. Într-o mărturie, o evaluare personală sau o declarație de angajament, poate fi centrală.

O regulă generală de etichetare a fiecărei propoziții ar fi greu de aplicat și poate produce zgomot. Mai utilă este obligația de dezvăluire când asistența modifică semnificativ atribuirea de competență, intenție sau responsabilitate. Instituțiile pot defini praguri după context.

Proveniența tehnică nu rezolvă autenticitatea. O persoană poate scrie singură o scuză manipulative și poate folosi AI pentru una sincer asumată. Totuși, cunoașterea procesului este relevantă când audiența decide meritul sau intimitatea.

Normele trebuie să evite o castă a „purității umane” care ascunde colaborarea și o cultură a indiferenței în care orice revendicare de autorat este acceptată. Trasabilitatea proporțională și responsabilitatea sunt o cale intermediară.

12.9 Protecția împotriva optimizării afective ascunse

Persuasiunea nu este nouă. Publicitatea, ritualul și politica au modelat emoții. AI schimbă adaptarea și viteza. Un sistem poate identifica momentul de vulnerabilitate și formula mesajul potrivit pentru persoana concretă.

Reglementarea ar trebui să distingă asistența cerută de manipularea ascunsă. Utilizatorul poate solicita un exercițiu care îl motivează. Este diferit ca platforma să inducă anxietate pentru a crește achiziția. Consimțământul trebuie să fie specific și posibil de retras.

Unele practici ar putea fi interzise: simularea suferinței agentului pentru a împiedica plecarea, folosirea datelor intime pentru presiune comercială, personalizarea politică opacă sau exploatarea minorilor. Altele necesită evaluare de risc și audit.

Este dificil de demonstrat intenția manipulatorie. De aceea, regulile trebuie să privească funcțiile de obiectiv, experimentele, datele și rezultatele, nu doar textul unei conversații. Guvernanța afectivă este în parte guvernanță a metricilor.

12.10 Un program de cercetare pentru emoțiile viitorului

Prima prioritate este construirea unor cohorte longitudinale. Utilizatorii ar trebui urmăriți înainte și după schimbări de muncă, adopția AI și formarea relațiilor cu agenți. Măsurile trebuie să includă venit, calitatea muncii, contact social, identitate, agenție și sănătate mintală. Fără design longitudinal, selecția va rămâne o confuzie majoră.

A doua prioritate este comparația culturală. Termenii propuși în această carte provin dintr-un mediu intelectual european și nord-american. Comunități cu alte norme despre muncă, familie și

tehnologie pot organiza experiența diferit. Cercetarea trebuie realizată în limbi locale și să permită apariția unor categorii neanticipate.

A treia este analiza mecanismelor. Experimentele pot varia contribuția umană, transparența, controlul și atribuirea meritului pentru a testa rușinea delegării și proprietatea psihologică. Actualizările simulate pot testa doliul de versiune în condiții etice. Sistemele de recomandare pot fi comparate după efectul asupra diversității afective și autonomiei.

A patra este dezvoltarea măsurilor care nu depind exclusiv de auto-raportare. Comportamentul de evitare, dezvăluirea, persistența, alegerea dificultății și relațiile efective oferă date suplimentare. Măsurile fiziologice pot ajuta, dar nu trebuie tratate drept detector de adevăr emoțional.

A cincea este evaluarea intervențiilor instituționale. Consultarea lucrătorilor, modurile colaborative, portabilitatea memoriei și zonele fără recomandare sunt ipoteze, nu soluții demonstrate. Ele trebuie testate pentru efecte și distribuție.

12.11 Limitele proiectării

Nu putem proiecta o societate fără frică, rușine sau invidie și nici nu este clar că ar fi dezirabil. Aceste emoții pot informa despre pericol, devalorizare și diferențe. Obiectivul este reducerea situațiilor inutile și a exploatării, nu eliminarea semnalului.

Nu putem anticipa toate categoriile care vor apărea. Cultura este creativă, iar tehnologia poate evolua în direcții neașteptate. Orice constituție afectivă trebuie să poată fi revizuită.

Nu există neutralitate. A nu reglementa lasă proiectarea în mâna pieței și a infrastructurii existente. A reglementa poate concentra puterea și poate îngheța norme premature. Răspunsul este guvernanță contestabilă, cercetare deschisă și pluralitate de instituții, nu un minister al emoțiilor.

În cele din urmă, distribuția materială rămâne fundamentală. O societate nu poate vindeca prin vocabular afectiv exproprierea și insecuritatea. Dar nici redistribuirea nu va organiza singură sensul, apartenența și competența. Cele două niveluri trebuie tratate împreună.

Concluzie. Continuitatea biologică și discontinuitatea lumii

Emoțiile umane nu sunt reziduuri iraționale ale unei minți altfel calculatoare. Ele coordonează atenția, prioritățile, corpul și acțiunea în situații relevante. În interacțiune, semnalează evaluări și modifică comportamentul altora. În grup, contribuie la statut, cooperare, sancțiune, reparație și apartenență. Aceste funcții sunt plauzibile și, pentru multe emoții, susținute de dovezi convergente; istoriile evolutive precise rămân însă dificil de demonstrat.

Nici imaginea opusă, potrivit căreia cultura inventează liber emoțiile, nu este suficientă. Oamenii împărtășesc vulnerabilități și sisteme de reglare vechi. Cultura nu lucrează pe o pagină goală. Ea selectează situații, concepte, practici și norme în interiorul unui spațiu constrâns biologic.

Întrebarea despre emoții noi are astfel un răspuns diferențiat. Apariția în câteva decenii a unor mecanisme biologice afective complet noi este foarte puțin plauzibilă. Apariția unor obiecte noi pentru emoții vechi este aproape sigură. Apariția unor configurații recurente și a unor categorii culturale noi este plauzibilă și, în unele cazuri, deja observabilă. Cât de mult aceste categorii schimbă fenomenologia, nu doar limbajul, rămâne deschis.

Inteligența artificială este relevantă nu doar fiindcă poate prelua sarcini, ci fiindcă separă componente care anterior erau mai strâns asociate. Rezultatul se poate separa de efort, competența afișată de competența internă, expresia empatică de experiență, memoria relațională de continuitatea unui agent și decizia de responsabilitatea resimțită. Aceste separări creează probleme de atribuire și noi obiecte pentru rușine, mândrie, doliu, anxietate și încredere.

Atlasul propus nu prezice o listă de emoții inevitabile. Doliul de competență, rușinea delegării, vertijul abundenței, claustrofobia predictivă sau doliul de versiune sunt ipoteze despre configurații. Unele au indicii actuale; altele sunt extrapolări. Niciuna nu necesită deocamdată postularea unui nou circuit biologic. Valoarea termenilor va depinde de cercetarea viitoare și de capacitatea lor de a explica mai mult decât conceptele existente.

O lume în care AI realizează o parte mare a muncii nu este o consecință sigură a stadiului actual al tehnologiei. Munca poate fi augmentată, polarizată, redusă sau reorganizată în forme extractive. Instituțiile de proprietate și distribuție vor conta la fel de mult ca performanța modelelor. Orice analiză care vorbește despre „post-muncă” fără politică este incompletă.

Dacă timpul de muncă se reduce fără insecuritate, societatea va trebui să transfere funcțiile afective ale ocupației către alte instituții. Oamenii au nevoie de autonomie, competență, relații și contribuție recunoscută, dar nu trebuie obligați să își câștige demnitatea prin output. Această tensiune nu poate fi rezolvată prin venit singur și nici prin moralizarea activității comunitare.

Agenții conversaționali arată deja că mediul social se poate extinde către entități a căror reciprocitate este greu de evaluat. Beneficiile și riscurile sunt mixte. Atașamentul nu trebuie ridiculizat, iar simularea relației nu trebuie lăsată fără reguli. Continuitatea, portabilitatea, protecția datelor și evitarea exclusivității manipulative sunt probleme de proiectare actuale, nu doar speculații despre conștiință artificială.

Conceptul cel mai general al cărții este cel al celor trei viteze. Biologia se schimbă lent și continuă să furnizeze constrângeri. Dezvoltarea și cultura pot reorganiza într-o generație declanșatoarele, conceptele și normele. Selecția algoritmică poate modifica în ore distribuția mesajelor și recompenselor. Acest al treilea ritm nu este autonom; este controlat de organizații și alimentat de predispoziții umane. Tocmai de aceea poate fi guvernat.

Nu este realist să cerem o lume fără inginerie afectivă. Școala, religia, piața, arta și dreptul au cultivat întotdeauna stări. Noutatea constă în personalizare, măsurare și optimizare continuă. Răspunsul nu este interzicerea emoției, ci transparența obiectivelor, dreptul de ieșire, protecția opacității și instituții care nu reduc valoarea umană la indicatori.

Cea mai prudentă concluzie este că oamenii viitorului pot avea aproximativ aceeași arhitectură afectivă de bază și totuși pot trăi într-o lume emoțională parțial străină nouă. Diferența va proveni din ceea ce pierd, din ceea ce consideră merit, din agenții de care se atașează și din instituțiile care le distribuie recunoașterea. Nu știm încă forma exactă. Știm suficient pentru a nu lăsa schimbarea exclusiv pe seama optimizării comerciale și a adaptării individuale.

Extinderea behavioristă schimbă și criteriul după care vom recunoaște o transformare afectivă reală. Nu este suficient ca oamenii să inventeze un cuvânt nou sau să raporteze o senzație familiară într-un chestionar. O schimbare culturală importantă ar trebui să fie vizibilă în comportamente stabile: regimuri de delegare, reguli de proveniență, ritualuri de comunicare fără asistență, noi forme de statut, practici de conservare a competenței sau moduri repetate de a consulta un agent înainte de acțiune.

În același timp, interacțiunea cu agenți care simulează emoții poate modifica nu doar conținutul comunicării, ci și parametrii ei temporali. Răspunsul instantaneu, răbdarea nelimitată și auto-dezvăluirea generată pot deveni referințe cu care oamenii compară interlocutorii umani. Dovezile actuale arată efecte locale asupra apropierii, stilului, satisfacției și comportamentului de echipă, dar nu justifică încă afirmația că normele umane au fost recalibrate la scară socială. Aceasta rămâne una dintre ipotezele centrale de urmărit longitudinal.

Concluzia generală rămâne astfel mai modestă decât o profeție despre „omul post-AI”. Este foarte probabil să vedem o selecție rapidă a strategiilor și normelor, plauzibil să vedem noi categorii afective și noi convenții temporale ale comunicării, și mult mai puțin plauzibil să vedem pe termen scurt mecanisme emoționale biologice complet noi. Ceea ce se poate schimba repede este relația dintre mecanisme vechi, situații noi și comportamentele prin care oamenii încearcă să-și păstreze agenția, apartenența și statutul.

Anexă. Matricea sintetică a scenariilor

Matricea nu oferă probabilități numerice. Ea rezumă gradul de sprijin actual și tipul de schimbare implicat. Calificativele trebuie revizuite pe măsură ce apar date longitudinale și comparative.

Scenariu	Statut actual și implicație afectivă
Augmentare largă a muncii	Plausibilă și deja observabilă, dar efectele sunt heterogene. Poate crește performanța și, dacă utilizarea este pasivă, poate reduce agenția, ownership-ul și sensul.
Polarizare ocupațională	Plausibilitate moderată spre ridicată, dependentă de sector și instituții. Ar intensifica anxietatea relevanței, comparația și conflictul distributiv.
Reducerea timpului de muncă cu distribuție sigură	Tehnic posibilă, politic incertă. Ar cere instituții alternative pentru contribuție, statut, apartenență și structură temporală.
Automatizare extractivă cu insecuritate	Plausibilă în absența negocierii și protecției sociale. Ar amplifica frica, furia și rușinea fără a rezolva funcțiile afective ale muncii.
Companion AI ca instrument complementar	Susținut ca posibilitate, nu ca rezultat universal. Poate oferi acces, reflecție și antrenament dacă extinde, nu înlocuiește, legăturile umane.
Companion AI ca substitut principal	Risc plauzibil pentru utilizatori vulnerabili, dar cauzalitatea și prevalența sunt încă neclare. Intensitatea și rețeaua socială offline sunt variabile centrale.
Selecție algoritmică a expresiei afective	Foarte plauzibilă și deja documentată pe platforme. Algoritmii nu inventează emoția, dar pot schimba rapid ce expresii sunt recompensate și imitate.
Categorie culturală emoțională nouă	Plausibilitate moderată. Ar necesita recurență, vocabular, norme, recunoaștere intersubiectivă și valoare explicativă peste termenii existenți.
Mecanism biologic emoțional nou în câteva decenii	Foarte puțin plauzibil. Nu există un mecanism credibil prin care schimbarea tehnologică apropiată să producă atât de repede un sistem afectiv genetic distinct.
Delegare maximală și de-skilling	Plausibilitate moderată și puternic dependentă de design. Poate crește performanța asistată și simultan reduce

AUTOPSIA EMOȚIILOR VIITOARE

	competența neasistată; sunt necesare date longitudinale.
Migrarea rolului spre verificare și arbitraj	Plausibilitate ridicată în domenii unde generarea devine ieftină, dar eroarea rămâne costisitoare. Necesită competență de audit și instituții care o recompensează.
Proveniență umană ca semnal	Plausibilitate ridicată în piețe simbolice, educație și decizii cu responsabilitate; probabil redusă pentru produse pur utilitare unde originea nu contează.
Comunicare afectivă accelerată de AI	Plausibilă la nivelul frecvenței consultării și auto-dezvăluirii; efectul asupra normelor umane de răspuns și intimitate rămâne insuficient demonstrat.
Spillover al stilului AI către interacțiuni uman-uman	Există indicii experimentale recente asupra limbajului și cogniției de echipă, dar literatura este tânără. Necesită replicări și studii longitudinale.
Ritualuri și spații „human-only”	Extrapolare culturală moderat plauzibilă. Analogii istorice există, dar amploarea depinde de valoarea simbolică a contribuției și de costul renunțării la AI.

Referințe

- Adolphs, R., & Anderson, D. J. (2018). *The Neuroscience of Emotion: A New Synthesis*. Princeton University Press.
- Albrecht, G., Sartore, G.-M., Connor, L., Higginbotham, N., Freeman, S., Kelly, B., Stain, H., Tonna, A., & Pollard, G. (2007). Solastalgia: The distress caused by environmental change. *Australasian Psychiatry*, 15(S1), S95–S98. <https://doi.org/10.1080/10398560701701288>
- Autor, D. H. (2015). Why are there still so many jobs? The history and future of workplace automation. *Journal of Economic Perspectives*, 29(3), 3–30. <https://doi.org/10.1257/jep.29.3.3>
- Barrett, L. F. (2017). *How Emotions Are Made: The Secret Life of the Brain*. Houghton Mifflin Harcourt.
- Barrett, L. F., Adolphs, R., Marsella, S., Martinez, A. M., & Pollak, S. D. (2019). Emotional expressions reconsidered: Challenges to inferring emotion from human facial movements. *Psychological Science in the Public Interest*, 20(1), 1–68. <https://doi.org/10.1177/1529100619832930>
- Bellaiche, L., Shahi, R., Turpin, M. H., Ragnhildstveit, A., Sprockett, S., Barr, N., Christensen, A., & Seli, P. (2023). Humans versus AI: Whether and why we prefer human-created compared to AI-created artwork. *Cognitive Research: Principles and Implications*, 8, 42. <https://doi.org/10.1186/s41235-023-00499-6>
- Boehm, C. (1999). *Hierarchy in the Forest: The Evolution of Egalitarian Behavior*. Harvard University Press.
- Boyd, R., & Richerson, P. J. (1985). *Culture and the Evolutionary Process*. University of Chicago Press.
- Brady, W. J., McLoughlin, K. L., Doan, T. N., & Crockett, M. J. (2021). How social learning amplifies moral outrage expression in online social networks. *Science Advances*, 7(33), eabe5641. <https://doi.org/10.1126/sciadv.abe5641>
- Brady, W. J., Wills, J. A., Jost, J. T., Tucker, J. A., & Van Bavel, J. J. (2017). Emotion shapes the diffusion of moralized content in social networks. *Proceedings of the National Academy of Sciences*, 114(28), 7313–7318. <https://doi.org/10.1073/pnas.1618923114>
- Brynjolfsson, E., Li, D., & Raymond, L. R. (2025). Generative AI at work. *Quarterly Journal of Economics*, 140(2), 889–942. <https://doi.org/10.1093/qje/qjae044>
- Cacioppo, J. T., Hughes, M. E., Waite, L. J., Hawkley, L. C., & Thisted, R. A. (2006). Loneliness as a specific risk factor for depressive symptoms: Cross-sectional and longitudinal analyses. *Psychology and Aging*, 21(1), 140–151. <https://doi.org/10.1037/0882-7974.21.1.140>
- Castelo, N., Bos, M. W., & Lehmann, D. R. (2019). Task-dependent algorithm aversion. *Journal of Marketing Research*, 56(5), 809–825. <https://doi.org/10.1177/0022243719851788>

- Chamberlain, R., Mullin, C., Scheerlinck, B., & Wagemans, J. (2018). Putting the art in artificial: Aesthetic responses to computer-generated art. *Psychology of Aesthetics, Creativity, and the Arts*, 12(2), 177–192. <https://doi.org/10.1037/aca0000136>
- Chandra, M., Hernandez, J., Ramos, G., Ershadi, M., Bhattacharjee, A., Amores, J., Okoli, E., Paradiso, A., Warreth, S., & Suh, J. (2025). Longitudinal study on social and emotional use of AI conversational agent. *arXiv:2504.14112*. <https://doi.org/10.48550/arXiv.2504.14112>
- Cheng, J. T., Tracy, J. L., & Henrich, J. (2010). Pride, personality, and the evolutionary foundations of human social status. *Evolution and Human Behavior*, 31(5), 334–347. <https://doi.org/10.1016/j.evolhumbehav.2010.02.004>
- Creanza, N., Kolodny, O., & Feldman, M. W. (2017). Cultural evolutionary theory: How culture evolves and why it matters. *Proceedings of the National Academy of Sciences*, 114(30), 7782–7789. <https://doi.org/10.1073/pnas.1620732114>
- Curtis, V., de Barra, M., & Aunger, R. (2011). Disgust as an adaptive system for disease avoidance behaviour. *Philosophical Transactions of the Royal Society B*, 366(1563), 389–401. <https://doi.org/10.1098/rstb.2010.0117>
- Damasio, A. R. (1999). *The Feeling of What Happens: Body and Emotion in the Making of Consciousness*. Harcourt.
- Darwin, C. (1872). *The Expression of the Emotions in Man and Animals*. John Murray.
- De Freitas, J., Oğuz-Uğuralp, Z., Uğuralp, A. K., & Puntoni, S. (2024). Lessons from an app update at Replika AI: Identity discontinuity in human–AI relationships. *arXiv:2412.14190*. <https://doi.org/10.48550/arXiv.2412.14190>
- Dehaene, S., & Cohen, L. (2007). Cultural recycling of cortical maps. *Neuron*, 56(2), 384–398. <https://doi.org/10.1016/j.neuron.2007.10.004>
- Dietvorst, B. J., Simmons, J. P., & Massey, C. (2015). Algorithm aversion: People erroneously avoid algorithms after seeing them err. *Journal of Experimental Psychology: General*, 144(1), 114–126. <https://doi.org/10.1037/xge0000033>
- Doshi, A. R., & Hauser, O. P. (2024). Generative AI enhances individual creativity but reduces the collective diversity of novel content. *Science Advances*, 10(28), eadn5290. <https://doi.org/10.1126/sciadv.adn5290>
- Ekman, P. (1992). An argument for basic emotions. *Cognition and Emotion*, 6(3–4), 169–200. <https://doi.org/10.1080/02699939208411068>
- Epley, N., Waytz, A., & Cacioppo, J. T. (2007). On seeing human: A three-factor theory of anthropomorphism. *Psychological Review*, 114(4), 864–886. <https://doi.org/10.1037/0033-295X.114.4.864>
- Fiske, A. P., Schubert, T. W., & Seibt, B. (2025). Seeking communal emotions in social practices that culturally evolved to evoke emotions: Worship, kitten videos, memorials, narratives of love, and more. *Annual Review of Psychology*, 76, 607–633. <https://doi.org/10.1146/annurev-psych-020124-023338>

- Gendron, M., Roberson, D., van der Vyver, J. M., & Barrett, L. F. (2014). Perceptions of emotion from facial expressions are not culturally universal: Evidence from a remote culture. *Emotion*, 14(2), 251–262. <https://doi.org/10.1037/a0036052>
- Gmyrek, P., Berg, J., Kamiński, K., Konopczyński, F., Ładna, A., Nafradi, B., Rosłaniec, K., & Troszyński, M. (2025). Generative AI and Jobs: A Refined Global Index of Occupational Exposure. ILO Working Paper 140. International Labour Organization. <https://doi.org/10.54394/HETP0387>
- Graham, J., Haidt, J., Koleva, S., Motyl, M., Iyer, R., Wojcik, S. P., & Ditto, P. H. (2013). Moral foundations theory: The pragmatic validity of moral pluralism. *Advances in Experimental Social Psychology*, 47, 55–130. <https://doi.org/10.1016/B978-0-12-407236-7.00002-4>
- Hacking, I. (1995). The looping effects of human kinds. In D. Sperber, D. Premack, & A. J. Premack (Eds.), *Causal Cognition: A Multidisciplinary Debate* (pp. 351–383). Clarendon Press.
- Haidt, J. (2001). The emotional dog and its rational tail: A social intuitionist approach to moral judgment. *Psychological Review*, 108(4), 814–834. <https://doi.org/10.1037/0033-295X.108.4.814>
- Hamilton, K., Hou, I., Patel, D., Nnam, S., Patel, H., & MacNeil, S. (2026). “Stuck in a spiral”: Shame and guilt as social regulators of AI use in computing education. *arXiv:2606.14920*. <https://doi.org/10.48550/arXiv.2606.14920>
- Henrich, J. (2016). *The Secret of Our Success: How Culture Is Driving Human Evolution, Domesticating Our Species, and Making Us Smarter*. Princeton University Press.
- Henrich, J., & Gil-White, F. J. (2001). The evolution of prestige: Freely conferred deference as a mechanism for enhancing the benefits of cultural transmission. *Evolution and Human Behavior*, 22(3), 165–196. [https://doi.org/10.1016/S1090-5138\(00\)00071-4](https://doi.org/10.1016/S1090-5138(00)00071-4)
- Hochschild, A. R. (1983). *The Managed Heart: Commercialization of Human Feeling*. University of California Press.
- Hoemann, K., Xu, F., & Barrett, L. F. (2019). Emotion words, emotion concepts, and emotional development in children: A constructionist hypothesis. *Developmental Psychology*, 55(9), 1830–1849. <https://doi.org/10.1037/dev0000686>
- Inzlicht, M., Shenav, A., & Olivola, C. Y. (2018). The effort paradox: Effort is both costly and valued. *Trends in Cognitive Sciences*, 22(4), 337–349. <https://doi.org/10.1016/j.tics.2018.01.007>
- Jackson, J. C., Watts, J., Henry, T. R., List, J.-M., Forkel, R., Mucha, P. J., Greenhill, S. J., Gray, R. D., & Lindquist, K. A. (2019). Emotion semantics show both cultural variation and universal structure. *Science*, 366(6472), 1517–1522. <https://doi.org/10.1126/science.aaw8160>
- Jahoda, M. (1982). *Employment and Unemployment: A Social-Psychological Analysis*. Cambridge University Press.

- Kashdan, T. B., Barrett, L. F., & McKnight, P. E. (2015). Unpacking emotion differentiation: Transforming unpleasant experience by perceiving distinctions in negativity. *Current Directions in Psychological Science*, 24(1), 10–16. <https://doi.org/10.1177/0963721414550708>
- Keltner, D., & Haidt, J. (1999). Social functions of emotions at four levels of analysis. *Cognition and Emotion*, 13(5), 505–521. <https://doi.org/10.1080/026999399379168>
- Keltner, D., Sauter, D., Tracy, J., & Cowen, A. (2019). Emotional expression: Advances in basic emotion theory. *Journal of Nonverbal Behavior*, 43, 133–160. <https://doi.org/10.1007/s10919-019-00293-3>
- Kleinert, T., Waldschütz, M., Blau, J., Heinrichs, M., & Schiller, B. (2026). AI outperforms humans in establishing interpersonal closeness in emotionally engaging interactions, but only when labelled as human. *Communications Psychology*, 4, 23. <https://doi.org/10.1038/s44271-025-00391-7>
- Kramer, A. D. I., Guillory, J. E., & Hancock, J. T. (2014). Experimental evidence of massive-scale emotional contagion through social networks. *Proceedings of the National Academy of Sciences*, 111(24), 8788–8790. <https://doi.org/10.1073/pnas.1320040111>
- Laland, K. N., Odling-Smee, J., & Myles, S. (2010). How culture shaped the human genome: Bringing genetics and the human sciences together. *Nature Reviews Genetics*, 11, 137–148. <https://doi.org/10.1038/nrg2734>
- Landy, J. F., & Goodwin, G. P. (2015). Does incidental disgust amplify moral judgment? A meta-analytic review of experimental evidence. *Perspectives on Psychological Science*, 10(4), 518–536. <https://doi.org/10.1177/1745691615583128>
- LeDoux, J. E. (2012). Rethinking the emotional brain. *Neuron*, 73(4), 653–676. <https://doi.org/10.1016/j.neuron.2012.02.004>
- Lee, E. H., Yin, Y., Jia, N., & Wakslak, C. J. (2026). Relying on AI at work reduces self-efficacy, ownership, and meaning while active collaboration mitigates the effects. *Scientific Reports*, 16, 13583. <https://doi.org/10.1038/s41598-026-42312-6>
- Lindquist, K. A., Satpute, A. B., & Gendron, M. (2015). Does language do more than communicate emotion? Current directions in the study of language and emotion. *Psychological Inquiry*, 26(1), 1–13. <https://doi.org/10.1080/1047840X.2015.988486>
- Lindquist, K. A., Wager, T. D., Kober, H., Bliss-Moreau, E., & Barrett, L. F. (2012). The brain basis of emotion: A meta-analytic review. *Behavioral and Brain Sciences*, 35(3), 121–143. <https://doi.org/10.1017/S0140525X11000446>
- McCullough, M. E., Kilpatrick, S. D., Emmons, R. A., & Larson, D. B. (2001). Is gratitude a moral affect? *Psychological Bulletin*, 127(2), 249–266. <https://doi.org/10.1037/0033-2909.127.2.249>
- Mikalonytė, E. S., & Kneer, M. (2022). Can artificial intelligence make art? Folk intuitions as to whether AI-driven robots can be viewed as artists and produce

- art. ACM Transactions on Human-Robot Interaction, 11(4), Article 43.
<https://doi.org/10.1145/3530875>
- Nesse, R. M. (2005). Natural selection and the regulation of defenses: A signal detection analysis of the smoke detector principle. *Evolution and Human Behavior*, 26(1), 88–105. <https://doi.org/10.1016/j.evolhumbehav.2004.08.002>
- Nesse, R. M. (2023). Evolutionary psychiatry: Foundations, progress and challenges. *World Psychiatry*, 22(2), 177–202. <https://doi.org/10.1002/wps.21072>
- Noy, S., & Zhang, W. (2023). Experimental evidence on the productivity effects of generative artificial intelligence. *Science*, 381(6654), 187–192.
<https://doi.org/10.1126/science.adh2586>
- OECD. (2023). *OECD Employment Outlook 2023: Artificial Intelligence and the Labour Market*. OECD Publishing. <https://doi.org/10.1787/08785bba-en>
- OECD. (2025). *Generative AI and the Future of Work: Perceptions and Prospects*. OECD/GPAI Global Dialogue on AI and the Future of Work.
<https://www.oecd.org/en/about/programmes/global-partnership-on-artificial-intelligence/generative-ai-and-the-future-of-work.html>
- Panksepp, J. (1998). *Affective Neuroscience: The Foundations of Human and Animal Emotions*. Oxford University Press.
- Parent-Rocheleau, X., & Parker, S. K. (2022). Algorithms as work designers: How algorithmic management influences the design of jobs. *Human Resource Management Review*, 32(3), 100838. <https://doi.org/10.1016/j.hrmr.2021.100838>
- Paul, K. I., & Moser, K. (2009). Unemployment impairs mental health: Meta-analyses. *Journal of Vocational Behavior*, 74(3), 264–282.
<https://doi.org/10.1016/j.jvb.2009.01.001>
- Phang, J., Lampe, M., Ahmad, L., Agarwal, S., Fang, C. M., Liu, A. R., Danry, V., Lee, E., Chan, S. W. T., Pataranutaporn, P., & Maes, P. (2025). Investigating affective use and emotional well-being on ChatGPT. *arXiv:2504.03888*.
<https://doi.org/10.48550/arXiv.2504.03888>
- Pierce, J. L., Kostova, T., & Dirks, K. T. (2003). The state of psychological ownership: Integrating and extending a century of research. *Review of General Psychology*, 7(1), 84–107. <https://doi.org/10.1037/1089-2680.7.1.84>
- Poonsiriwong, R., Archiwanguprok, C., & Pataranutaporn, P. (2026). “Death” of a chatbot: Investigating and designing toward psychologically safe endings for human–AI relationships. *arXiv:2602.07193*.
<https://doi.org/10.48550/arXiv.2602.07193>
- Przybylski, A. K., Murayama, K., DeHaan, C. R., & Gladwell, V. (2013). Motivational, emotional, and behavioral correlates of fear of missing out. *Computers in Human Behavior*, 29(4), 1841–1848. <https://doi.org/10.1016/j.chb.2013.02.014>
- Ravid, D. M., Tomczak, D. L., White, J. C., & Behrend, T. S. (2020). EPM 20/20: A review, framework, and research agenda for electronic performance monitoring. *Journal of Management*, 46(1), 100–126.
<https://doi.org/10.1177/0149206319869435>

- Reddy, W. M. (2001). *The Navigation of Feeling: A Framework for the History of Emotions*. Cambridge University Press.
- Rosenwein, B. H. (2006). *Emotional Communities in the Early Middle Ages*. Cornell University Press.
- Ryan, R. M., & Deci, E. L. (2017). *Self-Determination Theory: Basic Psychological Needs in Motivation, Development, and Wellness*. Guilford Press.
- Schaller, M., & Park, J. H. (2011). The behavioral immune system and why it matters. *Current Directions in Psychological Science*, 20(2), 99–103. <https://doi.org/10.1177/0963721411402596>
- Scheer, M. (2012). Are emotions a kind of practice (and is that what makes them have a history)? A Bourdieuan approach to understanding emotion. *History and Theory*, 51(2), 193–220. <https://doi.org/10.1111/j.1468-2303.2012.00621.x>
- Scheibehenne, B., Greifeneder, R., & Todd, P. M. (2010). Can there ever be too many options? A meta-analytic review of choice overload. *Journal of Consumer Research*, 37(3), 409–425. <https://doi.org/10.1086/651235>
- Scherer, K. R. (2009). The dynamic architecture of emotion: Evidence for the component process model. *Cognition and Emotion*, 23(7), 1307–1351. <https://doi.org/10.1080/02699930902928969>
- Selenko, E., Batinic, B., & Paul, K. I. (2011). Does latent deprivation lead to psychological distress? Investigating Jahoda's model in a four-wave study. *Journal of Occupational and Organizational Psychology*, 84(4), 723–740. <https://doi.org/10.1348/096317910X519360>
- Sell, A., Tooby, J., & Cosmides, L. (2009). Formidability and the logic of human anger. *Proceedings of the National Academy of Sciences*, 106(35), 15073–15078. <https://doi.org/10.1073/pnas.0904312106>
- Siegel, E. H., Sands, M. K., Van den Noortgate, W., Condon, P., Chang, Y., Dy, J., Quigley, K. S., & Barrett, L. F. (2018). Emotion fingerprints or emotion populations? A meta-analytic investigation of autonomic features of emotion categories. *Psychological Bulletin*, 144(4), 343–393. <https://doi.org/10.1037/bul0000128>
- Stroebe, M., Schut, H., & Boerner, K. (2017). Cautioning health-care professionals: Bereaved persons are misguided through the stages of grief. *OMEGA—Journal of Death and Dying*, 74(4), 455–473. <https://doi.org/10.1177/0030222817691870>
- Sznycer, D., Al-Shawaf, L., Bereby-Meyer, Y., Curry, O. S., De Smet, D., Ermer, E., Kim, S., Kim, S., Li, N. P., Lopez Seal, M. F., McClung, J., O, J., Ohtsubo, Y., Quillien, T., Schaub, M., Sell, A., van Leeuwen, F., Cosmides, L., & Tooby, J. (2018). Cross-cultural invariances in the architecture of shame. *Proceedings of the National Academy of Sciences*, 115(39), 9702–9707. <https://doi.org/10.1073/pnas.1805016115>
- Sznycer, D., Tooby, J., Cosmides, L., Porat, R., Shalvi, S., & Halperin, E. (2016). Shame closely tracks the threat of devaluation by others, even across cultures.

- Proceedings of the National Academy of Sciences, 113(10), 2625–2630.
<https://doi.org/10.1073/pnas.1514699113>
- Tangney, J. P., Stuewig, J., & Mashek, D. J. (2007). Moral emotions and moral behavior. *Annual Review of Psychology*, 58, 345–372.
<https://doi.org/10.1146/annurev.psych.56.091103.070145>
- Taylor, G. J., Bagby, R. M., & Parker, J. D. A. (1997). *Disorders of Affect Regulation: Alexithymia in Medical and Psychiatric Illness*. Cambridge University Press.
- Tooby, J., & Cosmides, L. (2008). The evolutionary psychology of the emotions and their relationship to internal regulatory variables. In M. Lewis, J. M. Haviland-Jones, & L. F. Barrett (Eds.), *Handbook of Emotions* (3rd ed., pp. 114–137). Guilford Press.
- Tracy, J. L., & Matsumoto, D. (2008). The spontaneous expression of pride and shame: Evidence for biologically innate nonverbal displays. *Proceedings of the National Academy of Sciences*, 105(33), 11655–11660.
<https://doi.org/10.1073/pnas.0802686105>
- Tracy, J. L., & Robins, R. W. (2007). The psychological structure of pride: A tale of two facets. *Current Directions in Psychological Science*, 16(3), 147–150.
<https://doi.org/10.1111/j.1467-8721.2007.00493.x>
- Van de Ven, N., Zeelenberg, M., & Pieters, R. (2009). Leveling up and down: The experiences of benign and malicious envy. *Emotion*, 9(3), 419–429.
<https://doi.org/10.1037/a0015669>
- Van Kleef, G. A. (2009). How emotions regulate social life: The emotions as social information model. *Current Directions in Psychological Science*, 18(3), 184–188. <https://doi.org/10.1111/j.1467-8721.2009.01633.x>
- Van Kleef, G. A., & Côté, S. (2022). The social effects of emotions. *Annual Review of Psychology*, 73, 629–658. <https://doi.org/10.1146/annurev-psych-020821-010855>
- Wanberg, C. R. (2012). The individual experience of unemployment. *Annual Review of Psychology*, 63, 369–396.
<https://doi.org/10.1146/annurev-psych-120710-100500>
- Warr, P. (2007). *Work, Happiness, and Unhappiness*. Lawrence Erlbaum Associates.
- Zhang, Y., Zhao, D., Hancock, J. T., Kraut, R., & Yang, D. (2026). Interaction with AI companions and psychological well-being. *Nature Human Behaviour*.
<https://doi.org/10.1038/s41562-026-02516-2>
- Blevins, T. (2026). Accommodation Goes Both Ways: Studying Linguistic Convergence Between Humans and Language Models. *arXiv:2605.29278*. <https://doi.org/10.48550/arXiv.2605.29278>
- Chu, M. D., Wu, Y., Chen, Z., Hwang, A. H.-C., & Luceri, L. (2026). When Chatbots Accommodate: What AI Companions Optimize for in Vulnerable Conversations. *arXiv:2606.04431*.
<https://doi.org/10.48550/arXiv.2606.04431>

Derner, E., Kučera, D., Gulati, A., Bagheri, A., & Oliver, N. (2026). Mind the Style: Impact of Communication Style on Human-Chatbot Interaction. arXiv:2602.17850. <https://doi.org/10.48550/arXiv.2602.17850>

Farach, A., Cambon, A., Tankelevitch, L., Hsueh, C., & Janssen, R. (2026). Scaffolding Human-AI Collaboration: A Field Experiment on Behavioral Protocols and Cognitive Reframing. arXiv:2604.08678. <https://doi.org/10.48550/arXiv.2604.08678>

Giles, H., Coupland, J., & Coupland, N. (Eds.). (1991). Contexts of Accommodation: Developments in Applied Sociolinguistics. Cambridge University Press.

Lee, H.-P. (H.), Sarkar, A., Tankelevitch, L., Drosos, I., Rintel, S., Banks, R., & Wilson, N. (2025). The impact of generative AI on critical thinking: Self-reported reductions in cognitive effort and confidence effects from a survey of knowledge workers. Proceedings of CHI 2025. <https://doi.org/10.1145/3706598.3713778>

Levinson, S. C., & Torreira, F. (2015). Timing in turn-taking and its implications for processing models of language. *Frontiers in Psychology*, 6, 731. <https://doi.org/10.3389/fpsyg.2015.00731>

Reeves, B., & Nass, C. (1996). *The Media Equation*. Cambridge University Press.

Riedl, C., Savage, S., & Zvelebilova, J. (2026). Cognitive Spillover in Human-AI Teams. arXiv:2407.17489v3. <https://doi.org/10.48550/arXiv.2407.17489>

Rimé, B. (2009). Emotion elicits the social sharing of emotion: Theory and empirical review. *Emotion Review*, 1(1), 60–85. <https://doi.org/10.1177/1754073908097189>

Stivers, T., Enfield, N. J., Brown, P., Englert, C., Hayashi, M., Heinemann, T., Hoymann, G., Rossano, F., de Ruiter, J. P., Yoon, K.-E., & Levinson, S. C. (2009). Universals and cultural variation in turn-taking in conversation. *Proceedings of the National Academy of Sciences*, 106(26), 10587–10592. <https://doi.org/10.1073/pnas.0903616106>

Sun, Y., Han, Y., & Lv, L. (2026). “Haste makes waste”: How delayed responses enhance customers’ usage intention of chatbot in

tourism services. *International Journal of Hospitality Management*, 137, 104673. <https://doi.org/10.1016/j.ijhm.2026.104673>

Székely, É., Miniota, J., & Hejná, M. (2025). Will AI shape the way we speak? The emerging sociolinguistic influence of synthetic voices. *Proceedings of IWSDS 2025*, 357–368. <https://doi.org/10.48550/arXiv.2504.10650>

Vaccaro, M., Almaatouq, A., & Malone, T. (2024). When combinations of humans and AI are useful: A systematic review and meta-analysis. *Nature Human Behaviour*, 8, 2293–2303. <https://doi.org/10.1038/s41562-024-02024-1>