

Wissensrepräsentation Die Erstellung von Stellvertretern (Repräsentanten oder ↗ Surrogaten), die die Aboutness (den Inhalt im Kontext) eines ↗ Dokuments abbilden. Die so entstehenden inhaltsbezogenen ↗ Metadaten eines Dokuments ermöglichen oder verbessern im ↗ Information Retrieval deren Auffindbarkeit. W. arbeitet ausschließlich mit explizitem objektivem ↗ Wissen. Ein Dokument enthält Wissen über Objekte (repräsentiert durch ↗ Begriffe, z.B. «Köln») oder Sachverhalte (repräsentiert durch Aussagesätze, z.B. «Köln hat im Jahr 2012 1 Mio. Einwohner.»). Der erste Aspekt führt zu den Informationsfiltern (Methoden der ↗ Wissensorganisation), der zweite zur ↗ Informationsverdichtung (↗ Referat). Die Methoden der Wissensorganisation lassen sich nach den beteiligten Akteuren in 3 Gruppen einteilen. 1. Intermediator-orientierte Informationsfilter: ↗ Indexierungssprachen (Knowledge Organization Systems, KOS) wie ↗ Nomenklatur (Schlagwortmethode), ↗ Klassifikation und ↗ Thesaurus benötigen Spezialisten sowohl für den Aufbau von Werkzeugen als auch für deren praktischen Einsatz, die ↗ Indexierung. ↗ Ontologien dienen dazu, das Wissen aus den Dokumenten in formalisierter Form direkt (ohne Umweg über Dokumente) in ein Informationssystem einzupflegen. 2. Autor-orientierte Informationsfilter: Ist das Dokument ein Text, so können spezifische Sucheingänge durch Markierung bestimmter Schlüsselwörter (im einfachsten Fall: der Titelsterme (Titelindexierung); aber auch im Rahmen der Textwortmethode weitaus elaborierter; ↗ Informationsextraktion) kreiert werden. Sind im Dokument Referenzen vorhanden, so werden diese im Rahmen der Zitationsindexierung ausgewertet. Bei wiss. Artikeln liegen häufig ↗ Autorenschlagwörter vor, die in Retrievalsysteme übernommen werden. 3. Nutzer-orientierte Informationsfilter: ↗ Folksonomies arbeiten mit freier Schlagwortvergabe, wobei die Nutzer ohne jede Regel Dokumente «taggen» (↗ gemeinschaftliches Indexieren). Nach dem Umfang der ↗ semantischen Relationen steigt die Aus-

drucksstärke der Methoden der Wissensorganisation von der Titelinde­xierung, Zitationsindexierung, Autorenver­schlagwortung und den Folksonomies (alle arbeiten nur mit der ↗ syntagmatischen Relation) über Nomenklatur (↗ Äquivalenzrelation), Klassifikation (zusätzlich ↗ hierarchische Relation), Thesaurus (zusätzlich ↗ Assoziationsrelation, ggf. Unterteilung der Hierarchie nach ↗ generischer Relation und ↗ partitiver Relation) zur Ontologie (frei definierbare Relationen) an. Es zeigt sich in der Praxis, dass der Umfang der repräsentierbaren ↗ Wissensdomäne umgekehrt proportional zur Ausdrucksstärke kleiner wird. Indexierung (oder Inhaltser­schließung) meint die praktische Tätigkeit des Abbildens der thematisierten Objekte in einem Dokument auf das Surrogat. Hierbei macht der Indexer (d. h. einer der gen. Akteure Intermediator, Autor oder Nutzer) von den Methoden der W. und – soweit vorhanden – den jeweiligen spezifischen Werkzeugen Gebrauch. Er legt damit diejenigen Begriffe fest, bei denen das Surrogat in einem Informationsfilter (wie ein Goldklumpen in einem Sieb) «hängenbleibt». Intermediator-orientiertes Indexieren wird entweder durch menschliche Indexer intellektuell (↗ intellektuelle Indexierung) oder automatisch durch Systeme durchgeführt (↗ automatische Indexierung). Inhaltsverdichtung geschieht durch Referieren. Die thematisierten Sachverhalte des Dokuments werden kurz, aber möglichst erschöpfend, in Form von Sätzen z. B. als ↗ Abstract abgebildet. Akteure sind entweder die Autoren, professionelle Abstractors oder maschinelle Systeme (beim Extracting oder beim Erstellen der Snippets bei den Trefferlisten von Suchmaschinen). Das Referieren verfolgt das primäre Ziel, den Nutzer bei den durch die Informationsfilter gefundenen Surrogaten bei der Ja/Nein-Entscheidung zu unterstützen, ob das Dokument tatsächlich seinen ↗ Informationsbedarf befriedigt.

↗ Informationsaufbereitung

Stock, W. G.; Stock, M.: Wissensrepräsentation. München 2008.

W. Stock