

# Analiza imaginarii din texte

**Lidia Trăușan-Matu**

Universitatea București,  
Facultatea de Istorie

Bd. Regina Elisabeta 4-12,  
lidia@racai.ro

**Valentin-Andrei Canciu**

Universitatea „Politehnica”  
București

Splaiul Independenței nr.  
313, București

**Ștefan Trăușan-Matu**

Universitatea „Politehnica”  
București

Splaiul Independenței, nr.  
313, București  
trausan@gmail.com

## REZUMAT

În lucrare este abordată problematica detectării automate a imaginarii din texte. Problema este foarte importantă deoarece rezolvarea ei permite identificarea conotațiilor unor texte, cu implicații majore asupra analizei interacțiunii textuale. Abordarea pleacă de la teoria imaginarii introdusă de istoricul Lucian Boia. Sunt folosite tehnici de prelucrare a limbajului natural. Au fost făcute și sunt prezentate în lucrare mai multe analize de texte de referință.

## Cuvinte cheie

Imaginar, Prelucrarea limbajului natural, înțelegerea textelor, conotații, colocații, wordnet

## Clasificare ACM

I.2.7 Natural Language Processing. J.5 Arts and Humanities

## INTRODUCERE

Interfațarea om-calculator în limbaj natural, un subdomeniu al înțelegerii limbajului natural este încă departe de a fi rezolvată [6], datorită mai multor cauze, una din ele fiind dificultatea detectării conotațiilor, a înțelesurilor ascunse în texte. În această direcție, în orice text un rol extrem de important în înțelegere îl au structurile arhetipale care formează imaginarii.

Cercetările în prelucrarea limbajului natural a considerat până acum în special aspectele lingvistice, dar a neglijat aspectele antropologice, istorice și contextuale. Lucrarea de față prezintă niște cercetări interdisciplinare, în care se încearcă folosirea tehnicilor de prelucrare a limbajului natural pentru a detecta dimensiunile imaginarii predominante într-un text. Importanța acestui demers rezultă din posibilitatea folosirii acestor informații pentru a crea profilul sau intențiile celui care emite un text și a adapta în consecință dialogul cu acesta.

Lucrarea continuă cu prezentarea principalelor definiții ale imaginarii și, în special cu teoria istoricului Lucian Boia [1]. Secțiunea care urmează introduce câteva dintre ideile și tehnicile de prelucrare a limbajului natural folosite. A patra secțiune analizează rezultatele obținute în cazul mai multor texte și a mai multor variante de analiză.

## O DEFINIȚIE A IMAGINARULUI

De-a lungul timpului au fost propuse mai multe definiții pentru imaginar. Potrivit Evelynei Patlagean, “domeniul imaginarii este constituit din ansamblul reprezentărilor care depășesc limita impusă de constatările experienței și de înlăturirile deductive autorizate de acestea” [1, pag.12]. Problema acestei definiții este imposibilitatea stabilirii unei delimitări clare între realitate și imaginar.

O alternativă ne este oferită de antropologul francez Gilbert Durand într-o lucrare devenită clasică [2]. Acesta afirmă că imaginarii este „ansamblul imaginilor și al relațiilor dintre imagini care constituie capitalul gândit al lui *homo sapiens*” și că el „ne apare ca marele denominator fundamental la care vin să se ralieze toate procedurile gândirii umane” [2, pag.19].

Jacques Le Goff evită să dea o definiție directă a imaginarii, preferând să evidențieze ce nu este imaginarii. Astfel, potrivit lui, imaginarii nu trebuie confundat nici cu “reprezentarea” realității exterioare, nici cu “simbolicul”, nici cu “ideologia”. Lucian Boia respinge o astfel de limitare, în primul rând deoarece nu există o reprezentare identică cu obiectul reprezentat (“reprezentarea” realității exterioare), în al doilea rând pentru că el consideră universul “simbolurilor” ca aparținând exclusiv imaginarii, și în cele din urmă deoarece chiar “ideologiile” pot fi interpretate ca mitologii secularizate.

Lucian Boia propune reconcilierea celor două tipuri de abordare a imaginarii ce vin dinspre discipline unde acest concept cultural a fost utilizat: pe de o parte antropologia, în special școala lui Gilbert Durand care pune accentul pe anumite structuri atemporale creatoare de imaginar (arhetipurile), dar care neagă istoricitatea imaginarii; iar pe de altă parte istoria, care abordează imaginarii într-un mod fragmentat (istoricilor li se datorează istorii ale imaginarii și nu o istorie a imaginarii) [1].

Mediind între cele două modele epistemologice, Lucian Boia reia trăsăturile esențiale ale conceptului în încercarea de a demonstra că „imaginarii este un produs al spiritului” și că „modul nostru de cunoaștere a lumii, rațiunea noastră și știința noastră se hrănesc din imaginar”. În esență, consideră autorul, imaginarii „constituie o realitate independentă, dispunând de propriile sale structuri și de propria sa dinamică ... și

poate fi folosit ca un barometru foarte sensibil al evoluției istorice” [1, pp. 14 și 26].

### **Structuri arhetipale ce descriu imaginarul**

Plecând de la ideea că există structuri durabile, Lucian Boia propune opt ansambluri sau structuri arhetipale care pot să acopere esențialul imaginarului.

#### *Conștiința unei realități transcendente*

Conștiința unei realități invizibile, insesizabile însă cu atât mai importantă și mai purtătoare de semnificații față de realitatea tangibilă. De acest prim arhetip țin supranaturalul, miraculosul și mai ales sacrul, el definind o caracteristică mentală intrinsecă condiției umane. Sacrul este prezent într-o mare diversitate de sisteme mitice de la cele mai simple până la cele mai complexe (religii monoteiste și politeiste).

Acest arhetip este aparent într-un declin în lumea modernă desacralizată, însă numai aparent deoarece pe rând știința, națiunea, rasa, sexul, “viitorul luminos” au fost sacralizate. Caracteristic pentru acest fenomen este proliferarea sectelor (care se strâng în jurul noilor adevăruri) sau succesul parapsihologiei care caută în om puteri miraculoase. De asemenea, dacă în trecut sacrul promova o categorie de *aleși*, oameni destinați a servi ca intermediari între lumea obișnuită și lumea transcendentă (regi, preoți, regi-zei, regi-preoți), acum acești *aleși* au fost înlocuiți de conducători *carismatici* de *staruri de campioni*.

#### *“Dublul”, moartea și viața de apoi*

Acest arhetip descrie convingerea că trupul material al omului este dublat de un element independent și imaterial (spirit, suflet etc.) Acesta este nemuritor și se poate desprinde de corp chiar în timpul vieții. Spiritul poate ajunge într-un loc deschis lumii celor vii (cultul strămoșilor la popoarele primitive), fie într-un loc îndepărtat și inaccesibil (Hadesul la greci). El poate fi răsplătit pentru meritele din timpul vieții (creștinism) sau nu (infernul clasic); sau, poate rămâne legat de lumea materială prin reîncarnări succesive. Nu există limite precise între lumea celor vii și a celor morți, acestea putând fi depășite, anumiți *aleși* putând străbate dincolo de această limită (călătorie inițiativă, extaz mistic); spiritele pot fi chiar contactate, credința perpetuată din vechime, însă prezentă și acum sub forma spiritismului.

#### *Alteritatea. Legătura dintre Eu și Ceilalți dintre Noi și Ceilalți*

Acest arhetip funcționează între limite foarte largi, de la diferențe minore, până la alteritatea radicală când Celălalt este împins dincolo de limitele umanului într-o zonă apropiată de divin sau de animalitate. Generalizând, alteritatea se poate referi la un întreg ansamblu de diferențe: spații, societăți, ființe, peisaje diferite.

#### *Unitatea*

Acest arhetip este reflectarea năzuinței oamenilor de a supune lumea unui principiu unificator prin care să poată fi explicat orice fenomen, pentru a crea imaginea unui univers omogen și inteligibil. Religiiile, filozofiile, științele, istoria, ideologiile, toate încearcă să creeze un sistem care să dea coerență diversității fenomenelor.

Unitatea se manifestă la orice nivel, de la legile ce guvernează Universul, la scara comunităților umane cărora o serie de mituri trebuie să le asigure coerența.

#### *Actualizarea originilor*

Originile sunt puse la mare preț în orice comunitate umană. Miturile de origine au rolul de a arunca o punte între trecut și prezent, de a da specificitate unui grup uman și de a-i conferi garanția unei perenități. Se poate referi la originile Universului (cosmogonie), ale omului, ale religiilor, națiunilor, statelor etc.

#### *Descifrarea viitorului*

Acest arhetip cuprinde o mare varietate de metode și practici ce urmăresc cunoașterea și controlul viitorului. Se poate referi la destinul individual al unui om, sau la destinul omenirii și al lumii (incluzând și sensul istoriei). Ocultismul, astrologia, profețiile arată o căutare obsesivă și nesatisfăcută.

#### *Evadarea, consecința refuzului condiției umane și a istoriei*

Omul aspiră să se elibereze de orice constrângeri și să se transforme în orice variantă imaginabilă: ascensiune spirituală, cunoaștere, puteri supranaturale și supraomenești, sfințenie sau regresie (către natura primordială). Se caută soluții fie în trecut, prin exaltarea începuturilor (nostalgia vârstei de aur) fie într-un viitor purificat (ideologii seculare sau religioase). Ca loc, acestea se pot manifesta fie în spațiul cotidian (utopii) fie pe tărâmurile îndepărtate (insule, planete, galaxii).

#### *Lupta (și complementaritatea) contrariilor*

Imaginarul este polarizat. Dialectica contrariilor este caracteristica tuturor religiilor și interpretărilor lumii, omului și ale istoriei. Polii opuși pot fi supuși fie principiului Unității fie disociați în părți contradictorii (idealism/materialism, clasicism/romantism).

## **DESCRIEREA GENERALĂ A APLICAȚIEI**

Cercetarea prezentată în lucrare a avut drept scop investigarea posibilității analizei asistate de calculator a structurii imaginarului dintr-un text prin prisma celor opt structuri arhetipale propuse de Lucian Boia. Pentru aceasta, mai întâi s-a construit un repertoriu de concepte pentru domeniul imaginarului, o așa numită ontologie (în termenii sistemelor informatice bazate pe cunoștințe). Această ontologie a fost creată plecând de la un set inițial de cuvinte relevante, care este apoi extins pe baza descoperirii de cuvinte similare folosind ontologia lexicală WordNet [5].

Analiza textelor investigate începe cu unele prelucrări prealabile, care elimină unele elemente nerelavante. În continuare, programul studiază, folosind relațiile de similaritate din WordNet, în ce măsură textul de analizat se apropie de fiecare din cele opt structuri ale imaginarului. Pentru a realiza acest lucru, mai întâi programul extrage locațiile (perechi de cuvinte care apar de obicei împreună, de exemplu, pentru una din cărțile analizate în lucrare: “working class”, “bourgeois society”, “bourgeois property”) din documentul de analizat.

Analiza textului este efectuată asupra colocațiilor și nu asupra cuvintelor individuale (se putea folosi în analiză și frecvența individuală a cuvintelor din text) deoarece s-a considerat că astfel se ajunge la rezultate mai interesante (putem mări semnificația unei colocații față de una – sau mai multe - din structurile imaginarului dacă în componența acesteia intră mai multe cuvinte ce manifestă aceeași apropiere față de una – sau mai multe – din cele opt structuri).

Pentru extragerea colocațiilor, programul a folosit pachetul NLTK [4] pentru prelucrarea limbajului natural, implementat în limbajul Python. Pentru determinarea părții de vorbire a unui cuvânt din document (determinare necesară pentru procesul de filtrare a colocațiilor după tipul părților de vorbire al cuvintelor componente) a fost folosit un adnotator ("tagger") antrenat pe o porțiune din corpus-ul Brown având același gen cu al documentului de analizat.

Colocațiile extrase sunt formate din două sau trei cuvinte (bigrame și trigrame) și au fost determinate la nivelul grupurilor de cuvinte folosind o fereastră (o secvență) de șase cuvinte deplasată de-a lungul fiecărei fraze. A fost necesară separarea în grupuri deoarece altfel ar fi putut apărea colocații formate din grupuri de cuvinte care nu se găsesc alăturate în document, alăturarea fiind una artificială dată de prezența unor cuvinte (din componența colocației) la sfârșitul unei fraze și de prezența altora (din componența colocației) la începutul frazei următoare.

Modul în care programul evaluează prezența în document a fiecărei structuri din cele opt ale imaginarului este următorul: Se procesează pe rând fiecare colocație, în ordinea descrescătoare a numărului de apariții și i se atribuie un scor. Acest scor este un vector cu opt componente, reprezentând scorul sau apropierea acestei colocații față de fiecare din cele opt structuri ale imaginarului. Scorul pentru fiecare colocație este adunat (pe componente) într-un vector ce conține scorul global pentru întreg documentul. Modul de calcul pentru scorul fiecărei colocații ia în calcul scorul pentru fiecare cuvânt component (care este amplificat pentru a scoate în evidență vârfurile – scorul foarte mare pentru anumite componente – și a atenua restul componentelor), ia în calcul prezența în colocație a mai multor cuvinte ce conțin vârfuri de semnificație în aceleași poziții ale vectorului de scor (scorul din aceste poziții va fi amplificat suplimentar) precum și frecvența de apariție a acestei colocații în document.

Apropierea unui cuvânt față de una din cele opt structuri arhetipale este determinată folosind WordNet astfel: se calculează similaritatea cuvântului de analizat față de fiecare cuvânt din toate cele opt categorii (cuvinte din ontologia imaginarului având aceeași parte de vorbire ca a cuvântului de analizat). Scorul față de o structură a imaginarului este o medie a primelor patru similarități (în ordine descrescătoare) față de cuvintele (având aceeași parte de vorbire cu a cuvântului de analizat) ce descriu acea structură a imaginarului. Programul dispune de o funcție ce afișează un grafic ce descrie scorul pentru fiecare din cele opt structuri.

Ieșirea furnizată de acest program a fost apoi studiată, selectând manual acele cuvinte care păstrează sensul comun al cuvintelor din lista de intrare; aceste cuvinte selectate manual vor fi adăugate la lista de intrare, urmând o nouă rulare a scriptului folosind această nouă listă de intrare.

## REZULTATELE ANALIZEI

S-a efectuat analiza a nouă documente scrise în limba engleză, care au fost descărcate de pe web:

- Karl Marx, "Manifestul Partidului Comunist"
- Nisargadatta Maharaj, "I Am That" (convorbire cu un guru indian)
- Rachel Cooper, Calsificând nebunia ("Classifying Madness", o carte științifică despre bolile psihice)
- C.S. Lewis, "Mere Christianity" (carte de apologetică a creștinismului)
- Andrew McDonald, "The Turner Diaries" (carte rasistă, împotriva populației de culoare, din anii '70)
- Adolf Hitler, "Mein Kampf"
- Diane Stein, "We Are Angels" (carte despre vindecare/spiritualitate)
- Karl Marx, "Capitalul"
- Secțiunea "D" ("RELIGION") a corpus-ului Brown

Asupra fiecărui document a fost efectuată o analiză folosind cinci variante de calcul a parametrilor pentru funcțiile de amplificare ale anumitor valori din program, după cum urmează:

1. Nu se efectuează nici o amplificare asupra valorilor
2. Se anulează amplificarea asupra factorului dat de deviația vectorului de scor; se acționează asupra scorului (astfel neamplificat) și se păstrează doar componentele foarte mari din scor.
3. Se acționează doar asupra factorului dat de deviația vectorului, diminuând amplificarea asupra componentelor vectorilor cu o deviație foarte mare, și mărinnd foarte mult amplificarea asupra componentelor vectorilor cu o deviație foarte mică (puține vârfuri, restul valori similare).
4. Se elimină complet amplificarea dată de prezența în colocație a unor cuvinte similare (în ce privește similaritatea față de aceleași componente ale ontologiei imaginarului)
5. Se amplifică foarte mult colocațiile ce au în ele cuvinte ce au un factor de similaritate între 0.4 și 0.5 (minimul este 0 și maximul 1).

Rezultatele obținute au fost reprezentate grafic, în corespondență cu cele 8 dimensiuni arhetipale ale imaginarului (notate în figurile următoare astfel: alteritate – "alterity", dublul – "double", evadare – "escape", lupta – "fight", viitorul – "future", originile – "origins", transcedența – "transcend" și unitate – "unity"). Pentru fiecare dimensiune este reprezentată o bară verticală de înălțime proporțională cu scorul obținut de ea în textul analizat.

Pentru a obține cea mai bună analiză s-au încercat toate cele 5 variante de prelucrare a parametrilor prezentate mai sus. În acest mod se vizualizează perspectiva ce mai relevantă, similar cu colorarea în diverse moduri a unor preparate puse pe lamelele pregătite pentru a fi vizualizate la microscop cât mai bine anumite detalii.

În continuarea lucrării vom prezenta o selecție din graficele rezultate în urma analizei asupra celor nouă texte, cu cele mai relevante cazuri din cinci variante de factori de amplificare.

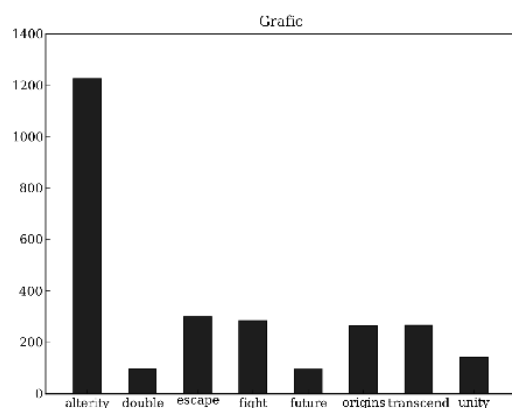


Figura 1. Mein Kampf, varianta 1

În "Mein Kampf", cartea de o tristă celebritate scrisă de Adolf Hitler, în urma prelucrărilor folosind varianta 1, se observă o valoare foarte mare pentru alteritate (conform figurii 1). Această caracteristică a fost găsită și în alte texte rasiste analizate. Următoarele două valori ca mărime aparțin dimensiunilor pentru luptă și evadare (caracteristice unor ideologii utopice).

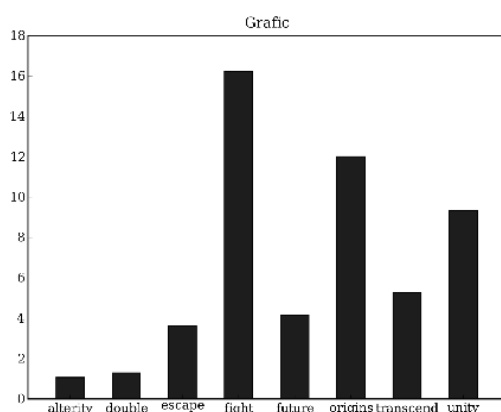


Figura 2. Mein Kampf, varianta a 2-a

În varianta a 2-a, în care se face o prelucrare care nu ține cont de deviația vectorului și păstrează doar componentele mari ale vectorului neamplificat, putem vedea că "lupta" ("fight") obține un scor considerabil mai mare (figura 2).

Aceeași valoare foarte mare a alterității o constatăm și în cealaltă carte rasistă (împotriva populației de culoare, din anii 1970) analizată, "The Turner Diaries", folosind varianta 1 (figura 3).

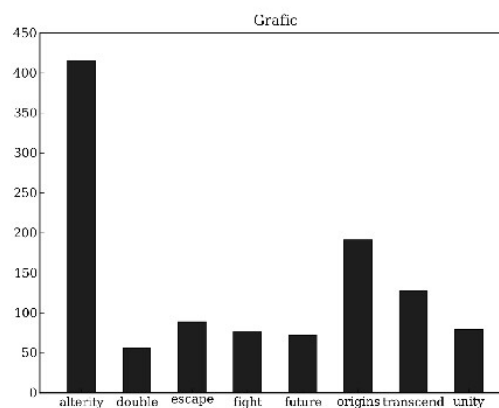


Figura 3. The Turner Diaries, varianta 1

Păstrând doar componentele mari ale scorului (și eliminând amplificarea dată de deviație) putem observa creșterea dimensiunii "evadării", în varianta a 2-a de analiză a aceleiași lucrări (figura 4).

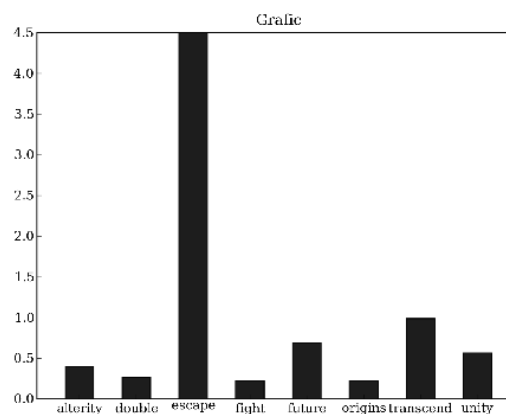


Figura 4. The Turner Diaries, varianta a 2-a

În "We Are Angels", carte spirituală, putem observa o valoare mare pentru transcendență (figura 5). În acest caz putem observa și o valoare mare pentru unitate (dorința de a supune toate lucrurile unui principiu unificator).

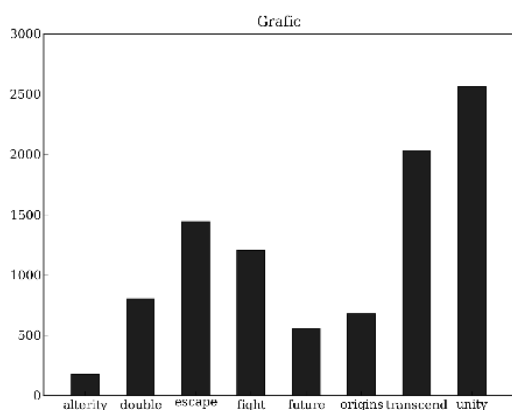


Figura 5. We Are Angels, varianta 1

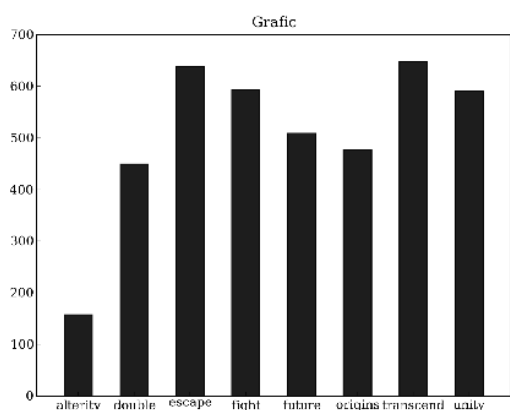


Figura 6. We Are Angels, varianta a 4-a

Varianta a 4-a (figura 6) elimină complet amplificarea dată de prezența în colocație a unor cuvinte similare obținem rezultate mult mai puțin interesante (mai puțin diferențiate).

În Manifestul Partidului Comunist, o ideologie utopică, putem observa valori foarte mari pentru evadare și luptă (varianta 1, figura 7).

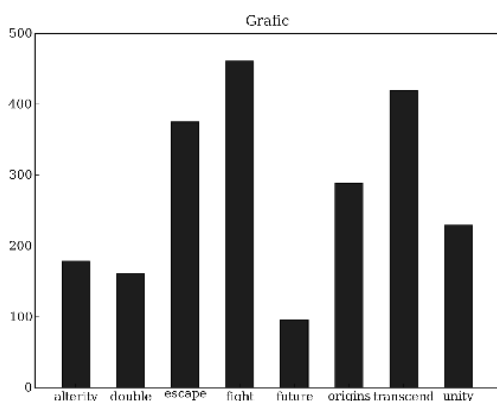


Figura 7. Manifestul Partidului Comunist, varianta 1

Păstrând doar componentele mari ale scorului și eliminând factorul dat de deviație, putem observa o prezență constantă a valorilor mari pentru luptă (figura 8).

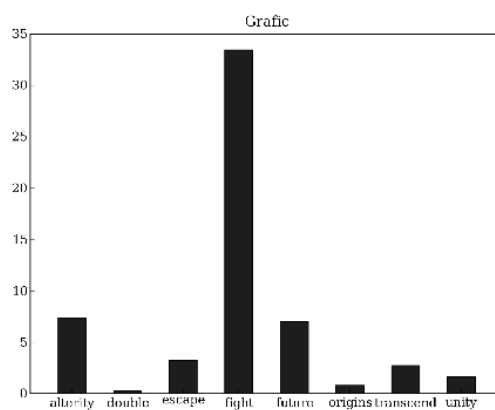


Figura 8. Manifestul Partidului Comunist, varianta a 2-a  
În 'I Am That' (convorbiri cu un guru spiritual indian), conform variantei 1 obținem graficul din figura 9.

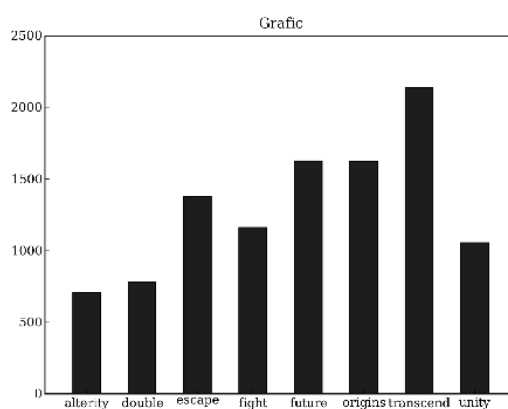


Figura 9. I Am That, varianta 1

Ca și în cazul altor cărți spirituale analizate, observăm o valoare mare pentru transcendență, dacă alegem varianta 3 a amplificării (figura 10).

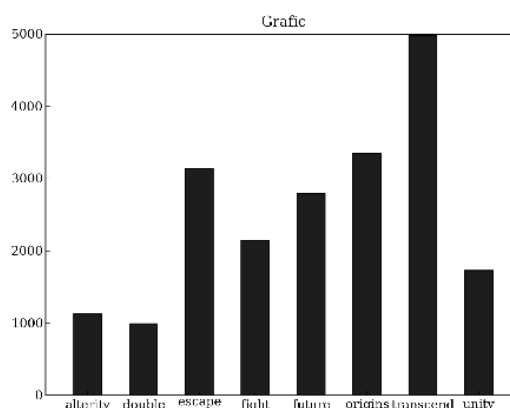


Figura 10. I Am That, varianta 3

Amplificând scorurile pentru vectorii cu deviație foarte mică putem observa o creștere substanțială a transcendenței. Acest lucru poate indica faptul că multe cuvinte fac parte net din sfera transcendentului (valoare mult mai mare de similaritate față de acesta, restul valorilor fiind toate coborâte).

“Capitalul” de Karl Marx, varianta 1, acoperă cele 8 dimensiuni după cum este ilustrat în figura 11.

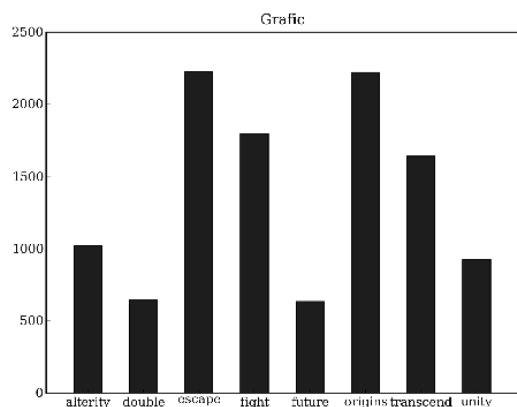


Figura 11. Capitalul, varianta 1

Trebuie remarcat faptul că graficul pentru ”Capitalul” este aproape identic cu cel al “Manifestul Partidului Comunist”, cealaltă carte a lui Marx analizată, evadarea, lupta, originile și transcendența fiind primele patru valori.

”Clasificând nebunia”, carte științifică despre bolile psihice, conform variantei 1 are graficul din figura 12.

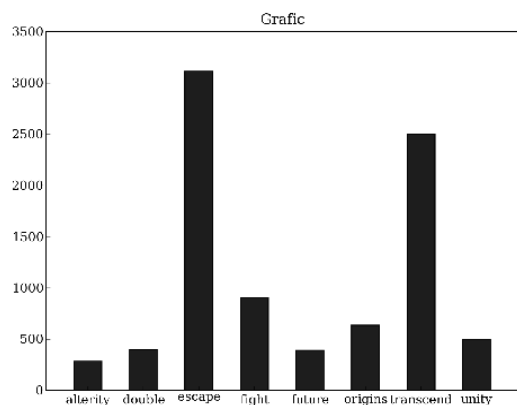


Figura 12. Clasificând nebunia, 1

Se observă o valoare foarte mare pentru evadare. În acest text ea este dată (în mare parte) de frecvent folosită expresie “natural kinds” – un concept filosofic folosit aici în studiul științific al bolilor psihice. Această creștere poate fi văzută ca artificială, textul neavând neapărat astfel de conotații. Poate aceasta este o limitare a metodei folosite, imaginarul nepretându-se evaluării textelor pur științifice fără o analiză mai amănunțită a conținutului acestora. Pentru a discerne între texte științifice într-un anumit domeniu ar trebui probabil să construim o ontologie mai strâns legată de acel domeniu.

Pentru ”Mere Christianity” (carte de apologetică a creștinismului a lui C.S. Lewis) avem graficul din figura 13.

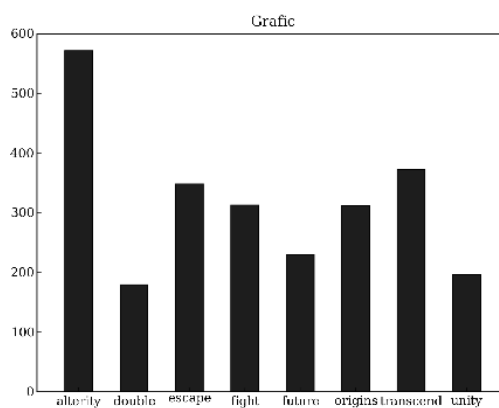


Figura 13. Mere Christianity

Observăm o valoare mare pentru alteritate; valoarea următoare este pentru transcendență și apoi evadare. Valoarea mare pentru alteritate poate fi explicată ca fiind caracteristică pentru creștinism - creștinii fac parte din lume dar nu sunt din lume, cologațiile ca “we-them”, “you-them”, “us-them” sunt frecvente în această carte. Multe din cologațiile frecvente conțin mai multe cuvinte ale alterității (uneori chiar aceleași) lucru care poate arăta o preocupare foarte mare pentru modul în care cei din afară privesc creștinismul (lucru de altfel normal pentru o carte de apologetică). Notă: în analiza tuturor documentelor, cuvântului “you” (frecvent în această carte) i-a fost atribuit un scor de 0 (în vectorul de scoruri) pentru toate componentele (considerându-l prea frecvent și mai puțin semnificativ); dacă ar fi avut un scor mai mare ca 0 pentru alteritate, ar fi condus la o valoare și mai mare a alterității pentru acest text cu atât mai mult cu cât este prezent în cologații cu alți termeni ai alterității.

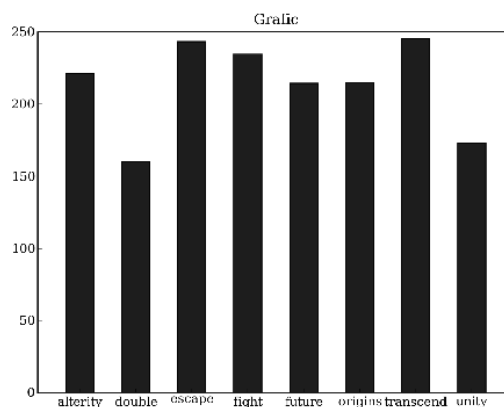


Figura 14. Mere Christianity, varianta a 4-a

Eliminând complet amplificarea dată de prezența în cologație a unor cuvinte similare (variantea 4), putem observa o netezire a valorilor, iar alteritatea nu mai este prima (cuvintele caracteristice alterității apărând adesea împreună, în cologații precum cele despre care am vorbit mai sus – “we-them”, “we-they”, “you-them”, “us-them”). Dacă “you” – vezi explicația pentru varianta 1 – (frecvent în acest text) ar fi avut un scor mai mare ca

zero pentru alteritate efectul asupra graficului ar fi fost probabil și mai mare.

Dacă analizăm secțiunea de religie a corpusului Brown conform variantei 1 obținem distribuția din figura 15.

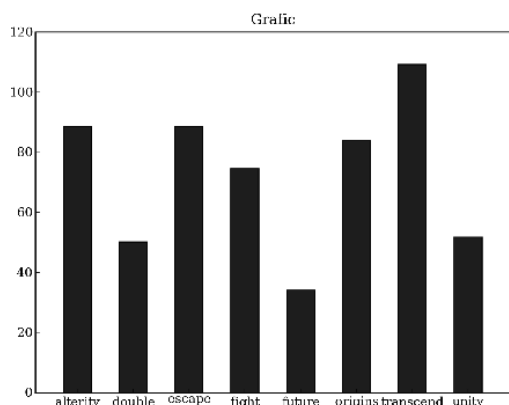


Figura 15. Corpusul Brown, secțiunea religie

Se observă o valoare mare pentru transcendență și, la fel ca și în cazul celuilalt text religios valori mari pentru alteritate și evadare. Deci primele trei componente ale imaginarului sunt aceleași ca și în textul lui C.S. Lewis.

Eliminând prima amplificare (varianta 2) asupra scorului (dată de deviație și distanța față de medie) și păstrând doar componentele mari ale scorurilor neamplificate, transcendența rămâne net deasupra celorlalte componente. De asemenea și dublul (componentă a domeniului religios) rămâne deasupra altor componente. Acest lucru poate arăta faptul că multe cuvinte au constant o apropiere mai mare de aceste componente.

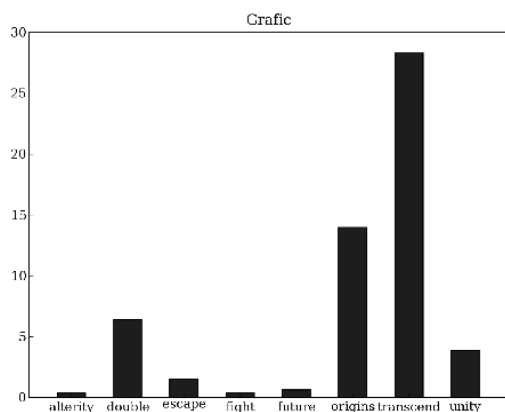


Figura 16. Corpusul Brown, secțiunea religie, varianta 2

Ca și pentru textul lui C.S. Lewis, dacă eliminăm amplificarea dată de prezența în colocații a unor cuvinte similare (varianta a 4-a), alteritatea cade din pozițiile fruntașe. În cazul documentului de față, de pe poziția a doua alteritatea ajunge pe poziția a șaptea din opt posibile; deci în ambele texte, cuvintele caracteristice alterității apar frecvent împreună. Alteritatea din "Mein Kampf" și "Turner Diaries" (cărți rasiste) nu pierde poziția fruntașă la folosirea variantei a 4-a.

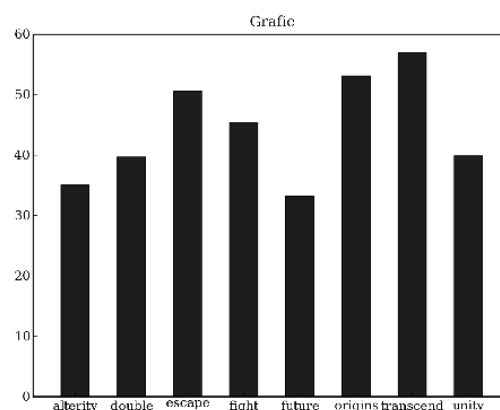


Figura 17. Corpusul Brown, secțiunea religie, varianta 4

## CONCLUZII

Programul a arătat că putem determina prezența – și măsura prezenței – într-un text a unor concepte dintr-o ontologie proprie. De exemplu, structura ce ține de ideea de *transcendență* o întâlnim în toate lucrările ce stau la baza studiului nostru într-un procent destul de semnificativ. Acest lucru demonstrează că există o preocupare constantă a omului pentru căutarea unor soluții transcendente. Progresul tehnologic (ce poate fi văzut ca o construcție secularizată a perfecționării omului) nu poate modifica în esență acest dat fundamental. Credința omului în existența unei realități de esență superioară marchează profund legăturile noastre cu Universul, necunoscutul, timpul și spațiul. Tendința de desacralizarea (propusă de filosofii iluminați) a pus în locul lăsat liber de credințele religioase propriu-zise o inepuizabilă colecție de credințe și milenarisme laice (comunismul, teoriile revoluționare, teoria progresului etc.) ce se înțeleg foarte bine cu știința și tehnologia. În esență, „desacralizarea” n-a schimbat nimic esențial, societatea tehnologică a pus în locul eroilor cu puteri supranaturale robotul.

Analiza pe lucrările menționate scoate în evidență o altă structură a imaginarului *alteritatea*. Frecvența mare a cuvântului în textele studiate sugerează existența unui raport dintre identitate și alteritate. Ne definim mereu în raport cu ceilalți. Într-un sens mai larg, alteritatea se referă la un întreg ansamblu de diferențe: societăți diferite, ființe diferite, spații și peisaje diferite. Acest mecanism conduce la o lume foarte fragmentată, fascinantă dar în același timp neliniștitoare.

Programul poate fi modificat ușor pentru a folosi o altă ontologie cu oricâte componente. O serie de amplificări asupra unor valori numerice folosite în calcularea scorurilor (apropieri față de componente ale ontologiei) pot releva alte informații interesante despre un text (vezi secțiunea "Rezultate").

Ontologia imaginarului construită pentru analiza noastră este potrivită probabil cel mai mult identificării

psihologiei unui text (sau a psihologiei autorului acestuia) și poate să nu fie potrivită pentru analiza unor texte pur științifice.

Folosirea unei alte ontologii mai specifice poate să dea rezultate mai bune dacă dorim clasificarea unor documente în funcție de apartenența la un anumit domeniu sau o anumită categorie a unui domeniu.

#### REFERINȚE

1. Lucian Boia, *Pentru o istorie a imaginarului*, Traducere de Tatiana Mochi, București, Editura Humanitas, 2000.
2. Gilbert Durand, *Structurile antropologice ale imaginarului. Introducere în arhetipologia generală*, Traducere de Marcel Aderca, București, Editura Univers Enciclopedic, 1998.
3. Jacques Le Goff, *L'imaginaire médiéval*; Éditions Gallimard, Paris, 1985
4. NLTK, [http://nltk.org/index.php/Main\\_Page](http://nltk.org/index.php/Main_Page)
5. WordNet, <http://wordnet.princeton.edu>
6. Winograd, T., Flores, F., *Understanding computers and cognition*, Ablex Publishing Corp., 1986