

# EVOLUȚIA PRELUCRĂRII ȘI TRANSMITERII DATELOR ÎN SOCIETATEA UMANĂ

Lenuța ALBOAIE<sup>1</sup>

adria@info.uaic.ro

**ABSTRACT:** In this paper we realize an overview along a direction that allows us to capture the role of data, information and knowledge in human society. From the first data stored in stone tables and papyrus, and from the first processing performed by abacus, we reached at complex systems for data storage and management. In the present context, it is recognized that IT (Information Technologies) has become an omnipresent factor in our lives and its importance generated a very fast development of society. Human communities, their groupings, led to the current conditions where a huge amount of data exists on the Web (e.g. social networks). This paper discusses some aspects as trust and reputation: necessary conditions in order to interact with appropriate data. It also focuses on data structures and mechanisms of modeling and transmission, in order to control the social impact of these aspects.

**KEYWORDS:** information management, social networks, trust, reputation

## Introducere

Omul este o ființă din universul fizic ce simte nevoia de cunoaștere. Cunoașterea este obținută prin observație, percepție, intuiție și experimentarea diferitelor lucruri. Omul este de asemenea o ființă socială, ceea ce a condus la apariția comunicării. Sintetic putem considera:

- ◆ primele instrumente folosite: limbajul, semnele și vorbirea;
- ◆ primele forme de comunicare scrisă: scrierea pe tăblițe de piatră, pereții peșterilor sau papirus.
- ◆ primele forme de comunicare la distanță, precursori ale Internetului actual: semne cu foc sau codul morse.

În contextul evoluției societății umane, și ne vom referi în lucrarea de față mai puțin la aspecte economice și politice și mai degrabă la cele sociale și tehnologice, perspectiva oamenilor asupra datelor, informațiilor a cunoscut o schimbare radicală odată cu apariția societății informaționale.

---

<sup>1</sup> Lector, dr., Alexandru Ioan Cuza University, Computer Science Department, Iași, colaboratoare a Diviziei de Istoria Științei a CRIFST al Academiei Române

Tehnologiile actuale au permis o creștere a interacțiunii între persoane care nu se cunosc personal. În acest caz interacțiunile online înlocuiesc de multe ori interacțiunile umane și astfel se simte nevoia de creare a unui cadru virtual potrivit care să fie ca o oglindă a lumii reale. În acest demers s-a observat ca există doi factori importanți care catalizează interacțiunile între persoane, atât în lumea reală cât și în lumea virtuală (online): încrederea și reputația.

Încrederea și reputația sunt două concepte interconectate. Încrederea există în mod normal la nivel personal. Reputația exprimă o opinie colectivă rezultată din părerile membrilor dintr-o comunitate. Aceste două concepte împreună cu noțiunile de dată, informație, cunoștințe le vom analiza în secțiunea 2 a prezentei lucrări.

În lumea reală oamenii interacționează cu alți oameni dar și cu date, informații cu scopul de a căpăta mai multe cunoștințe. În lumea online lucrurile stau într-un mod similar, diferența ar fi că spațiul Web stochează cantități foarte mari de date care sunt în continuă creștere, și care constituie pentru oameni o adevărată bază de cunoaștere. Însă în această mare de date apare o problemă esențială: nu lipsa de date calitative așa cum am putea crede la o primă impresie, ci găsirea și organizarea acestora. Această problemă a supraîncărcării cu informație (eng. *information overload*), a fost prevăzută de Alvin Toffler încă din 1970 în cartea sa *Șocul Viitorului* [Toffler, 1970] în care se precizează și efectul acestei probleme: iluzia imaginii complete asupra domeniului de interes. Domeniul filtrării informației (eng. *information filtering* vine cu o serie de propuneri care aduc soluții la diferite clase de probleme pe această direcție. Pe această axă, implicând conceptele de încredere și reputație ne vom încadra în discuțiile abordate de aceasta lucrare.

Sociologul Kate M. Gordon a realizat un experiment prin care a solicitat unui număr de 200 de studenți să evalueze o serie de itemi. S-a observat că răspunsul estimat pe grup a fost corect în proporție de 94%, rezultat mult mai relevant decât cel al studenților care au dat răspunsuri corecte în mod individual. În [Surowiecki, 2005] se specifică faptul că, în general, deciziile de grup sunt „mai inteligente decât decizia celui mai inteligent din grup”. Acest tip de evaluare poate conduce însă la riscuri de tipul penalizării ideilor inovative, minoritare, problemă descrisă în [Tocqueville, 1840] sub denumirea de „tirania majorității” (eng. *tyranny of the majority*). În mod firesc, și opiniile grupurilor minoritare contează și trebuie văzute ca oportunități. Dacă însă acestora li se acordă o prioritate totală, ajungem la cealaltă extremă, așa numitul fenomen de „*echo chamber*”. În acest caz, se arată în [Sunstein, 1999] se va ajunge la o fragmentare a societății

în micro-grupuri care tind să își ducă la extrem propriile păreri. Mai mult, membrii acestora vor interacționa doar cu cei asemănători lor și nu se vor mai confrunta cu idei și opinii noi. Se vor forma grupuri care vor găsi greu căi de intercomunicare. Vom discuta în secțiunea 3 modul în care un sistem de încredere și reputație asigură membrilor comunităților online o îmbinare echilibrată a unei viziuni personalizate cu o viziune globală asupra sistemului. Așa cum vom vedea din exemplul propus, această viziune, se concretizează în spațiile diferitelor comunități online prin detectarea esențialului (a utilizatorilor sau a informațiilor cele mai relevante).

În secțiunea 4 a lucrării vom exemplifica îmbunătățirile pe care conceptele de încredere și reputație le poate aduce în mecanismul de filtrare al informației, mecanism extrem de important pentru fericirea sau de ce nu, supraviețuirea unui individ.

Secțiunea 5 este dedicată concluziilor asupra acestei lucrări și tot aici prezentăm direcțiile pentru viitoarele oportunități de cercetare.

### **Date, informații și sisteme care ne conduc la îmbunătățirea cunoașterii**

De la începuturile omenirii nu credem că a existat un segment de activitate care să nu implice în mod implicit sau explicit concepte de date, informație, cunoștințe etc. O viziune generală acceptată asupra acestor concepte este următoarea [Ackoff, 1989], [Liew, 2007], [Davenport, 1998], [Garvin, 1996], [Krogh et.al., 2000], [Amrit, 2001]:

- ◆ Datele sunt văzute ca o colecție de fapte în forma brută, care pot exista sub orice formă, fără un înțeles sau utilizare specială. Astfel datele pot exista dar nu au o semnificație. Un exemplu ar fi un formular care conține date diverse.

- ◆ Informațiile sunt rezultatul analizei și agregării datelor și sunt văzute ca grupări cu sens care transformă data din forma brută în colecții utile, cu înțeles. Informațiile sunt date între care s-au stabilit niste legături relaționale. Aceste legături nu asigură neapărat o utilitate informațiilor respective. Un exemplu de informații ar fi bazele de date relaționale, care transformă datele stocate din interiorul lor în informații.

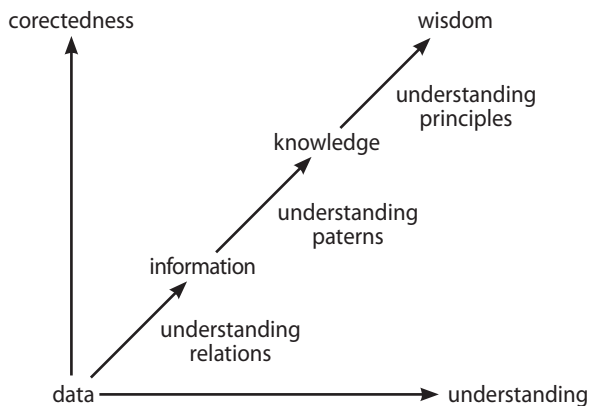
- ◆ Cunoștințele, rezultate în urma studiului și acumulării de experiență, reprezintă o însumare a ceea ce a fost perceput, învățat sau descoperit și care duce la sensibilizare, familiaritate și/sau înțelegere. Cunoașterea este un proces determinist, care asigură un înțeles pentru cel care a acumulat informații și poate implica uneori un proces de inferență de noi cunoștințe.

Conform lui Russell Ackoff [Ackoff, 1989], conținutul minții umane poate fi împărțit în cinci categorii: data, informație, cunoștințe, înțelegere, înțelepciune. Primele patru sunt strâns corelate cu trecutul iar înțelepciunea integrează aspecte legate de viziune și proiectare ce sunt în strânsă legătură cu viitorul.

În accepțiunea lui Ackoff, înțelegerea este procesul prin care cineva pe baza cunoștințelor existente sintetizează noi cunoștințe. „Diferența dintre înțelegere și cunoștințe este similară diferenței între învățare și memorare.” [Ackoff, 1989]

Înțelepciunea este un proces de extrapolare, nedeterminist și non-probabilistic care implică cunoașterea nivelurilor anterioare și putem sintetiza că reprezintă procesul de evaluare a înțelegerii, de a discerne între corect și incorect, între bine și rau.

În [Bellinger et. al., 2004] este prezentată o viziune ușor modificată față de Ackoff asupra conceptelor de dată, informație, cunoștințe, înțelepciune – DIKW (*Data, Information, Knowledge, Wisdom*).



**Fig. 1** – Ierarhia DIKW (*Data, Information, Knowledge, Wisdom*)  
[Bellinger et. al., 2004]

În această viziune înțelegerea asigură liantul între cele patru concepte analizate. Datele reprezintă un fapt sau o declarație fără legătură cu altceva. Informațiile implică înțelegerea unei relații de un anumit fel (e.g. între cauză și efect). Cunoașterea reprezintă un model care asigură un nivel ridicat de înțelegere cu privire la ceea ce este descris. Înțelepciunea este sistemică și implică mai mult decât înțelegerea principiilor fundamentale integrate în cunoștințe.

Sistemele de calcul au apărut având ca și model procesorul uman. S-a urmărit o imitare a modului în care omul ajunge să ia

decizii cât mai înțelepte pornind de la datele brute de intrare [Drăgănescu, 1979].

În viitorul apropiat sistemele de calcul nu vor depăși faza de asistare a omului în luarea unor decizii majore în multe domenii (medical, comerț electronic, etc.), dar ele vor fi de un real folos atunci când vor oferi rezultate inteligente în analiza unor cantități imense de date. Altfel spus nivelul de „înțelepciune” al acestor sisteme de calcul datorat diferitelor încercări ale lumii IT (e.g. Inteligența Artificială, Semantic Web, etc.) se va apropia din ce în ce mai mult de înțelepciunea umană. Este cert faptul că sistemele existente pot constitui un factor important în îmbunătățirea procesului de cunoaștere.

În lucrarea de față, privim la situația actuală în care IT-ul a devenit un factor omniprezent în viața fiecăruia dintre noi, și care a determinat o ușoară translație a activităților noastre din lumea reală imediată într-un spațiu virtual, online. Ca și în lumea reală, și aici interacționăm cu alți oameni (pe care îi vom numi mai departe în mod generic utilizatori) și avem de-a face cu date, informații care să ne conducă pe fiecare din noi la noi nivele de cunoaștere, sau și mai mult, la noi nivele de înțelepciune. Scopul discuției noastre este cel de a contribui la creșterea calității acestor interacțiuni, în condițiile în care avem utilizatori care nu se cunosc personal. Dacă în lumea reală, alegem cu mai mare ușurință cu cine dorim să interacționăm sau ce informații dorim să selectăm, în lumea virtuală avem nevoie de un mecanism care să ne asigure un nivel automat de selecție care să țină cont de acțiunile și preferințele noastre.

Și în acest context intervin două concepte, încrederea și reputația care joacă un rol esențial în lumea reală și care așa cum vom vedea au un rol definitoriu în asigurarea calității interacțiunilor online.

Încrederea și reputația sunt două concepte strâns legate între ele. Încrederea se găsește la nivel personal. Reputația exprimă în schimb, o opinie colectivă obținută pe baza părerilor membrilor din comunitate.

În fiecare dintre domeniile în care încrederea joacă un rol important, de exemplu: sociologia, psihologia, științe politice, economia, filosofia și informatica, au fost date numeroase definiții conceptului de încredere. Discuția în această lucrare se va purta din perspectiva modului în care încrederea este folosită și modelată în acest moment în sistemele online disponibile pe Web. În [Gambetta, 1990] se găsește următoarea definiție:

*Definiție: Încrederea este un nivel particular de probabilitate subiectivă cu care un agent va efectua o acțiune particulară, atât*

*înainte ca noi să putem monitoriza o asemenea acțiune, cât și într-un context în care este afectată propria noastră acțiune.*

Altfel spus o entitate A (utilizator) are încredere într-o alta entitate B dacă realizează o acțiune pe baza încrederii ca acțiunile viitoare a lui B vor conduce la obținerea unor rezultate bune.

Conceptul de încredere are asociate patru proprietăți [Golbeck, 2005]:

- ◆ tranzitivitate (engl. *transitivity*);
- ◆ compozabilitate (engl. *composability*);
- ◆ personalizarea (engl. *personalization*);
- ◆ asimetria (engl. *asymmetry*).

Aceste proprietăți fac posibil să privim încrederea ca și concept computațional și astfel să putem avea mecanisme de calcul a încrederii.

Încrederea nu este perfect tranzitivă în sens matematic, ci reprezintă faptul că noțiunea de încredere poate fi transmisă între entități. De exemplu: A are încredere mare în B, B are încredere mare în C, dar e posibil ca A să aibă o încredere medie în C.

Cea de-a doua proprietate a încrederii, *compozabilitatea* asigură calcularea încrederii de-a lungul lanțurilor de utilizatori. Această proprietate specifică de fapt cum sunt propagate ratingurile asociate încrederii între entități care nu sunt în contact direct. Putem avea situații de tipul celei considerate mai sus în care A are încredere în B, B are încredere în C și atunci între A și C se stabilește un nivel de încredere. Poate exista situația și în care A poate beneficia de mai multe păreri asupra lui C și în acest caz se face o compunere a acestor valori pentru obținerea unei evaluări potrivite.

Cea de-a treia proprietate a încrederii o reprezintă *caracterul personal al încrederii*. De exemplu, asupra entității C, A și B pot avea păreri diferite. Aceasta varietate a încrederii conduce la calculul încrederii din perspectiva unei entități, rezultate obținute vor trebui să reflecte interesele entității respective.

O altă proprietate a încrederii este *asimetria* și semnifică faptul că dacă A are încredere în B, este posibil ca B să nu aibă același nivel de încredere în A. Avem de-a face cu așa numita „încredere unidirecțională” (eng. *one-way trust*) [Hardin, 2002][Cook, 2001].

În esență, încrederea este reprezentată de judecățile unui utilizator asupra altuia, uneori realizate direct și explicit, alteori indirect prin evaluarea diferitelor acțiuni realizate de acel utilizator.

În lumea reală, avem în general persoane în care avem încredere mai mare sau mai mică. Aceasta formează de fapt așa numita rețea de

încredere. O rețea de încredere (engl. *trust network*) este un graf obținut prin agregarea aprecierilor utilizatorilor. Aceste aprecieri pot fi cuantificate, astfel încât putem avea diferite niveluri de încredere ce se pot stabili între utilizatori. Într-o astfel de rețea de încredere poate rula așa-numita *metrica de încredere* (eng. *trust metric*) care este de fapt un algoritm ce primește la intrare informațiile din rețea și calculează diferite valori ale încrederii între utilizatori [Massa, 2006]. În literatura de specialitate se folosește termenul de metrică de încredere, unde metrică nu are sens de metrică drept concept matematic (funcție distanță). O metrică de încredere este un algoritm care poate prevedea încrederea pe care un membru o poate acorda unui alt membru. De fapt, metricile de încredere au rolul de a exploata raționamentele făcute de alți utilizatori, chiar dacă un utilizator are o imagine locală parțială asupra sistemului. În general, toate metricile de încredere pornesc de la presupunerea că încrederea poate fi propagată (proprietatea de tranzitivitate a încrederii), astfel dacă A are încredere în B și B are încredere în C, atunci se poate stabili un nivel de încredere între A și C.

Metricile de încredere sunt împărțite în metrici cu domeniu local și metrici cu domeniu global. Metricile globale iau în calcul toate nodurile și legăturile de încredere existente. O valoare globală este asignată unui agent pe baza tuturor informațiilor din rețea. În literatura de specialitate o serie de cercetători susțin faptul că, în mod firesc, nu există metrică de încredere globală și pot exista doar metrici locale [Mui, Mohtashemi & Halberstadt, 2002]. În esență, metrica de încredere globală calculează reputația globală și nu încrederea personalizată. Discuțiile din această lucrare se vor încadra pe direcția stabilită de [Mui, Mohtashemi & Halberstadt, 2002].

Definirea conceptului de reputație, ca și a celui de încredere, a resuscitat interesul multor cercetători.

Pornim de la două definiții pe care le întâlnim în dicționarul Merriam-Webster și respectiv în dicționarul Compact Oxford și în care conceptul de reputație este definit astfel:

Definiția 1: *calitatea generală sau caracterul văzut sau judecat de oameni în general*

Definiția 2: *convingerile sau opiniile păstrate în general despre ceva sau cineva*

Altfel spus, reputația este ceea ce se spune sau crede despre o persoană.

În [Rahman et. al., 2000a] se specifică faptul că reputația încorporează toate așteptările legate de comportamentul unei entități pe baza informațiilor existente în istoria entității respective.

În comunitatea rețelelor sociale reputația este un cuantificator rezultat în urma diferitelor interacțiuni într-o rețea socială și reputația unui individ este vizibilă tuturor membrilor rețelei sociale. Multe alte sisteme cum ar fi eBay, Amazon folosesc o astfel de reputație globală.

În [Josang et. al., 2007] se explică în mod informal care este diferența dintre încredere și reputație. Să considerăm următoarele două afirmații:

(1) Am încredere în tine datorită reputației tale bune.

(2) Am încredere în tine în ciuda reputației tale mai puțin bună.

Presupunem faptul că cele două afirmații se referă la o aceeași tranzacție care are loc între doi membrii A și B al unei comunități. Prima afirmație reflectă faptul că A are încredere în B pe baza reputației acestuia. A doua afirmație sugerează faptul că A are încredere în B, pe baza unor experiențe directe sau indirecte cu B, și acestea din urma primează în fața reputației pe care B o are. De aici rezultă o caracteristică pe care am discutat-o și în secțiunea anterioară: încrederea este un fenomen personal și subiectiv care se bazează pe diferiți factori. Afirmația (2) arată că experiența personală are mare greutate, dar în cazul în care aceasta nu există atunci trebuie să se facă apel la opiniile altora.

Reputația poate fi considerată o măsură colectivă a încrederii pe care membrii unei comunități o acordă unui anumit membru [Josang et. al., 2007].

În [Mui, Mohtashemi & Halberstadt, 2002] se identifică tipul de relații dintre conceptul de încredere și reputație, astfel:

- ♦ reputația unei entități A crește dacă în comunitatea în care ea este, crește și încrederea pe care o au alte entități în A;

- ♦ creșterea încrederii lui A în entitatea B crește șansa ca B să evalueze pozitiv pe A;

- ♦ creșterea numărului de astfel de acțiuni reciproce de evaluare pozitivă va duce automat în comunitate la creșterea reputației.

Astfel, reciprocitatea este definită ca schimb reciproc de evaluări (favorabile sau nu). Scăderea oricărei din cele trei variabile duce automat la efectul invers.

Putem observa de altfel cum definiția informală a reputației este translatată din lumea reală în lumea virtuală, și dacă în primul caz nu avem o formulă de calcul efectivă în cel de-al doilea context reputația poate exista doar datorită unei modalități de a o cuantifica. Aceste



doua concepte a caror caracteristici le-am prezentat mai sus, încredere și reputație vor sta la baza modelului discutat în secțiunea următoare.

### **Rolul unui model de calcul al încrederii și reputației**

Încă de la început facem observația că scopul acestei secțiuni este cel de a pune în evidență modul în care conceptele ca încrederea și reputația pot fi utilizate ca și elemente de bază în interacțiunile dintre indivizi aflați în comunități sociale online. De asemenea modelul prezentat va arăta că utilizatorii au posibilitatea de a accesa informații care le sunt cu adevărat relevante, asigurându-se astfel un drum mai ușor spre cunoaștere în sistemul informațional existent.

Există în literatura de specialitate propuneri de metrici locale de încredere și pot fi menționate: [Massa & Avesani, 2006], [Golbeck, 2005], [Ziegler & Lausen, 2004], [Zhili et. al., 2009]. În această secțiune vom discuta asupra modelului StarTrust [Alboaie & Vaida, 2011], [Alboaie, 2008], care integrează atât o metrică locală de încredere cât și un mecanism de calcul al reputației.

#### Modelare informală

În orice context factorul cel mai important care exercită un impact important asupra deciziei de a interacționa cu o altă persoană sau de a utiliza sau nu o resursă îl reprezintă încrederea. Încrederea este o relație binară dintre două entități [Massa, 2006], [Josang, 1996]. Încrederea se manifestă/exprimă prin declarații de încredere (engl. *trust statements*) care pot fi:

- ◆ directe: „Eu A, am încredere în B”;
- ◆ indirecte: „Eu A, folosesc resursa lui B fiindcă am încredere în punctul lui de vedere” sau „Eu A, citez acest articol despre stele Halo a lui B fiindcă el este expert în domeniul astronomiei”.

Aceste declarații de încredere sunt cuantificate în sistemul nostru folosindu-se valori furnizate (în mod explicit) sau calculate (în mod implicit).

Pe Web, în general, sau în comunitățile online, în particular, pot exista rețele foarte mari de utilizatori și atunci încrederea bazată doar pe experiența personală nu mai este suficientă. Astfel în modelul discutat avem o componentă numerică explicită ce reprezintă încrederea directă (valoare obținută din evaluarea explicită a utilizatorilor) dar și o componentă implicită calculată în sistem ce reprezintă încrederea indirectă.

Metrica locală discutată asigură în sistem faptul că un utilizator are propria sa viziune asupra utilizatorilor sau resurselor existente. În propunerea noastră pornim de la premisa că nu există proprietatea

de „bun” sau „rău” asociată datelor sau informațiilor. Când vom folosi această terminologie ne vom referi de fapt la: „informație bună din punctul de vedere al unui utilizator”, „informație neinteresantă din punctul de vedere al unui utilizator”. Astfel, informațiile poate fi folositoare pentru o mulțime de utilizatori, dar poate fi total inutilă pentru o altă mulțime. De asemenea nu vor exista utilizatori care au păreri greșite sau păreri bune, vom folosi terminologia: utilizatorul A are o opinie diferită față de utilizatorul B.

Sistemul propus oferă de asemenea și calcularea unei reputații globale în sistem, aceasta fiind extrem de utilă unui utilizator nou care devine membru al comunității. Inițial, un utilizator nu are o viziune locală asupra sistemului iar reputația globală îi poate oferi o orientare inițială în sistem spre atingerea scopurilor lui (găsirea de utilizatori sau informații cu un grad ridicat de interes pentru acesta). Pentru asigurarea echilibrului între „*echo chamber*” și „*tirania majorității*” modelul propus are și componenta de calcul a reputației.

În aceasta lucrare nu vom intra în detalii privind modelarea formală a sistemului de încredere și reputație StarTrust, pentru mai multe detalii privind aceasta se poate consulta [Alboai & Vaida, 2011]. Suntem însă interesați să urmărim exemple care să ne poată conduce la concluzii asupra importanței implicării conceptelor de încredere și reputație în modelarea interacțiunilor online între utilizatori, dar și în procesul de furnizare de informații utile acestora (care duce implicit la susținerea procesului de acumulare de cunoștințe de către aceștia).

### Exemplificare

Considerăm o comunitate formată din 15 utilizatori și 20 de resurse în care integram modelul de încredere și reputație. Facem observația că aceste valori mici sunt considerate astfel din motiv de a putea urmări și exemplifica prin figuri lizibile exemplul discutat.

Ratingul pe care un utilizator îl dă altui utilizator sau altei resurse măsoară gradul de apreciere a utilizatorului, respectiv resursei. Acest parametru poate lua valori în intervalele următoare, care au asociate următoarele semnificații: (0, 1] = *useless/spam*; (1, 2] = *poor*; (2, 3] = *worth attention*; (3, 4] = *good*; (4, 5] = *exceptional*.

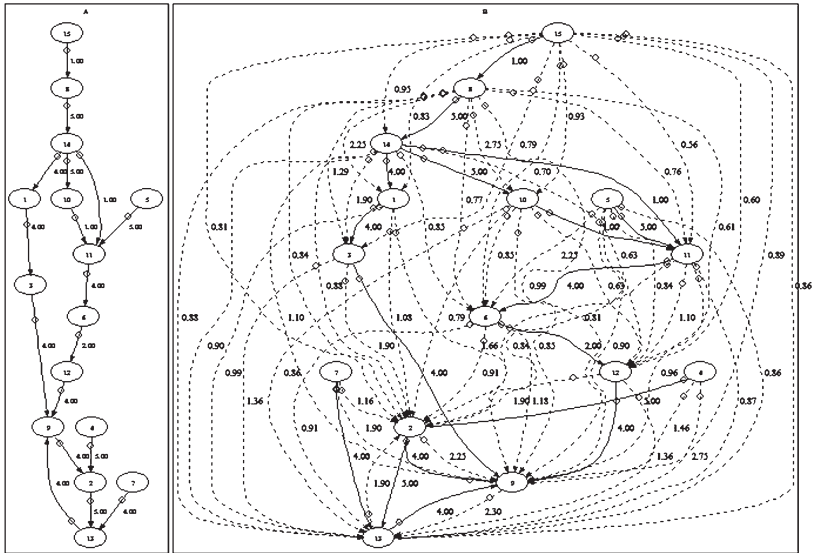
**Tabelul 1** – Evaluările explicite generate utilizator-utilizator

<b>ID Utilizator</b>	1	2	3	4	5	6	7	8	9	10	11	12	13	14	14	14	15
<b>ID Utilizator Evaluat</b>	3	13	9	2	11	12	13	14	2	11	6	9	9	1	10	11	8
<b>Rating</b>	4	5	4	5	5	2	4	5	4	1	4	4	4	4	5	1	1

**Tabelul 2** – Evaluările explicite generate utilizator-resursă

<b>ID Utilizator</b>	1	2	3	4	5	6	7	8	9	10	11	12	13	13	13	14	14	15	15	15	15			
<b>ID Resursa Evaluată</b>	20	11	16	16	13	18	14	9	2	12	14	2	6	4	8	9	8	12	16	5	6	13	14	20
<b>Rating</b>	4	4	3	3	5	4	4	5	4	2	2	4	4	4	4	4	5	1	1	4	4	5	5	2

Propagarea încrederii aplicând metrica de încredere va duce la obținerea următoarelor rezultate:



**Fig. 2** – Graful reprezintă comunitatea formată din 15 utilizatori; A este graful conținând ratingurile explicite; B este graful conținând ratingurile explicite și pe cele implicite; arcele continue reprezintă evaluările explicite, iar evaluările implicite sunt reprezentate cu arce punctate.

Folosind modelul de calcul al reputației obținem pentru cazul analizat valoarea reputației utilizatorilor din comunitate:

**Tabelul nr. 3** – Valorile reputației membrilor comunității

<b>ID Utilizator</b>	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
<b>Valoarea Reputației</b>	3,8	2,3	2,7	0	0	3	0	2	2,83	3,33	2	2	2,7	2	0

Să considerăm situația în care un utilizator nou A aderă la comunitatea de mai sus. Inițial, el nu are nici un fel de viziune proprie asupra resurselor din sistem. Pentru a putea accesa informații

într-un mod organizat, poate folosi viziunea generala (adica reputatia) asupra fiecărui utilizator din sistem. Dacă noul utilizator găsește date sau informații evaluate ridicat, dar de proastă calitate din punctul lui de vedere, atunci poate asocia un rating scăzut sa zicem utilizatorului B care a facut acea evaluare. Din modul în care are loc propagarea încrederii în sistem, se va asigura faptul ca sistemul nu va mai furniza utilizatorului A, ca și informații utile, pe cele evaluate foarte bine de B. Se asigură astfel o viziune personalizată a fiecărui utilizator din sistem.

Pentru exemplul nostru, modelul de încredere și reputație va calcula ierarhia de resurse pentru fiecare utilizator în parte.

**Tabelul nr. 4** – Viziunea asupra resurselor a fiecărui membru al comunității

ID Utilizator 1	20	16				
ID Utilizator 2	4	8	9	11		
ID Utilizator 3	3	16				
ID Utilizator 4	11	16				
ID Utilizator 5	14					
ID Utilizator 6	18	2	6			
ID Utilizator 7	14	4	8	9		
ID Utilizator 8	8	9	12	16		
ID Utilizator 9	3	11				
ID Utilizator 10	14	12				
ID Utilizator 11	18	14				
ID Utilizator 12	2	6	3			
ID Utilizator 13	4	8	9	3		
ID Utilizator 14	8	20	14	12	16	
ID Utilizator 15	13	14	5	6	9	20

### Aplicabilitatea modelului

Un model de încredere și reputație poate juca un rol important în orice context în care aspecte ca încrederea și reputația se pot manifesta și menționam doar cateva exemple: rețelele sociale, comunități științifice, comunități medicale [Alboaie & Buraga, 2009], etc. Ne vom opri cu discuția asupra rolului pe care un sistem de încredere și reputație îl poate avea într-un sistem de recomandare, și vom prezenta pe scurt o serie de rezultate obținute în urma unor experimente realizate cu ajutorul modelului StarTrust în aceste sisteme.

Sistemele de recomandare se încadrează în clasa sistemelor de filtrare a informației a căror scop este cel de a elimina informația

redundantă sau nedorită înainte ca aceasta să fie prezentată utilizatorului. Un sistem de recomandare are ca scop sugerarea utilizatorilor resurse (de exemplu cărți, imagini, pagini web etc.) care să fie potrivite cu preferințele acestora. Aceste recomandări se realizează pe baza unor tehnici, cea mai folosită fiind filtrarea colaborativă (eng. *collaborative filtering*) [Goldberg et. al., 1992].

Un algoritm de filtrare colaborativă realizează o estimare asupra ratingului pe care un utilizator l-ar putea da unei resurse (fără ca utilizatorul să fi evaluat explicit resursa respectivă). Aceste estimări se fac pe baza unor ratinguri anterior exprimate de utilizator.

În principal pașii de funcționare a unui sistem de recomandare bazat pe filtrare colaborativă urmează trei etape:

- 1) colectarea preferințelor;
- 2) găsirea utilizatorilor similari;
- 3) calcularea ratingurilor probabile;

Majoritatea sistemelor de recomandare utilizează mecanismul oferit de PC – *Pearson correlation* [Rodgers et. al., 1988] pentru a face recomandări și folosesc ca date de intrare ratingurile pe care utilizatorii le dau resurselor. Pe baza acestora se calculează coeficienții ce stabilesc o legătură între utilizatori. La pasul următor, sistemul de recomandare calculează un rating probabil pe care un utilizator l-ar asocia unei anumite resurse și pe baza acestor valori se pot face recomandări.

Discuția nu are ca scop detalierea diferitelor mecanisme de realizare a acestor trei pași. Dorim doar să creăm cadrul în care încrederea și reputația pot juca un rol esențial în asigurarea unei bune funcționari a sistemului.

Pentru aceasta, să considerăm două situații reale:

- ◆ integrarea rapidă a noilor utilizatori într-o comunitate care folosește un sistem de recomandare poate fi un proces îndelungat. Acest lucru este valabil și membrilor din comunitate care deși sunt mai vechi nu au făcut multe evaluări explicite resurselor. În [Massa & Avesani, 2006b] se arată, de exemplu, că în Epinion 53% din utilizatori au dat mai puțin de 5 ratinguri resurselor. Deci pentru acești utilizatori sistemul de recomandare are dificultăți în furnizarea de resurse potrivite, din lipsa de informații suficiente.

- ◆ promovarea unei resurse noi adăugate în sistem se face într-o perioadă lungă de timp.

Intuitiv, dacă avem un utilizator nou  $U$  care a evaluat o mulțime de resurse, atunci această evaluare nu se poate propaga. Dar dacă utilizatorul  $U$  a evaluat o mulțime de utilizatori, atunci pe baza unei

metrici de încredere va putea avea loc propagarea încrederii. În acest fel, în mod indirect se pot estima valori posibile pe care utilizatorul U le-ar putea asocia unui număr mult mai mare de resurse și în acest mod și sistemul de recomandare va fi mai eficient.

În studiul nostru vom efectua următoarea analiză:

- ◆ avem ca date de intrare:
  - un număr de ratinguri explicite asociate resurselor pentru fiecare utilizator în parte;
  - un număr de ratinguri explicite asociate altor utilizatori pentru fiecare utilizator în parte;
- ◆ avem ca date de ieșire:
  - folosind un sistem de recomandare bazat pe mecanismul oferit de PC, obținem numărul de utilizatori accesibili pentru fiecare utilizator;
  - folosind StarTrust obținem numărul de utilizatori accesibili pentru fiecare utilizator;
  - folosind StarTrust obținem numărul de resurse accesibile.

Pentru a exemplifica consideram o comunitate formată din 10 utilizatori și 200 de resurse.

**Tabelul nr. 5** – Rezultatele obținute în urma aplicării algoritmului de calcul bazat pe PC și a StarTrust

ID Utilizator	1	2	3	4	5	6	7	8	9	10
Nr. ratinguri Utilizator-Utilizator <sup>1</sup>	2	3	3	2	2	2	3	2	2	3
Nr. ratinguri Utilizator-Resursă <sup>2</sup>	9	6	9	8	6	9	8	6	17	20
PC – Nr. utilizatori accesibili <sup>3</sup>	1	2	1	1	1	1	1	1	2	3
PC – Nr. resurse accesibile <sup>4</sup>	9	24	9	8	6	9	8	6	34	37
StarTrust – Nr. utilizatori accesibili <sup>5</sup>	9	9	9	9	9	9	9	9	9	9
StarTrust – Nr. resurse accesibile <sup>6</sup>	25	31	39	18	23	36	34	31	32	42

**Legendă:**

(1) numărul de ratinguri pe care le-a dat în mod explicit fiecare utilizator altor utilizatori; (2) numărul de ratinguri care le-a dat în mod explicit fiecare utilizator unor resurse; (3) numărul de utilizatori accesibili aplicând algoritmul de calcul bazat pe PC; (4) numărul de resurse accesibile aplicând algoritmul de calcul bazat pe PC; (5) numărul de utilizatori accesibili aplicând StarWorth; (6) numărul de resurse accesibile aplicând StarWorth

Analizând rezultatele obținute pentru comunitatea formată din 10 utilizatori și 200 de resurse, obținem următorul grafic reprezentând numărul de utilizatori cu care există o legătură în comunitate folosindu-se: algoritmul bazat pe PC standard și StarTrust.

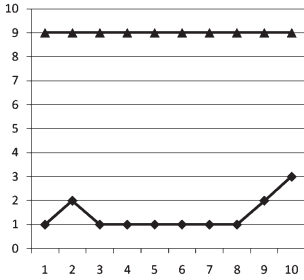


Fig. 3 – Comparație între numărul de utilizatori accesibili folosind algoritmul bazat pe PC și StarTrust

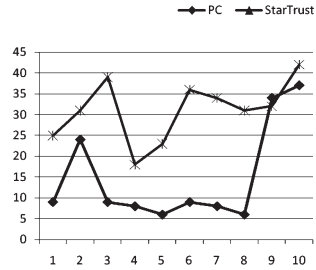


Fig. 4 – Comparație între numărul de resurse accesibile

Extindem experimentul descris anterior și vom rula algoritmul bazat pe PC și StarTrust asupra mai multor instanțe de comunități online formate din 10 de utilizatori și 200 de resurse.

Furnizăm mai jos graficele care arată legătura între numărul de evaluări date și numărul de utilizatori, respectiv resurse accesibile folosind algoritmul bazat pe PC și StarTrust. Graficele de mai jos a fost obținute în urma rulării algoritmilor pe 100 de instanțe de comunități online.

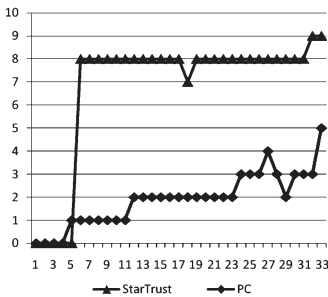


Fig. 5 – Comparație între numărul de utilizatori accesibili folosind algoritmul bazat pe PC și StarTrust

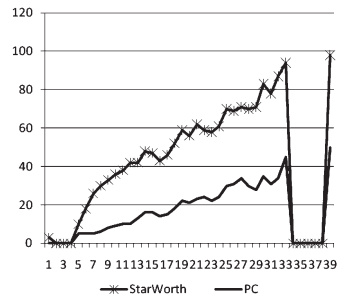


Fig. 6 – Comparație între numărul de resurse accesibile (pe axa Ox sunt reprezentate numărul de note de tip utilizator-resursă)

Argumentele aduse anterior conform căreia introducerea unui mecanism de propagare a încrederii conduce într-un sistem de recomandare la un acces mai rapid la un număr mai mare de utilizatori și resurse sunt susținute de experimentele realizate [Alboaie & Vaida, 2010].

## 5. Concluzii

Prelucrarea și transmiterea datelor între membrii societății umane a ridicat mereu probleme de credibilitate și oportunitate care mult timp au fost tratate doar într-un mod intuitiv.

De exemplu, suportul pe care stau activitățile economice sunt banii sau alte forme de mijlocire a schimbului de bunuri și servicii. Esențial ca aceste schimburi să aibă loc, este existența încrederii între participanții la tranzacții până la nivelul în care valorile materiale pot fi înțelese ca o măsură obiectivă a încrederii între oameni și societate.

Credem că inovațiile și aprofundarea înțelegerii în domeniul mecanismelor care stau la baza încrederii pot avea un impact important asupra activităților economice, sociale sau chiar a fundamentului moral pe care stă societatea noastră.

Problema reprezentării și folosirii informației despre încredere își face apariția chiar și în bazele biologice ale animalelor (dacă vedem genele ca un mecanism de comunicare de informații).

Lucrarea de față realizează o incursiune în domeniul prelucrării și transmiterii datelor în societatea umană, focalizându-se cu precădere asupra aspectelor legate de transmiterea datelor pe Web în cadrul diverselor comunități online. Aspectele legate de încredere și reputație sunt analizate în vederea integrării lor într-un proces de selecție care să aibă un impact pozitiv în luarea unor decizii majore.

Enumeram o serie de consecințe imediate, obținute din modul în care sistemul de încredere și reputație funcționează:

- ◆ Informațiile relevante pentru un utilizator (chiar și cele noi) sunt vizibile în topul listei de resurse
- ◆ Sistemul asigură că utilizatorul vede resursele prioritizate, în funcție de preferințele sale
- ◆ Utilizatorii care adaugă informații de tip spam, vor vedea informații de tip spam, fiindcă sistemul grupează utilizatorii în funcție de preferințele lor
- ◆ Utilizatorii sunt încurajați să facă evaluări corecte, și nu se va întâmpla ca în sisteme cum ar fi eBay unde multe din evaluările pozitive vin din teama unei potențiale răzbunări.

În literatura de specialitate, modelările actuale, ca și modelul tratat în cadrul lucrării au la bază o singură dimensiune a încrederii. Ca direcție viitoare de analiză, vom încerca să identificăm aspecte privind dihotomia conceptului de încredere în contexte diverse.



**Bibliografie**

- [1] Ackoff, R. L., „*From Data to Wisdom*”, Journal of Applies Systems Analysis, Volume 16, 1989, p. 3–9.
- [2] Amrit Tiwana, *The Essential Guide to Knowledge Management – E – Business and CRM Applications*, Prentice – Hall, 2001.
- [3] Alboaie, L., *PReS – Personalized evaluation System in a WEB Community*, ICE-B 2008: Proceedings of the Int. Conf. on E-Business, 2008, p. 64–69.
- [4] Alboaie, Lenuța, Vaida, Mircea-F., *Trust and Reputation Model for Various Online Communities*, Studies în Informatics and Control, ISSN 1220–1786, June 2011, Vol. 20, Issue 2, pp. 143–156
- [5] Alboaie, Lenuța, Vaida, Mircea-F., *Modeling of Trust to Provide Users Assisted Secure Actions in Online Communities*, NDT 2010, Network Digital Technologies, Prague, 7–9 July 2010, pp.369–382, Communications in Computer and Information Science Part I, (CCIS) of Springer Lecture Notes Series – 87 ([www.springer.com/series/7899](http://www.springer.com/series/7899)), ISI Proceedings and Scopus, ISSN 1865–0929, ISBN-10 3–642–14291–5, ISBN-13 978–3–642–14291–8
- [6] Alboaie, Lenuța, Buraga, Sabin, *Trust and Reputation in e-Health Systems*, IFMBE Proceedings Vol. 26., International Conference on Advancements of Medicine and Health Care through Technology, Meditech09, 23–26 September 2009, pp. 257–260, ISSN 1680–0737, ISBN 978–3–642–04291–1, e-ISBN 978–3–642–04292–8, DOI 10.1007/978–3–642–04292–8, Library of Congress Control Number: 2009934354, Cluj-Napoca, Romania, Indexare Springer
- [7] Baader, Franz, Deborah L. McGuinness, Daniele Nardi, Peter F. Patel-Schneider, *The description logic handbook: Theory, implementation, and applications*, 2007
- [8] Bellinger, Gene; Castro, Durval; and Mills, Anthony. *Date, Information, Knowledge, and Wisdom*, 2004, <http://www.outsights.com/systems/dikw/dikw.htm>.
- [9] Berners-Lee, Tim, Mark Fischetti, *Weaving the Web: The Original Design and Ultimate Destiny of the World Wide Web by its inventor*, Britain: Orion Business, 1999
- [10] Cook, K. (ed.), *Trust in Society*, New York, Russell Sage Foundation, 2001
- [11] Davenport, T.H. and L. Prusak, *Working Knowledge: How Organizations Manage What They Know*, Harvard Business School Press, Boston, MA. 1998
- [12] Drăgănescu, Mihai, *Profunzimile lumii materiale*, Bucuresti, Editura Politica, 1979
- [13] Gambetta, D., *Can We Trust Trust?*, în D. Gambetta, editor, *Trust: Making and Breaking Cooperative Relations*, pages 213–238. Basil Blackwell, Oxford, 1990
- [14] Garvin, Andrew, P., *The Art of Being Well Informed – What You Need To Know To Gain The Winning Edge in Business*, Avery Publishing Group, 1996
- [15] Goldberg, D., Nichols, D., Oki, B. M. Terry, D., *Using collaborative filtering to weave an information tapestry*, Communications of the ACM, Volume 35, Issue 12 (December 1992), Special issue on information filtering, 61–70, 1992
- [16] Golbeck, J., *Computing and Applying Trust in Web-based Social Networks*, PhD thesis, University of Maryland, 2005
- [17] Hardin, R., *Trust & Trustworthiness*, New York: Russell Sage Foundation, 2002
- [18] Harmon, A., *Amazon Glitch Unmasks War Of Reviewers*, The New York Times, 14 February 2004
- [19] Josang, A., *The Right Type of Trust for Distributed Systems*, Proceedings of the 1996 New Security Paradigms Workshop, 1996

- [20] Josang, A., Ismail, R., Boyd, C., *A survey of trust and reputation systems for online service provision*, Springer, 2007
- [21] Von Krogh, G., Ichijo, K., and Nonaka, I., *Enabling Knowledge Creation: How to Unlock the Mystery of Tacit Knowledge and Release the Power of Innovation*, Oxford University Press, New York, NY, USA, 2000
- [22] Liew, Anthony, *Understanding Data, Information, Knowledge And Their Inter-Relationships*, Journal of Knowledge Management Practice, Vol. 8, No. 2, June 2007
- [23] Massa, P., *Reputation is in the eye of the beholder: on subjectivity and objectivity of trust statements*, ENISA eID Workshop, May 2007;
- [24] Massa, P., Avesani, P., *Trust-aware Bootstrapping of Recommender Systems*, Proceedings of the ECAI'06 Workshop on Recommender Systems, 2006
- [25] Mui, L., Mohtashemi, M., Halberstadt, A., *A computational model of trust and reputation*, System Sciences, 2002, HICS. Proceedings of the 35th Annual Hawaii International Conference on, pp. 2431–2439, 2002
- [26] Rahman, A., Hailes, S., *Virtual Communities*, System Sciences, 2000. Proceedings of the 33rd Annual Hawaii International Conference, 2000
- [27] Rodgers, J. L., Nicewander, W. A., *Thirteen ways to look at the correlation coefficient*, The American Statistician 42: 59–66, 1988
- [28] Sunstein, Cass, *Republic.com*. Princeton University Press, 1999
- [29] Surowiecki, *The Wisdom of Crowds: Why the Many Are Smarter Than the Few*, 2005
- [30] Tocqueville, A., *Democracy in America*, Doubleday, New York, 1840
- [31] Toffler, Alvin, *Șocul viitorului*, Editura Politică, București, 1973
- [32] Zhili Wu, Xueli Yu, Jingyu Sun, *An Improved Trust Metric for Trust-Aware Recommender Systems*, First Int. Workshop on Education Technology and Computer Science, 2009, Vol. 1: 947–951.
- [33] Ziegler, C., Lausen, G., *Spreading activation models for trust propagation*, Proceedings of the IEEE International Conference on e-Technology, e-Commerce, and e-Service, Taipei, Taiwan, 2004