

## **Rolul inteligenței artificiale în crearea sistemelor de organizare a cunoașterii pentru bibliotecile digitale**

*Jeny Jordan*

Dezvoltarea fără precedent a tehnologiei informaționale a dat curs unor schimbări majore în paradigmele informației și comunicării, ale modurilor de organizare și comunicare a cunoașterii. Modificarea infosferei – prin utilizarea pe scară largă a computerelor, a inteligenței artificiale și a mijloacelor multimedia interactive – a dus la modificarea intelectuală și a cunoașterii în genere. Societatea modernă a secolului XXI se bazează pe relația indestructibilă dintre știință și tehnologie ceea ce a dus implicit la mutarea accentului de la participarea individuală la actul cunoașterii la participarea colectivă, la o cooperare participativă realizabilă cu noua tehnologie. Noua tehnologie a schimbat ordinea și organizarea cunoașterii, a făcut posibilă interdisciplinaritatea și apariția unor noi domenii ale cunoașterii, ne-a schimbat radical modul de gândire și instrumentele operaționale de organizare a cunoașterii. Abundența de medii informaționale și de formate a dus la regândirea sistemelor de ierarhizare și ordonare a cunoașterii și la impulsivitatea personalului din biblioteci la găsirea de practici inovatoare și fiabile în raport cu noua tehnologie pentru împachetarea, distribuirea și regăsirea informației.

În secolul XXI, Internetul, Intranetul și conceptul de conectare la rețea au început deja să joace un rol major ca viitoare paradigmă pentru schimbare. Este era în care tehnologia informațională devine pilonul central al fiecărei biblioteci. Rețelele de computere vor deveni structurile pe baza cărora cele mai multe organizații vor fi construite, iar comunicarea și furnizarea de informații vor fi teme dominante. Noua tehnologie permite ca informația să fie ușor personalizată și adaptată, agenții cunoașterii vor putea furniza, în acest sens, un acces facil la informația cerută. Imaginea potențială a bibliotecilor digitale se

prefigurează ca o „poartă” deschisă informației în care toți cetățenii vor avea acces, asincronic, de oriunde la multiplele baze de date informaționale pentru a încuraja învățarea activă și permanentă.

În era informațională discursul uman se înscrie într-o nouă paradigmă – a scrisului pe ecran ce este purtat în spațiu virtual – ducând la crearea unei lumi meta-informaționale unde ponderea inteligenței artificiale în dezvoltarea cunoașterii devine tot mai semnificativă. Numerizarea informațiilor (voce, date, imagini etc.) permite realizarea rețelilor integrând servicii tot mai sofisticate: codare, acces la bănci de date, la poșta electronică etc. Se consideră că tehnica de calcul, informatica și cibernetica devin instrumente ale gândirii, mijlocind relația acesteia cu realitatea și determinând tot mai mult *ritmul și caracterul dezvoltării cunoașterii*.

Bibliotecilor le revine rolul major de a clasifica, a orândui și a organiza informația, care este, în fond, un mod de generare a cunoașterii sau care reprezintă instrumente necesare în procesul cunoașterii. Și în bibliotecă începem să utilizăm tot mai des analogiile informaționale, începem să analizăm problemele în termenii teoriei informației bazându-ne pe noua tehnologie informațională cu ajutorul căreia creăm „meta-informații”, „meta-date”, adică informații cu valoare adăugată. Integrarea tehnologiei în biblioteci înseamnă utilizarea tehnologiei pentru crearea unor instrumente adecvate și eficiente de căutare, selectare, stocare, diseminare, regrupare a informațiilor disponibile pe diferiți suportați, nu numai informația digitizată sau stocată pe discuri optice, ci și cea tipărită a cărei evidență și clasificare este digitizată.

Ponderea inteligenței artificiale în domeniul selectării, achiziționării, structurării și diseminării informației este semnificativă și în redimensionarea cunoașterii obiective prin acte cognitive ce sunt influențate de prezența și manifestarea acesteia. Autoreflexia și comunicarea cunoașterii acustice au fost înlocuite de prezența ubicuă a computerului și a comunicării cibernetice. „Inteligența artificială are capacitatea de a se integra unor forme deschise de dialog iradiind întregul unor experiențe *sui-generis* de natură cognitivă”<sup>1</sup>, spune Cornelia Maria Bârliba. Inteligența artificială a devenit un instrument formal de cunoaștere. Dihotomia „intelență naturală – intelență artificială” duce

<sup>1</sup> MARIA CORNELIA BÂRLIBA, *Introducere în epistemologia informațională*, București: Editura Științifică, 1990, p. 9.

la descoperirea unor resurse cognitive în sfera „netradițională” a posibilităților de manifestare a cunoașterii, renunțând la ierarhiile consacrate.

Noul *organon* al societății informaționale și computerizate în care trăim ar fi transformarea inteligenței artificiale într-o metodă de cercetare. Iar pentru performanțele acesteia de a atinge nivelul inteligenței umane și chiar să-l depășească, este nevoie de înțelegerea naturii proceselor cognitive ale minții umane, deconstruindu-le, pe cât posibil, pentru a putea reface, tehnic, acea rețea neuronală care transformă informația în cunoaștere. „Dimensiunea diacronică constitutivă a structurii de cercetare om-computer reține prin sinteză criteriile, principiile cunoașterii umane care, asociate legilor calculatorului, se vor transforma în norme ale unității primare a actelor intelective în societatea informațională.”<sup>2</sup> Așanumitele abilități cognitive superioare nu pot fi înțelese dacă nu se subliniază procesele de bază ale fiecărui neuron, sinapse, porți și a altor componente care trebuie să fie înțelese.

Cercetările făcute în domeniul numit *științe cognitive (cognitive sciences)* au ajuns într-un punct provocator pentru tehnologia informațională avansată în sensul că se afirmă că imageria mentală rezultă dintr-o activitate cerebrală similiară ca operațiuni cu cea a computerului, această activitate determinând, prin calitatea procesării, acuratețea iconică a fiecărui individ: „Ipoteza este că imaginile vizuale sunt ca niște *display-uri* pe un tub catodic (CRT), generate de un program computerizat (*plus data*).”<sup>3</sup> Teza dată de Kosslyn se înscrie în *teoria computațională a minții umane*. Conceperea minții ca un sistem de organe computaționale a cărui alcătuire rezultă de pe urma procesului de selecție naturală (mintea este un soi de computer neuronal încastrat cerebral) și al cărui „program” principal ar fi acela de a rezolva orice fel de problemă intrată în atenție ține de ceea ce s-ar putea numi *deconstrucția inginerească* a minții și creierului uman. Teoria computațională a minții ar putea fi asociată înțelegerii stărilor și convingerilor umane ca *informație*, la rândul ei încapsulată în „program” sub formă de simboluri. În forma ei extremă, scopul ei este să explice cogniția în termenii fizicii. Realizarea fizică a sistemului cognitiv este

<sup>2</sup> *Ibidem*, p. 14.

<sup>3</sup> *Apud* STEPHEN MICHAEL KOSSLYN, *Image and Mind: The Resolution of the Imagery Debate*, MIT Press, 1980, p. 5.

descrișă speculativ de un mecanism construit artificial capabil să opereze prin intermediul unor reguli, de tip logic, transformând informația în informație cu valoare adăugată prin *manipulare simbolică*.

Teoria computațională a minții – ca rezultată a combinării metodelor și conceptelor din știința computerelor și neuroștiință – încearcă să realizeze o mai bună înțelegere a creierului uman și a proceselor cognitive, căutând să explice cum semnalele chimice și electrice procesează informația în sistemul nervos și cum funcția de reprezentare este îndeplinită. Sistemul nervos este interpretat ca un sistem de procesare a informației a cărui dinamică poate fi simulată pe computer. Știința neuronală furnizează teoriile de bază despre anumite aspecte ale neuronilor, rețelelor neuronale și proceselor sinaptice, iar știința computerului încearcă să le modeleze și să le simuleze pe computer<sup>4</sup>.

În epoca *inteligenței artificiale* s-a produs o diviziune semnificativă în rolul și în importanța celor două diviziuni cognitive și metodologice: o primă diviziune, cea *clasică*, în care cele două metode se raportează la lumea omului, la existență în genere, iar în cea de-a doua la lumea omului cu bogăția inerentă de sensuri și semnificații dar transpusă în cadrele inteligenței artificiale. Are loc un proces de transformare a cunoașterii naturale în mecanisme automate și o reevaluare, o reconsiderare a procesualității inteligenței naturale din perspectiva sistemelor de inteligență artificială. Modelarea inteligenței artificiale după modelul inteligenței naturale, a creierului uman în speță, duce la proiectarea unor modele tehnologice noi care au la bază principiile de funcționare ale sistemelor vii. Inteligența creează inteligență și tinde s-o depășească prin crearea unor astfel de sisteme inteligente care să fie în stare nu numai să colecteze și să selecteze informația, dar să prelucreze informația care să ducă la cunoaștere implicită. Cu inteligența artificială ne aflăm într-un context informațional simulat, nu autentic: „... în inteligența artificială nu ne aflăm în fața unei informații prin «geneză», ci a unei informații «de atribuire». În seria consecințelor cu rol de diminuare a specificului informațional se află transpunerea informației în sens uman (prin geneză și semnificație) pe suporturi mecanice, ceea ce estompează unele trăsături ale notei informației umane.”<sup>5</sup> Informația apare pe calculator ca un

<sup>4</sup> Apud \*\*\* *Encyclopedia of Semiotics* edited by Paul Bouissac, New York-Oxford : Oxford University Press, 1998, pp. 135-136.

<sup>5</sup> MARIA CORNELIA BĂRLIBA, *op. cit.*, p. 246.

obiect, ca materie primă pentru obținerea altei informații, ea nu se aseamănă nici cu informația programată, nici cu cea umană existentă în afara spațiului calculatorului; nu are sens pentru calculator, ci numai pentru subiectul uman. Știința computerelor este un nou fenomen cognitiv.

Specialiștii din domeniul științei neuronale computaționale nu folosesc computerele în experimentele lor numai pentru colectarea, transformarea și computerizarea datelor empirice, ci le folosesc de asemenea ca mecanisme care generează ele însele rezultate. Deoarece modelele neuro-științifice sunt destul de complexe, este imposibil să le prezici comportarea pur intuitiv, de aceea dinamica lor implicită trebuie să fie făcută explicită prin simularea<sup>6</sup> lor pe computer. Alternativa la complexitatea și limitările modelului creierului uman este simularea care capturează numai trăsăturile de bază ale structurilor neuronale naturale, cum ar fi procesarea paralelă, procesarea locală și un înalt nivel de interacțiune între procesele simple și răspândirea activărilor. Cel mai prominent exponent al modelului creierului simplificat realizat până acum este clasa de modele cunoscută sub numele de *conecționiste* sau *procesare paralel distribuită* (*PDP – parallel distributed processing*). Modelul conecționist simplifică modelul creierului uman, străduindu-se să simuleze procesele care au loc într-o rețea neuronală la un nivel foarte abstract. Această abordare computațională diferă de programele de computer convenționale în care informația este înmagazinată mai mult sau mai puțin (lingvistic) explicit și poate fi regăsită printr-o adresă de un algoritm (schimbând, stocând, duplicând, ștergând etc.) operând asupra acestor reprezentări. Evidența empirică din știința neuronală arată că nici un simbol, sintagmă, propoziție sau algoritm (în sensul tradițional) nu poate fi găsit în creier. Trebuie să existe o alternativă, mult mai primitivă, un mecanism pentru reprezentarea și procesarea cunoașterii. Pentru aceasta, abordarea conecționistă construiește un model simplificat al

---

<sup>6</sup> Prin *simulare* înțelegem că o problemă dintr-un anumit domeniu al cunoașterii este, în primul rând, transformată într-un spațiu reprezentational. Acest construct sau structură formală acționează ca un model al realității, căruia putem să-i aplicăm un algoritm de inferență sau un set de reguli și să-i schimbăm parametrii, realizând noi parametrii și simboluri. Realitatea este simulată în dinamica unui spațiu de reprezentare artificială iar rezultatele acestei simulări sunt «aduse înapoi la realitate» în urma unui act de interpretare de către utilizator.

creierului uman, componentele de bază ale unui sistem conexiunist fiind reduse la un set de unități de procesare care modelează neuroni singuri, procesarea urmând să aibă loc în interiorul fiecărei unități neuronale. Toate unitățile lucrează în paralel și complet local, iar unitățile care sunt „conectate” la mediu sunt numite unități de intrare sau de ieșire.

Modelele conexiuniste oferă de asemenea câteva aplicații interesante, ca de exemplu simularea memoriei asociative, sarcini-model și sarcini de recunoaștere a imaginii, procesarea limbajului, simularea unor simple circuite neuronale, control robotic, crearea unui model, descompunere suplă, aplicații-reguli-de-program, toleranța greșelii și altele<sup>7</sup>. Dintr-o perspectivă filosofică și epistemologică, rețelele conexiuniste au implicații interesante pentru înțelegerea noastră a cunoașterii și a reprezentării cunoașterii în sistemele cognitive naturale cât mai ales în cele artificiale. Din acest punct de vedere interesul s-ar putea contura, pentru bibliotecile digitale, în aplicații de perfecționare a sistemelor de organizare a cunoașterii integrate de bibliotecile digitale prin crearea unor complexe *ontologii* și *rețele semantice* capabile să „producă” informație cognitivă.

Procesele computoriale sunt considerate procese de transformare simbolică prin care informația este stocată, clasificată și indexată. Conceptele computoriale pot fi folosite pentru a specifica tipuri calitativ distincte ale reprezentării datelor și transformării lor. Componenta cognitiv-informațională dobândește valențe noi în contextul dimensiunii conceptuale a tehnologiilor de vârf. *Informația cognitivă* este un produs sofisticat al computerului, caracterizat prin aceea că este o *proiecție a viitorului* (informație ce este folosită pentru detectare și prognoză), este *logică* (ceea ce înseamnă că presupune existența unui scop, a unei relații între rezultate), și *acțional-selectivă* (în sensul în care informația este folosită pentru selectarea acțiunilor și mijloacelor cele mai adecvate pentru atingerea scopului propus). Dimensiunea entropică este absentă din structurile algoritmizate, ceea ce constituie un neajuns semnificativ în a defini inteligența artificială ca fiind genul proxim al inteligenței umane, care are această dimensiune. Entropia este o ordine puternică, funcțională în măsura în care deține rolul de principiu fundamental al explicării

<sup>7</sup> Apud \*\*\* *Encyclopedia of Semiotics* edited by Paul Bouissac, New York-Oxford : Oxford University Press, 1998, pp. 141-144.

aparitiei noului, al fenomenului creației. Prin dezvoltarea și perfecționarea sistemelor conexiunilor se țintește depășirea acestei dificultăți încercând modelarea unui sistem neuronal artificial care să poată trasa inferențe asupra unor fapte, date și reguli sau într-o rețea semantică, pe baza anumitor reguli și algoritmi manipulând simboluri.

Noul val al comunicării în spațiul cibernetic a luat cu asalt orice formă tradițională de conservare a cunoașterii pentru a remodela-o într-o structură ca adjuvant al cunoașterii. Biblioteca digitală, ca paradigmă a bibliotecii viitorului, este o noțiune care prefigurează o bibliotecă inovativă în structură, în procesarea informației și în disponibilizarea ei prin aplicații complexe ale tehnologiei și se bazează pe un sistem de organizare a cunoașterii care poate fi creat local sau prin importare de date dintr-o altă sursă a unui astfel de sistem pe care, ulterior, îl modifică, îi adaugă sau șterge anumite informații. Termenul de sistem de organizare a cunoașterii (*knowledge organization system*) este proiectat pentru a cuprinde toate tipurile de scheme pentru organizarea informației și promovarea managementului cunoașterii. Sistemele de organizare a cunoașterii cuprind scheme de clasificare care organizează materialele la un nivel general (cum ar fi aranjarea într-o anumită ordine a cărților pe rafturi), vedete de subiecte care furnizează mai multe informații și permite un acces mai detaliat și fișierele de autoritate care controlează variantele diferite ale informațiilor cheie (cum ar fi numele geografice și numele de persoane, de autori). Ele includ de asemenea și scheme mai puțin tradiționale, cum ar fi rețelele semantice și ontologiile. Din această cauză sistemele de organizare a cunoașterii sunt mecanisme pentru organizarea informației, ele sunt „inima” fiecărei biblioteci, muzeu sau arhivă.

Sistemele de organizare a cunoașterii sunt folosite pentru a organiza materialele pentru a le regăsi mai ușor și pentru a administra colecția. Un astfel de sistem funcționează ca o punte de legătură între nevoile informaționale ale utilizatorului și documentele colecției. Cu ajutorul lui, utilizatorul este capabil să identifice sursa dorită fără o cunoaștere anterioară a existenței acelei surse. Sistemul de organizare a cunoașterii ghidează utilizatorul prin teme de paginilor Web sau ale motoarelor de căutare a site-urilor prin baleiere sau căutare directă. În plus, acest sistem permite organizatorilor să răspundă la întrebări privitoare la scopul colecției și la ceea ce este nevoie să se găsească în ea.

Toate bibliotecile digitale utilizează unul sau mai multe astfel de sisteme. Exact ca într-o bibliotecă fizică, sistemul de organizare a

cunoașterii, într-o bibliotecă digitală, furnizează o vedere de ansamblu asupra conținutului colecției și sprijină regăsirea informațiilor. Schema unui astfel de sistem poate fi una tradițională relevantă pentru scopul materialului (cum ar fi sistemul zecimal Dewey sau *INSPEC Thesaurus*), o schemă dezvoltată comercial cum ar fi site-urile *Yahoo*, *Amazon*, *Hotmail*, *Alapage* și *Excite* sau o schemă dezvoltată local pentru un intranet colectiv.

Odată cu creșterea puterii de computerizare a informației și cunoașterii ca și creșterea motoarele de căutare sofisticate și a numărului de utilizatori care utilizează individual aceste instrumente de căutare, s-a ajuns la schimbări majore în sistemele de organizare a cunoașterii chiar în mediul electronic, al Web-ului. Pentru o lungă perioadă de timp, sistemele de organizare a cunoașterii relaționate cu serviciile de indexare și abstractizare erau aplicate numai de catalogatori și indexatori antrenați în utilizarea indexării sistemelor de organizare a cunoașterii pentru un produs particular sau produse. Primii utilizatori ai unor astfel de instrumente au fost bibliotecarii și alți căutători profesioniști. Odată cu proliferarea fără precedent a datelor electronice, a exploziei de publicații electronice și a creșterii preocupării pentru facilitarea localizării informației au condus la un interes reînnoit în dezvoltarea sistemelor de organizare a cunoașterii pentru folosirea acestora nu numai de către bibliotecari și de profesioniștii informației, ci și de utilizatorii finali.

Una din caracteristicile deosebite și de mare interes ale unei biblioteci digitale este cea disponibilitate de a face resursele accesibile și altor comunități. Și acest lucru este posibil datorită Internetului care este un fel de cameră imensă umplută cu cărți, un loc de întâlnire care aduce laolaltă diferite comunități, persoane, informații etc. Sistemele de organizare a cunoașterii, pe Internet, sunt mijloace de conectare a acestor disparate comunități și pot fi folosite pentru a furniza acces alternat la subiecte, pentru a adăuga moduri de înțelegere a resurselor bibliotecilor digitale, a sprijini accesul multilingual (prin adăugarea unor dicționare și tezaure multilinguale) și a aproviziona cu termeni pentru extinderea căutării textelor libere în domeniile care sunt relativ necunoscute utilizatorului. Căutarea textelor libere este principala metodă de căutare pe web. Numai un procentaj mic al resurselor web au metadate și chiar un și mai mic procentaj au precizat un vocabular de control. Totuși, limbile diferite în care circulă informația fac problematică căutarea

textelor libere. Chiar un utilizator experimentat în astfel de căutări nu poate să știe toată terminologia (sinonime sau termeni de relație) care să poată fi utilizați în literatura de specialitate pentru a exprima un concept. Problema este exacerbată când utilizatorul nu este familiar cu domeniul sau este interesat de o zonă interdisciplinară a cunoașterii. Un ajutor poate veni din partea sistemelor de organizare a cunoașterii pentru selectarea cuvintelor cheie pentru textele libere. Sistemele de organizare a cunoașterii pot fi foarte valoroase în sprijinirea căutării textelor libere în cadrul bibliotecii digitale și în integrarea resurselor web într-o bibliotecă digitală existentă. Totuși, aceste sisteme trebuie să fie utilizate cu atenție. Sistemele de organizare a cunoașterii au fost în general dezvoltate pentru o disciplină specifică, o sarcină sau o funcție anumită sau pentru indexarea unei colecții specifice sau baze de date. De aceea, în funcție de domeniul în care sistemul de organizare a cunoașterii este folosit și de complexitatea sistemului, el poate sau nu poate sugera termeni relevanți pentru căutarea textelor libere. Extinzând o căutare cu termeni corelați, mai degrabă decât prin pure sinonime, putem să avem niște rezultate care sunt numai marginal relevante pentru utilizator.

Deoarece bazele de date online s-au mutat pe web, sistemele de organizare a cunoașterii au început să-și furnizeze produsele lor incluzând uneltele de vocabular în acest mediu. Pentru a ușura căutarea și afișarea informației din sisteme de organizare a cunoașterii necunoscute anterior, sistemul trebuie să aibă identificatori unici și durabili pentru fiecare din conceptele din sistem. O a doua cerere este un protocol pentru o interogare distribuită și răspunsul sistemului de organizare a cunoașterii. Aceasta este o problemă critică pentru sistemele structurate complex ca tezaure, rețele semantice și ontologii. Tezaurele se bazează pe concepte și arată relațiile dintre termeni. Relațiile exprimate de obicei într-un tezaur include o ierarhie, sinonimie (echivalență de sens) și o asociere sau corelare. Aceste relații sunt în general reprezentate prin notări de tipul descriptorilor generici (*BT – broader term*), termenilor specifici (*NT – narrower term*), sinonimelor (*SY – synonym*) sau termenilor asociați sau corelați (*RT – associative term* sau *related term*). Multe tezaure sunt foarte mari; ele pot include mai mult de 50.000 de termeni. Multe dintre ele au fost dezvoltate pentru o disciplină specifică sau pentru un produs specific. Relațiile semantice nu sunt sub formă de ierarhii, ci mai degrabă sub formă de rețele sau „țesătură”. Conceptele

sunt gândite ca noduri sau ca legături relaționale din afara lor. Ontologiile sunt cele mai noi etichete ce pot fi atașate unor anumite sisteme de organizare a cunoașterii. Ele au fost dezvoltate ca modele conceptuale specifice și reprezintă relații complexe între obiecte, incluzând reguli și axiome ce lipsesc din rețelele semantice. Ontologiile care descriu cunoașterea într-un domeniu specific sunt adesea conectate cu sisteme de căutare a datelor și de management al cunoașterii<sup>8</sup>.

Teoria computațională a minții ar putea juca un rol deosebit de important în realizarea unor sisteme conexiuniste perfect aplicabile unui sistem de organizare a cunoașterii, făcând posibilă construirea de rețele semantice și ontologii complexe care să ajute utilizatorul să regăsească rapid și eficient informația dorită din noianul de surse existent pe Internet și, mai mult decât atât, să-l ajute efectiv în procesul cunoașterii prin conexiunile și asocierile inedite de termeni și/sau surse realizabile cu ajutorul unor astfel de instrumente de căutare și relațiile pertinente ce se pot crea între diverse câmpuri conexe ale cunoașterii care să producă plusul de informație interdisciplinară. Noile sisteme de organizare a cunoașterii ar putea deveni producătoare de cunoaștere, mijloace interactive și performante care să accelereze progresul cunoașterii. Astfel comunicarea om-mașină ar depăși statutul de simplă sintagmă și ar deveni o realitate vie, o simbioză perfectă între natural și artificial.

<sup>8</sup> \*\*\**Systems of Knowledge Organization for Digital Libraries*: <http://www.clir.org/pubs/>