

Richard P. SMIRAGLIA. *Domain analysis for Knowledge organization. Tools for ontology extraction.* Waltham, MA (USA); Kidlington (UK), Chandos Publishing (Elsevier), 2015, p. 116, ISBN 978-0-08-100150-9, € 56,95.

Richard P. Smiraglia, professore presso la School of Information Studies dell'Università di Wisconsin-Milwaukee (USA) e direttore, da diversi anni, della rivista «Knowledge Organization», organo ufficiale della International Society for Knowledge Organization (ISKO), in questo volume di 105 pagine presenta la *Domain Analysis* come paradigma metodologico per la Scienza dell'organizzazione della conoscenza e come insieme di tecniche e di strumenti per l'individuazione delle basi di conoscenza specifiche per i diversi settori scientifici.

La denominazione *Domain Analysis* in realtà si deve a Ruben Prieto-Diaz, *computer scientist*, che la utilizzò nel 1990 nell'ambito di un progetto per la realizzazione di una classificazione dei componenti software, allo scopo di proporre lo sviluppo e il riuso, che venne modellata sullo schema analitico-sintetico di Ranganathan. In quel contesto, *Domain Analysis* prevedeva l'esplorazione da diverse prospettive, e tenendo conto dei diversi aspetti, del dominio di conoscenze dell'Ingegneria dei software. Dopo questo primo utilizzo, *Domain Analysis* ha assunto la fisionomia di un modello generale per l'analisi dei diversi ambienti e contesti socio-culturali di produzione della conoscenza, ed è stata proposta come paradigma interpretativo per la Library and Information Science alla metà degli anni Novanta del secolo scorso,

anche in opposizione all'orientamento cognitivista che, sulla scia degli studi di Psicologia cognitiva e di Intelligenza artificiale, aveva trovato spazio all'inizio dello stesso decennio anche nell'ambito della LIS. L'orientamento mentalista e cognitivista, fondato sullo studio della conoscenza come processo mentale solipsistico, isolato dal contesto socio-culturale, aveva indicato una linea di ricerca volta ad analizzare le modalità della comprensione, della memorizzazione a breve e a lungo termine, dei processi del pensiero durante la ricerca e del recupero dei concetti da parte dei singoli esseri umani, per elaborare sistemi di informazione basati sui meccanismi cognitivi degli utenti.

In un fondamentale articolo del 1995 Birger Hjørland e Hanne Albrechtsen (*Toward a New Horizon in Information Science: Domain Analysis*, «JASIS», 46, n. 6, p. 400-425) hanno presentato *Domain Analysis* come il nuovo paradigma interpretativo e il possibile fondamento teoretico della Scienza dell'informazione, il cui compito avrebbe dovuto quindi essere proprio quello di elaborare i contenuti che si presentano all'interno dei differenti campi di conoscenza, cioè nell'ambito delle diverse comunità di produzione scientifica. I domini vengono analizzati come comunità di discorso e di pensiero, come strutture cognitive collettive, che in parte determinano anche la percezione della rilevanza o meno dei documenti utilizzati dai ricercatori. Negli anni successivi ancora Hjørland delinea le prospettive offerte dalla *Domain Analysis* (*Domain Analysis in Information Science. Eleven Approaches Traditional as Well as Innovative*, «Journal of Documentation», 58, 2002, n. 4, p. 422-462) e cura, insieme a Jenna Hartel, un numero speciale di «Knowledge Organization» (30, 2003, n. 3/4 p. 125-245) interamente dedicato alla riflessione sui diversi aspetti della prospettiva *domain-analytic*, che emerge qui come un modello di approccio alla Information Science che privilegia le dimensioni socio-culturali e storiche del concetto di informazione. Nel 2015 ancora un numero speciale di «Knowledge Organization» (42, n. 8), a cura dello stesso Smiraglia e di Maria José Lopez-Huertas, fa il punto sugli sviluppi recenti degli studi *domain-analytic*. Smiraglia illustra breve-

mente la *Domain Analysis* per la *Knowledge Organization*, e negli altri contributi vengono discusse le problematiche relative allo studio dei domini interdisciplinari, viene analizzato l'*Information retrieval* come dominio scientifico, e vengono delineate le prospettive dell'applicazione della D. A. anche al campo archivistico.

Smiraglia, autore di un altro recente volume – *The elements of Knowledge organization* (Heidelberg, Springer, 2014) – nel quale ha presentato gli aspetti principali dell'organizzazione della conoscenza e i suoi strumenti, e nel quale ha già esposto sinteticamente anche il paradigma della *Domain Analysis*, in questo volume edito da Chandos nel 2015 ne ripercorre brevemente la genesi, a partire dalla metà degli anni Novanta del secolo scorso, e gli sviluppi, e ne cita la letteratura pertinente. Obiettivo specifico del volume è tuttavia l'analisi delle tecniche empiriche volte ad individuare la base di conoscenza di un dominio scientifico o di una comunità di discorso, e cioè i concetti condivisi da un gruppo disciplinare e le relazioni tra i concetti stessi. L'assunto di base è il fatto che in un dominio scientifico gli scienziati manifestano un fondamentale consenso epistemologico e condividono gli approcci metodologici e la terminologia.

S. si concentra sull'applicazione della *Domain Analysis* al settore della *Knowledge Organization* e sulla necessità di adottarla come paradigma metodologico per sviluppare gli strumenti per l'organizzazione della conoscenza – thesauri, sistemi di classificazione, ontologie, bibliografie ragionate – specifici per ciascun dominio. Nell'ambito della K. O., infatti, il concetto di dominio si applica in particolare ai gruppi che hanno sviluppato un sistema specifico per l'organizzazione della conoscenza. L'obiettivo della D. A. è mappare e visualizzare la conoscenza condivisa da una comunità di ricerca, i concetti e le relazioni tra i concetti, e utilizzare i risultati della mappatura per elaborare specifici *knowledge organization systems*. L'analisi dei domini scientifici fonde in un unico tronco la dimensione ontologica, cioè quali sono i concetti su cui un settore scientifico lavora, e quella epistemologica, cioè in che modo essi vengono recepiti e utilizzati nel settore specifico.

Nei capitoli centrali del volume S. raccoglie la presentazione dei metodi empirici e delle tecniche per mappare e visualizzare le basi di conoscenza dei singoli domini. Si possono individuare i concetti che costituiscono gli elementi fondamentali delle basi di conoscenza adottando metodi empirici, ad esempio individuando gli attori della comunità scientifica e i loro ruoli, gli obiettivi, le metodologie e la produzione del dominio scientifico, e si possono empiricamente estrarre i termini scientifici utilizzati nei testi messi in circolazione. In particolare nel capitolo quarto, il più lungo e ricco di informazioni e di dimostrazioni pratiche, vengono presentate le tecniche per l'individuazione e la visualizzazione dei concetti, cioè per realizzare una mappa ontologica di ciascun dominio: *citation analysis*, cocitazione degli autori e *network analysis*, *co-word analysis*.

La *citation analysis*, tecnica bibliometrica nota a tutti gli esperti di Scienza dell'informazione, viene esibita come metodo per rappresentare la semantica sociale di una comunità di discorso. Si presenta nei dettagli il funzionamento di due notissimi servizi commerciali, Web of Science (WoS) di Thomson Reuters e Scopus di Elsevier, entrambi utilizzabili per creare mappe degli articoli più citati, sulla base dei concetti che vengono cercati, o elenchi degli autori più produttivi in un settore disciplinare, e infine per analizzare la tipologia dei lavori citati (articoli o monografie) e la loro *half-life*, cioè l'età mediana della letteratura in essi citata (età del 50% dei lavori citati), come indicatore di aggiornamento scientifico. S. non evidenzia, tuttavia, il fatto che, poiché i *temi* possono essere ricercati attraverso le parole presenti nei titoli, negli abstract e negli elenchi di parole chiave fornite nei record, questo non assicura l'individuazione dei concetti. Sottolinea molto opportunamente, invece, la mancanza della completa copertura delle aree dello scibile da parte dei due strumenti commerciali, e la necessità, che da ciò deriva, di ricorrere alla creazione manuale di indici citazionali per i molti domini che non sono coperti. Altrettanto giustamente viene rilevato da S. come la *co-word analysis*, l'insieme delle tecniche e degli strumenti per l'estrazione e la visualizzazione dei ter-

mini più frequentemente utilizzati nei lavori scientifici di un dominio, sebbene utile, non sia in grado di mettere in evidenza i concetti sui quali il dominio lavora, e che solo l'analisi dei termini all'interno dei contesti nei quali essi vengono impiegati potrebbe restituire la mappa semantica condivisa dal settore.

L'analisi degli autori, e dei lavori cocitati (*cocitation analysis*), tecnica anch'essa molto conosciuta e usata in LIS, da quando Henry Small (ISI), che tuttavia S. non cita, la introdusse nel 1973 (*Co-citation in the scientific literature* «JASIS», 24, p. 265-269) come nuova forma di *document coupling*, viene presentata nei dettagli. La cocitazione è la misura della frequenza con cui due (o più) autori o lavori vengono citati insieme da un altro. Questa misura permette di stabilire *cluster*, raggruppamenti di articoli (e autori) frequentemente citati insieme, e realizzare mappe di insiemi di articoli, o di autori, o di riviste, cocitati, utili a rappresentare un campo disciplinare, una scuola di pensiero o di ricerca. Infine, permette di analizzare all'interno di ciascun *cluster*, ma anche di reti di collegamento tra autori (*author cocitation networks*), quanto siano frequentemente citati gli autori, o gli articoli, e monitorare nel tempo i cambiamenti avvenuti nella citazione. Rientra tra le tecniche qui presentate la proposta personale di S. di arricchire lo studio dei domini scientifici con l'analisi delle opere che hanno avuto un ruolo centrale in un paradigma scientifico-culturale. S. la definisce *instantiation* delle opere, e la presenta come una tecnica volta a quantificare le monografie che costituiscono il nucleo scientifico di ciascun dominio e ad analizzarne le diverse edizioni e le eventuali opere da esse derivate.

Dopo aver dedicato i capitoli centrali alla presentazione delle tecniche empiriche per la individuazione dei domini scientifici, nel capitolo finale S. presenta brevemente un ulteriore approccio alla *Domain Analysis*, eminentemente qualitativo, basato sulla *Cognitive work analysis* (CWA), teoria che trova il fondamento nei lavori di Jens Rasmussen dedicati allo sviluppo di sistemi informativi, tra cui anche le biblioteche, nei quali egli coniugava modelli concettuali propri della *computer*

*science*, della psicologia e delle scienze cognitive (J. Rasmussen, A.M. Pejtersen, L.P. Goodstein, *Cognitive Systems Engineering*. New York, Wiley, 1994). CWA si offre come approccio multidisciplinare; è una teoria volta ad analizzare la dimensione cognitiva del lavoro collettivo di una comunità scientifica, ed è basata sullo studio delle interazioni che avvengono tra i membri di un dominio e sull'analisi dei limiti e dei vincoli imposti alla ricerca dai singoli domini (stato della tecnologia, criteri di distribuzione dei fondi, leggi sul copyright, ecc.). S. mette in risalto l'apporto della CWA al settore della *Knowledge organization*, e in particolare sottolinea l'importanza delle interpretazioni simboliche e l'analisi dei contesti per interpretare correttamente i concetti e i significati coinvolti in un dominio scientifico, realizzare la base di conoscenza del dominio stesso e utilizzarla quindi per elaborare specifici sistemi per l'organizzazione della conoscenza espressa da quel settore. Tuttavia, CWA costituisce un paradigma teoretico troppo complesso per poter essere apprezzato dal lettore nelle poche pagine che qui sono dedicate alla sua analisi, ed avrebbe avuto bisogno di un esame più approfondito e di una presentazione dettagliata.

Nel complesso il volume, in linea con la politica di Chandos Publishing, volta a privilegiare la pubblicazione di monografie dal taglio teorico-pratico a beneficio soprattutto dei professionisti del settore della Library and Information Science, risulta utile soprattutto per la minuziosa presentazione delle tecniche e degli strumenti adatti a stabilire le basi di conoscenza dei diversi domini scientifici. Tuttavia, l'inquadramento teoretico della *Domain Analysis* e la sua applicazione al dominio della *Knowledge organization*, nonché i collegamenti di quest'ultima con la più generale Scienza dell'informazione, e infine le interessanti prospettive offerte dalla CWA, risultano piuttosto compressi e troppo sinteticamente presentati al lettore.

Maria Teresa Biagetti