

ANALIZA DOCUMENTARĂ DE CONȚINUT: ABORDĂRI TEORETICE

Natalia GOIAN,
dr. conferențiar universitar, USM

Organizarea activității documentare la nivel internațional e legată de numele marelui savant Paul Otlet. În 1934 Paul Otlet publică lucrarea intitulată *Tratat de Documentare*, unde în unul din capitole vorbește și despre „operații, funcții, activități, care ne-ar permite stabilirea punctelor de acces al documentului”. Tot în această lucrare, autorul specifică una din etapele documentării - cea de descriere, fără a numi termenul „analiză documentară”. Astăzi, mai mulți autori, referindu-se la procesele de selectare a datelor principale ce caracterizează documentul, numesc aceste operații **analiză documentară**.

Principalele dicționare în știința informării întăresc acest punct de vedere.

Pia Atanasiu și Natalia Greco, punând în relief principalele aspecte în tratarea „analizei documentare” o prezintă ca „examinarea unui document și extragerea datelor bibliografice și a caracteristicilor de conținut ale acestuia”. Mai sunt și alte opinii, conform cărora analiza documentară se consideră ca o etapă în funcționarea sistemului de informare, care începe cu colectarea documentelor primare și se termină cu înmagazinarea informațiilor.

Analizând diferitele puncte de vedere asupra noțiunii de **analiză documentară** putem concluziona că nici unul dintre autori nu echivalează analiza do-

cumentară numai cu caracteristica formală a documentului, dar și cu caracteristica conținutului acestuia.

Până nu de mult, centrele documentare pledau mai mult pentru acumularea documentelor, precum și pentru căutarea soluțiilor de gestionare a proceselor de acumulare, de multe ori în detrimentul proceselor de selectare a informației.

Și atunci e necesar să situăm analiza documentară de conținut în sensul tradițional acceptat de documentariști și lingviști, puțin diferit de concepțiile teoreticienilor în domeniul comunicării, care tratează analiza documentară ca „analiza frecvențelor în conținutul manifest al unui mesaj, folosind identificarea și numărarea unităților-cheie de conținut ca bază a metodei ei” /3/. Putem considera că analiza documentară de conținut se polarizează între două universuri de referință, cel al autorului și cel al documentaristului și se încheie pe o unitate texto-contextuală determinată. Ținând cont de faptul că punctul de plecare (autor-text-context) este inamovibil, putem pătrunde doar în alegerea documentaristului, al cărui nivel de cunoștințe va condiționa sensibil rezultatele.

Deoarece obiectivele documentare constituie un fapt real, un punct de plecare în procesul analizei documentare sunt cinci variabile ce definesc cadrul conceptului de referință al analizei do-

cumentare de conținut: *textul, contextul, orizontul, viziunea documentaristului, obiectivele documentare și procedura de validare a rezultatelor.*

Dacă admitem pluralismul interpretărilor la care se prestează un document textual (în funcție de cine îl solicită, de ce și pentru ce), este necesar să stabilim diferența între doi termeni polisemantici: conținut și semnificație. Pentru a simplifica panorama complexă în jurul conceptului de conținut, cu un astfel de cuvânt ne referim la materia intrinsecă a documentului, care este într-o măsură oarecare independentă de uzul său temporal. În Dicționarul Explicativ Uzual conținutul este tratat ca „latura conceptuală a unei lucrări artistice sau conținutul unei cărți”. Toate documentele au un conținut relativ permanent, dar un număr variabil de semnificații, în funcție de utilizarea concretă a acestui conținut substanțial într-un moment determinat. Evident, că există o relație directă între conținutul substanțial al unui document și semnificațiile sale potențiale pentru utilizatori. După Pia Atanasiu și Natalia Grecu, conceptul *semnificație* este „înțeles (conținut semantic) al unui cuvânt” /1/. În Dicționarul explicativ al limbii române, această noțiune, *semnificația*, este tratată drept conținutul semantic al unui cuvânt; înțeles, sens, accepție... /4/.

Vom distinge în conținut cel puțin trei dimensiuni:

- Informativă, care ne trimite la sensul (semnificatul, forma fizică) semantic pragmatic al documentelor, la caracterul său descriptiv conform căruia semnificația unei expresii este aceea la care se referă sau o reprezintă.

- Enunțiativă, care prezintă o situație discursivă, fiind determinante contextele emițător (locutorul care emite mesajele) și, mai cu seamă, receptor (sau destinata-

rul, cel care primește mesajul transmis).

- Documentară, care se delimitează adaptând conținutul enunțiativ la o situație documentară determinată, implicând atribuirea de proprietăți unui document în scopul descrierii sale.

În urma acestor explicații, concepem analiza documentară de conținut ca pe un proces cognitiv de recunoaștere, descriere și reprezentare a conținutului documentar.

Analiza documentară de conținut se fundamentează pe un ansamblu de paradigme-bază înglobate în diferite domenii științifice: psihologie cognitivă, sociologie, logică, lingvistică textuală și semiotică discursivă, teoria comunicării, documentare. Paradigma psihologiei cognitive este un domeniu relativ tânăr, care concepe omul ca pe o ființă capabilă să producă informație în termeni ce depășesc fizicul pur pentru a atinge niveluri superioare. Una din variantele sale, procesarea informației, încearcă să explice procesele mintale și structurile implicate în dobândirea, înțelegerea, procesarea și folosirea informației.

Logica contribuie, de asemenea, într-un mod esențial la investigarea și dezvoltarea proceselor de analiză a conținutului, mai ales la nivel structural sau sintactico-semantic, folosind două tipuri de raționamente: deductiv și inductiv. Una din variantele cele mai eficiente este logica formală, care s-a transformat într-un auxiliar inestimabil al analizei conținutului, deoarece ne permite să operăm în mod logic cu diferite unități semantice de comunicare și argumente cu unicul scop de a le atribui o categorie sintactică specifică. Transformările care derivă din aceste operații logice asupra unităților logico-pragmatice garantează integritatea textului ca unitate de conținut.

O limbă este un sistem de semne sau

uniuni inseparabile din două componente, constituite în mod reciproc ca structuri și dependente una de alta, fiind ambele de natură convențională. Lingvistul elvețian F. de Saussure le numea semnificat și semnificant. În lingvistica saussuriană, limba este un sistem abstract de semne și convenții care stă la baza actelor individuale de enunțare /3/. Astfel, lingvistica se constituie ca un domeniu special în cadrul unei teorii generale a semnelor, în care se evidențiază lingvistica textuală, cele mai interesante atribuții ale căreia se centrează în două fronturi-cheie pentru cercetarea documentelor: structuri și tipologii textuale. În realitate, analiza documentară de conținut utilizează lingvistica ca univers teoretic capabil să subzideze activitatea sa practică.

Fără îndoială că diferitele teorii ale comunicării contribuie la configurarea unei doctrine a analizei documentare de conținut. Înțelegem prin comunicare întregul proces de transmitere a informației care are loc prin intermediul emisie, transportării și recepționării unui mesaj.

Termenul *comunicare*, fiind utilizat din secolul 14, pe parcurs își schimbă sensul. „În contextul noilor mijloace de comunicare - radio, cinema, televiziunea - termenul adecvat ar fi cel de „comunicare - difuzare” /5/. Făcând o trecere în revistă a opiniilor diferitor autori, în continuare Ion Drăgăan scrie că „Apariția comunicării tehnice, pe lângă cea naturală, implică deci o schimbare a sensului termenului: ideea de „împărțire”, „împărțășire” (a unor mesaje) va fi dublată de cea de „transmitere” (comunicare instrumentală)” /5/.

Există două mari tipuri de definiții ale comunicării. Prima vede comunicarea ca fiind un proces prin care A trimite un mesaj lui B, mesaj ce are un efect asupra celui din urmă. A doua definiție

vede comunicarea drept negocierea și schimbul de semnificație, proces în care mesajele, persoanele determinate cultural și „realitatea” interacționează astfel încât să ajute înțelesul ca să fie produs, iar înțelegerea - să apară /3/. Dacă prima abordare are ca scop identificarea etapelor prin care trece comunicarea fixată de Lasswell (1948) prin modelul său “Cine spune ce / pe ce canal / cui / și cu ce efect?” /3/, a doua abordare este cea structuralistă, care accentuează relația dintre elementele constitutive necesare pentru ca înțelesul să apară. Aceste elemente se pot încadra în trei mari categorii: textul, semnele și codurile lui:

1. persoanele care „citesc” textul, experiența socială și culturală care i-a format atât pe ei, cât și semnele /codurile pe care le folosesc;

2. conștiința unei „realități externe” la care se referă atât textul, cât și oamenii /3/.

Această, ultimă, abordare se adaptează în întregime la interesele noastre, în analiza documentară de conținut, deoarece se sprijină pe semiotică, pe teoria enunțării, pe pragmatică și pe teoria textului.

Unele autorități în domeniu, cum ar fi F. de Saussure, pun în prim - plan gruparea elementelor care alcătuiesc un „text” (semne / coduri / limbaj) /3/. Dar bazele interdisciplinare ale analizei documentare includ, de asemenea, aspecte matematice (modelul Shannon și Weaver), statistice (legile Zipf, Bradford etc.) și, în cele din urmă, toate acestea relaționate cu științele umanistice și sociale, în sânul cărora se întâlnesc sociologia și toate ramurile psihologiei, fără a uita de alte domenii adiacente precum sunt informatica, cibernetica, inteligența artificială și sistemele experte.

Procesul unitar al analizei documentare de conținut poate fi fragmentat în

anumite secvențe, și anume: *lectură, analiză, sinteză*. Ținând cont de problema în cauză, putem afirma că ***lectura este unica formă posibilă de acces la conținutul documentului***.

Lectura este tratată de Lazăr Șăineanu ca „faptul de a citi, citire, citit; text citit” /8/. Lectura este un proces de natură interactivă care depinde atât de text, cât și de persoana care citește. Lectura constă într-un șir de procedee coordonate care includ operațiuni perceptuale lingvistice și conceptuale. Înțelegerea unui pasaj scris implică o inserare continuă a unei doze mari de cunoștințe prealabile. Un rol important în procesul lecturii îi revine viziunii (persoana care citește construiește în minte tablouri de înțelegere); acțiunii duble ale memoriei (pentru a putea mobiliza tablourile, punând în relație cunoscutul și necunoscutul); intervenția rațiunii (de analiză și sinteză). Deci, putem afirma că lectura implică nu numai semne vizuale care compun textul ci și aspecte cognitive în strânsă relație cu cunoștințele depozitate în memoria cititorului.

Odată înțeles, textul trebuie analizat, analiză ca metodă de investigare constând în descompunerea subiectului cercetat în elementele lui componente, fiecare din acestea din urmă putând constitui ulterior un obiect de studiu separat /1/. Analizând textul trebuie să luăm în considerare că sunt procese care se interpenetrează: în timpul lecturii documentaristul decodifică, interpretează structurile, reprezintă informația. Fără memorie n-am putea nici cel puțin să integrăm informația în sistemul nostru de cunoștințe, chiar pentru un timp foarte scurt.

În procesul analizei documentare e important să ținem cont de două operații: *selectarea și interpretarea*.

Selectarea este procesul fizic consecutiv comparării unei cereri de informare cu un masiv de informații, respectiv, separarea din masiv a acelor informații care *corespund caracteristicii de regăsire a cererii/1/*. Referindu-ne la cazul concret, anume, la analiza documentară de conținut, selectarea constă în eliminarea unităților de înțeles – fraze și cuvinte care sunt considerate irelevante pentru analiză.

Din punctul de vedere al conținutului documentelor, selectarea trebuie să aibă în vedere actualitatea informațiilor, cât și semnificația lor. La fel de importante sunt personalitatea analistului, calificarea lui pentru a trata subiectul dat. Textul trebuie supus în primul rând, procesului de comprimare (condensare). Condensarea textului (comprimarea textului) constituie un proces de reducere a textului printr-o nouă redactare sau prin prelucrarea documentară, astfel urmărindu-se păstrarea celor mai importante informații. Gradul de reducere poate fi diferit în funcție de scopul pe care-l urmărim.

Interpretarea textului depinde, în mare măsură, de orizontul documentaristului, de context și de alte obiective documentare. Din aceste motive interpretarea textului poate fi subiectivă. După ce textul a fost interpretat, este necesară elaborarea unor materiale de informare care să redea concentrat toate informațiile cuprinse în document. Acest rol îi revine procesului de sinteză documentară. Sintezele documentare prezintă sub o formă generalizată informațiile cuprinse în documentele primare apărute pe un anumit interval de timp, pentru o tematică prestabilită. Ele au ca scop să orienteze pe specialiști în fluxul de informații, să-i dea o imagine clară asupra stadiului atins și a perspectivelor de dezvoltare în temele tratate,

eliminând detaliile secundare. Sinteza este pasul cel mai delicat și dificil căci, deși activitățile analitice prealabile se pot supune unor tehnici mai mult sau mai puțin variate, este practic imposibil de a stabili tehnici de sinteză care să fie valide pentru toate tipurile de documente și de documentariști.

Analiza documentară de conținut poate fi efectuată prin intermediul diferitor operații analitico-sintetice și anume: *clasificarea*, *indexarea*, *procesul de rezumare*. Aceste operații sunt tratate în diferite dicționare de specialitate, de către diferiți autori. Unii autori /6/ caracterizează *clasificarea* ca "transformarea sau reprezentarea datelor prin aplicarea unui cod". Alții /1/ tratează *clasificarea* ca pe un „proces de încadrare a unui subiect într-o subdiviziune a unui sistem de clasificare, pe baza existenței sau lipsei unor anumite caracteristici sau gruparea noțiunilor în clase și subclase cu caracteristici identice sau similare evidențiind relațiile semantice dintre ele.

Indexarea este „reprezentarea cu ajutorul limbajului documentar sau natural a datelor care rezultă din analiza conținutului unui document sau a unei părți dintr-un document în vederea regăsirii acestuia” /6/.

Rezumat (documentar) este „redarea pe scurt a principalelor idei și concluzii cuprinse într-un document” /1/.

În continuare, ne vom referi la **clasificările de bibliotecă**.

Clasificarea este un ansamblu ordonat de concepte distribuite în mod sistematic în clase, formând o structură bine determinată. De regulă, clasificările sunt compuse din: *tabele principale*; *tabele auxiliare*; *index alfabetic pe materii*.

Tabelul principal conține clasele care ne prezintă universul cunoașterii umane, diviziunile și subdiviziunile în care se

descompune fiecare clasă, divizate ierarhic de la general la particular.

Tabelul auxiliar îmbină o serie de concepte suplimentare, necesare pentru a evidenția (oglinzi) forma de prezentare a documentului, punctul geografic, limitele cronologice, limba în care este scris documentul etc., care urmează să fie alăturate conform anumitor reguli de alte concepte din tabelul principal.

Indexul alfabetic pe materii este o listă de termeni aranjați în ordine alfabetică. Fiecare termen este urmat de o notație corespunzătoare care ne trimite la clasa, diviziunea, subdiviziunea respectivă din tabelul principal.

Cuvântul *clasificare* este polisemantic și poate fi interpretat în sens dublu: a) ca proces de ordonare și repartizare în clase a obiectelor, ținând seama de anumite caracteristici; b) ca un sistem propriu de clase, dinainte prestabilit, în care aceste obiecte se ordonează. Înțelegem *clasificarea* ca pe o operație intelectuală proprie cunoașterii umane, care încearcă să dezvăluie conținutul documentelor pentru a-l formaliza și reprezenta cu ajutorul unui limbaj prestabilit. Obiectivul principal al clasificării constă în a permite gruparea materiilor în clase cu scopul de a depozita și recupera ulterior informația. *Clasificarea* afectează, de asemenea, conținutul necesităților informaționale ale utilizatorilor, pentru că în același mod determină în ce clase și subclase ale uneia sau alteia clasificări s-a situat informația care poate satisface aceste necesități. Procesul de clasificare este supus anumitor reguli formale de logică, deoarece materiile trebuie divizate de fiecare dată doar în corespundere cu o caracteristică de divizare, iar clasele și subclasele rezultante trebuie să se excludă reciproc și simultan. Divizionarea în clase trebuie să fie continuă, fără salturi.

Spre deosebire de indexare, care este o operație de natură analitică, clasificarea cere un efort notabil de sinteză orientat spre detectarea și izolarea temei principale și nu a cuvintelor, prin care se prezintă documentul. Pentru a stabili o tipizare a limbajelor de clasificare trebuie să ținem cont de componența lor lexicală și de structura lor.

Așadar, *clasificările pot fi deosebite după conținut și după structură.*

După conținut, deosebim *clasificări enciclopedice și clasificări specializate.*

Clasificările enciclopedice se prezintă sub formă de liste de termeni normalizați din toate domeniile cunoașterii. De tip enciclopedic, putem numi, sistemele mari de clasificare: CDD, CZU.

Clasificările specializate se prezintă sub formă de liste de termeni normalizați într-un anumit domeniu. Clasificările specializate se concentrează pe o singură temă, urmată de alte subteme dependente de tema principală. Printre cele mai cunoscute clasificări speciale putem numi Clasificarea zecimală în domeniul astronauticii (CDA), Clasificarea în domeniul medicinei a Bibliotecii Naționale de Medicină (National Library of Medicine) etc.

Din punctul de vedere a structurii, deosebim *clasificări enumerative, clasificări pe fațete și clasificări mixte.* Cele enumerative sunt sistemele tradiționale de clasificare bibliotecar-bibliografice, ele enumeră sau introduc în liste toți termenii necesari pentru tratarea unei teme aparte (clasificările specializate) ori încearcă să cuprindă tot universul cunoașterii - în cazul clasificărilor enciclopedice. După structură putem numi: CZU, CDD, Clasificarea Bibliotecii Congresului (LCC) etc.

Clasificările pot fi, de asemenea, *științifice* – sistematizând fenomenele lumii naturale, ca un ajutor la studiul și considerarea lor logică – sau *documen-*

tare, ajutând la organizarea documentelor în scopul facilitării găsirii informației. CZD și CZU sunt clasificări documentare. Clasificările de bibliotecă tind a fi ori *de bibliotecă* – pentru aranjarea fizică a colecțiilor ca atare și ca un ghid pentru cuprinderea acestora – ori *bibliografice*, pentru indexarea și descrierea detaliată a documentelor, fără referință la o anumită colecție. CZD s-a format ca o clasificare de bibliotecă, iar CZU ca una bibliografică/2/.

Metodica tradițională de prezentare a acestor clasificări constă în divizarea domeniilor cunoașterii umane în clase, clasele în subclase etc., se divid de atâtea ori, de cât este nevoie de detaliere în procesul de clasificare. Unele clasificări ierarhice au și particularitatea de a prezenta universul informației în grupe, care se împart în zece clase, iar acestea, divizează la rândul lor, în alte zece subclase, fiecare subclasă în alte zece etc. până la atingerea unei mari precizii, reducând extensiunea unui concept la limite foarte precise.

Un element caracteristic pentru toate sistemele de clasificare este *notația*. Notația reprezintă un sistem de simboluri, numere sau combinații de simboluri și numere care sunt atribuite termenilor de clasificare și sunt capabile să reprezinte clasele, diviziunile, subdiviziunile. Scopul principal al notației este de a determina succint conceptele cât și relațiile dintre ele. Când se utilizează un singur tip de simboluri, este o notație pură. Ca exemplu poate servi CDD, în care indicii de clasificare sunt exprimați prin cifre arabe. Un sistem care utilizează mai mult de un tip de simboluri se numește *sistem de notație mixtă*. Clasificarea Colon are o notație mixtă de tip alfa-numerică, în exprimarea căreia sunt utilizate literele alfabetului latin, literele alfabetului grecesc și numere (cifre).

Un interes deosebit, astăzi, îl prezintă discuțiile despre o altă clasificare care nu prea este abordată în Republica Moldova. Este vorba de o „clasificare în interesul cititorului” [7]. „Clasificarea pe interesul cititorului” este o clasificare simplă care

își propune să reflecte mai degrabă interesele speciale ale utilizatorului, decât conținutul cărții. Plasând utilizatorul în centrul preocupărilor, clasificarea în interesul cititorului încurajează și promovează accesul la colecții.

Referințe bibliografice:

1. Atanasiu, P. *Termeni din domeniul informării și documentării: Îndrumar metodologic* / Pia Atanasiu, Natalia Grecu. – București, 1988. – 482 p.
2. *Clasificarea Zecimală Universală: Ediție medie internațională în limba română: În 2 părți. P.1 Tabele sistematice* / Ediție îngrijită de G. Clinca. – București, 1998. – 368 p.
3. *Concepte fundamentale din științele comunicării și studiile culturale* / Tim O’Sullivan, John Hartley, Danny Saunders / București: Polirom, 2001
4. *Dicționarul explicativ al limbii române*. – Ediția a 2-a. – București: Univers Enciclopedic, 1996. – 1192p.
5. Drăgan, Ion. *Paradigme ale comunicării de masă: Orizontul societății mediatice: Partea 1* / Drăgan Ion. – București, 1996. – 303 p.
6. Regneală, Mircea. *Dicționar explicativ de biblioteconomie și știința informării* / Mircea Regneală. – București: FABR, 2001. – 425 p.
7. Sapiie, J. *O clasificare în interesul cititorului: scheme ușor accesibile utilizatorului* // *Biblioteconomie: culegere de traduceri prelucrate*. – 1997. – V. 33. – Nr. 4. – P. 77-83.
8. Șăineanu, Lazăr. *Dicționar universal al limbii române: 77.000 de articole* / Lazăr Șăineanu; Coordonator: A. Dobrescu. – Chișinău: Litera, 1998. – 1362 p.

O BAZĂ DE DATE ÎN CONTINUĂ EVOLUȚIE

Ludmila PÎNZARU,
director adjunct BM „B. P. Hasdeu”

WorldCat creat de bibliotecile membre ale OCLC¹ există mai mult de 30 de ani și este una din cele mai funcționabile baze de date bibliografice. Astăzi el se transformă în full-fledged „service platform”, avînd o structură bazată pe web, care este utilizată de biblioteci în managementul colecțiilor prin separarea funcțiilor.

Ce face WorldCat-ul în continuare?

Furnizează un singur punct de intrări pentru bibliotecile care caută o expoziție pentru întreaga colecție. Atrage parteneri de business care vor să unească mii de biblioteci, ca parte a mecanismelor lor de căutare sau alte resurse. Extinde promisiunea bibliotecii de a se include 24/7 în customer-driven, serviciu perso-

¹ OCLC – Online Computer Library Center