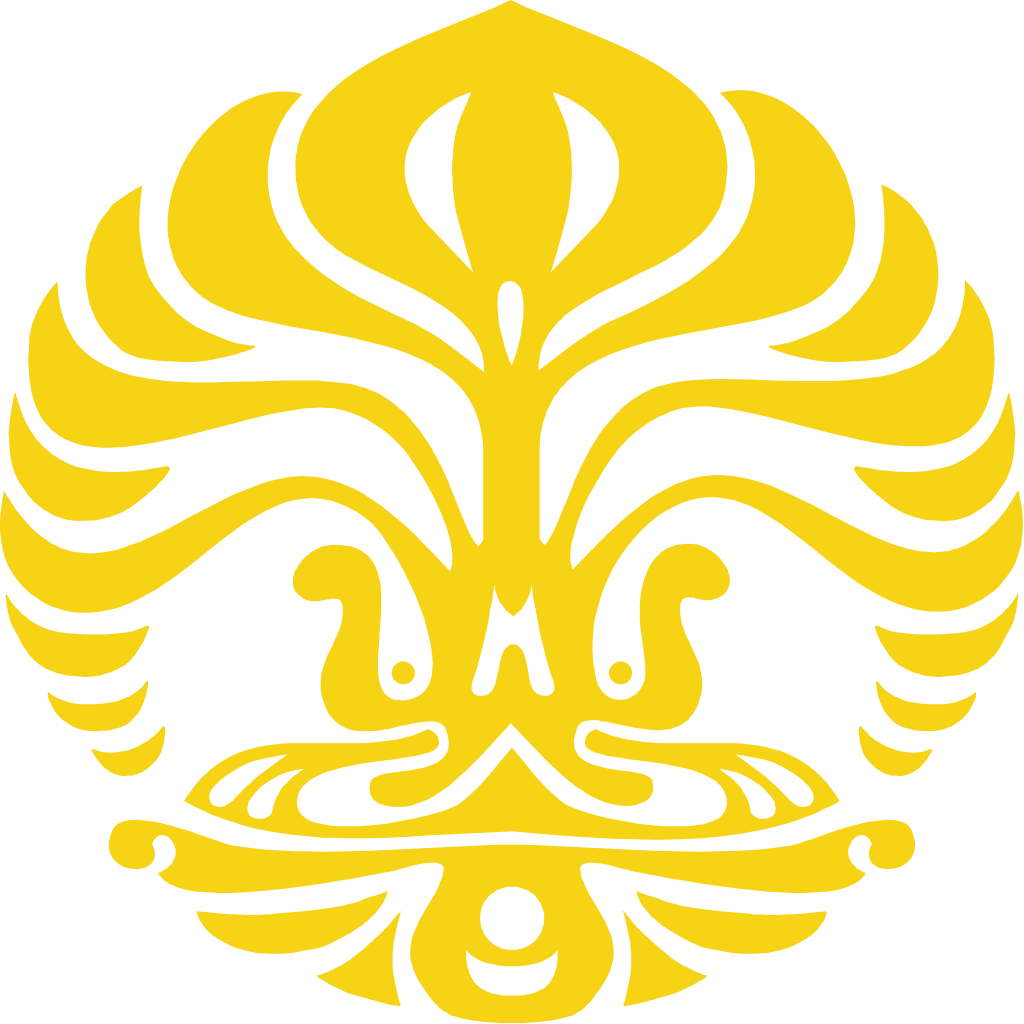
**RESUME DASAR-DASAR ORGANISASI INFORMASI**

**KATEGORISASI**



* **Aldi Rizki Padila (1906390001)**
* **Anis Mufida (1906293663)**
* **Fauzan Eka Kusuma (1906392285)**
* **Fina Idamatussilmi (1906304912)**
* **Hanifah Basirah Azizah Fahimah (1906390084)**
* **Muhammad Fadhil Sulaiman (1906364035)**
* **Syahrani Nur Azizah (1906293796)**

**KATEGORISASI**

1. **APA ITU KATEGORISASI, KLASIFIKASI, TAKSONOMI**

Beberapa definisi yang disediakan oleh Oxford English Dictionary (OED):

* Kategorisasi: untuk ditempatkan dalam kategori atau kategori; untuk mengklasifikasikan.
* Kategori: istilah..diberikan pada kelas umum istilah, hal, atau gagasan; kelas, atau divisi, dalam skema klasifikasi umum apa pun
* Klasifikasi: tindakan mengklasifikasikan atau mengatur kelas, sesuai dengan karakteristik bersama atau afinitas yang dirasakan; distribusi sistematis, alokasi, atau pengaturan hal-hal dalam sejumlah kelas yang berbeda
* Kelas: satu set atau kategori hal yang memiliki beberapa sifat atau atribut yang sama, dikelompokkan

Merriam-Webster mendefinisikan klasifikasi sebagai "pengaturan sistematis dalam kelompok atau kategori sesuai dengan kriteria yang ditetapkan." 2 Pada akhirnya, tidak ada sumber yang membantu membedakan antara klasifikasi dan kategorisasi

Secara umum, kategorisasi dianggap lebih luas dan lebih abstrak daripada klasifikasi; kategorisasi adalah fungsi kognitif yang digunakan untuk mengelompokkan konsep, daripada proses terstruktur yang digunakan untuk mengatur sumber daya fisik secara sistematis. Kategorisasi dapat dilihat sebagai pengelompokan yang amorf atau kurang jelas; sedangkan klasifikasi dapat dipandang sebagai struktur hierarki yang komprehensif untuk mengatur sumber daya informasi di rak linear. Elin Jacob menunjukkan bahwa dalam bidang LIS klasifikasi istilah dapat memiliki beberapa pengertian. Istilah ini dapat merujuk pada tiga konsep yang berbeda tetapi terkait:

* Sistem kelas, yang dipesan sesuai dengan serangkaian prinsip yang telah ditentukan dan digunakan untuk mengatur serangkaian entitas;
* Grup atau kelas dalam sistem klasifikasi; dan
* Proses penugasan entitas ke kelas dalam sistem klasifikasi.3

Jacob juga membandingkan kategorisasi dan klasifikasi: Meskipun sistem klasifikasi dan kategorisasi keduanya merupakan mekanisme untuk menetapkan urutan melalui pengelompokan fenomena terkait, perbedaan mendasar di antara mereka memengaruhi bagaimana urutan dipengaruhi-perbedaan yang membuat perbedaan dalam konteks informasi yang ditetapkan oleh masing-masing sistem ini. Sementara klasifikasi tradisional sangat ketat karena mengharuskan suatu entitas adalah anggota atau bukan dari kelas tertentu, proses kategorisasi bersifat fleksibel dan kreatif dan menarik asosiasi tidak mengikat antara entitas-asosiasi yang didasarkan bukan pada seperangkat prinsip yang telah ditentukan sebelumnya. tetapi pada pengakuan sederhana atas kesamaan yang ada di set entitas.

Dent menyimpulkan, kelas "adalah pengelompokan entitas yang didefinisikan di mana anggota memenuhi definisi kelas dan dapat didaftar" dan kategori "adalah label kognitif yang diterapkan pada pengelompokan entitas yang tidak dapat dihitung. di mana keanggotaan ditentukan oleh kekhasan di antara anggota dan bukan definisi menyeluruh. " Meskipun sedikit yang setepat Dent atau Jacob dalam membedakan antara dua konsep, perbedaan-perbedaan ini sangat membantu dalam memahami diskusi kategori dan kelas yang mengikuti

Kata taksonomi berasal dari taksi Yunani (pengaturan, tata tertib) dan nomos (hukum). Taksonomi telah ada di dunia hierarki ilmu yang ketat sejak Aristoteles, dalam Historia Animalium, menciptakan taksonomi kerajaan hewan. Taksonomi ilmiah paling terkenal Namun, adalah Linnaeus's Systema Naturae, klasifikasi tumbuhan dan hewan dari tahun 1735. Taksonomi yang sedang dibangun saat ini untuk digunakan dalam intranet, meskipun demikian, tidak harus mengikuti aturan ketat seperti itu, juga tidak selalu hierarkis. tampaknya memiliki kesamaan adalah kategori. Oleh karena itu, diskusi kategorisasi dalam bab ini juga berlaku untuk taksonomi.

Daftar taksonomi yang ditemukan online sering kali mencakup skema klasifikasi tradisional, daftar judul subjek, peopleonomies, dan ontologi, serta alat khusus subjek yang benar-benar menyebut diri mereka taksonomi Contoh yang terakhir adalah Taksonomi Teknologi Pendidikan12 dan Taksonomi GRIN13 dari Sistem Plasma Nutfah Tumbuhan Nasional AS. Sejumlah taksonomi merupakan hak milik organisasi yang membuatnya dan berada di bawah hak cipta, sehingga tidak dapat digunakan secara bebas dalam situasi lain.

1. **TEORI KATEGORISASI**

Teori Kategori ini sudah ada sejak masa Yunani Kuno. Misalnya, Filsuf kuno dan ahli matematika Phytagoras mengkategorikan siklus kehidupan yang terdiri dari kelahiran, pertumbuhan, pembusukan, kematian, penyerapan, dan metamorposis. Selain itu, Empedocles menempatkan setiap benda fisik ke dalam salah satu dari empat elemen diantaranya, bumi, udara, api, dan air.

Teori Kategori yang diciptakan oleh Aristoteles paling berpengaruh dalam diskusi pengkategorisasian dan merupakan upayanya mengatur objek serta ide dunia.

**Kebangkitan dan Runtuhnya Teori Kategori Klasik**

Akar dari sistem klasifikasi kontemporer dapat ditelusuri kembali ke teori kategorisasi klasiknya Aristoteles. Kata kategori berasal dari bahasa yunani *“kategorein”* yang berarti menuduh, menegaskan, dan predikat.

Teori Aristoteles mencerminkan definisi lainnya, ada 10 status makhluk atau 10 hal yang bisa diungkapkan mengenai suatu objek atau ide. Berikut pengkategoriannya:

* Zat
* Kuatitas
* Hubungan
* Tempat
* Waktu
* Posisi
* Negara
* Aksi
* Kasih sayang

Teori tersebut menempatkan objek atau ide ke dalam kategorisasi yang sama berdasarkan apa yang mereka miliki. Pendekatan ini tidak tertandingi sampai pertengahan abad kedua puluh. Seperti kata George Lakoff, kategori dianggap dipahami dengan baik. Sampai saat itu kategori dianggap sebagai wadah abstrak keketatan keanggotaan biner, baik barang-barang kepunyaan yang ada di dalam maupun di luar wadah.

Kategori juga didefinisikan sebagai kelompok yang saling ekslusif, yaitu segala hal menjadi satu kepunyaan ataupun satu kategori, serta atribut umum dari anggota kelompok. Pandangan sempit mengenai kategorisasi ini mulai berubah sekitar 65 tahun yang lalu.

**Keretakan dalam Teori Kategori Klasik**

Dalam sejarahnya singkat yang dirangkum oleh *Lakoff’s Women, Fire, and Dangerous Things, What Categories Reveal about the Mind.*

* Ludwig Wittgenstein

Menunjukan bahwa kategorisasi seperti permainan yang tidak cocok dengan cetakan klasik. Kategori ini tidak memiliki satu koleksi properti pada umumnya.

* J.L Austin

Ia memperluas analisis Wittgenstein untuk mempelajari kata-kata. Dia mempertanyakan “mengapa kita memanggil hal-hal yang berbeda dengan nama yang sama”, misalnya foot of a mountain, foot of a list, person’s foot. Bukankah seharusnya gunung *(mountain),* daftar *(list),* dan orang *(person)* berada dalam kategori yang sama jika mereka semua memiliki kaki *(foot)?*

* Lotfi Zadeh

Menyumbangkan teori himpunan tidak jelas dari pengelompokan teori klasik kategori. Dia mencatat bahwa beberapa kategori didefinisikan dengan baik sementara yang lainnya, tidak. Baik itu anggota klub atau bukan, misalnya apakah seseorang tinggi atau tidak tergantung sampai batas tertentu dari pengamat. Kategori “tall” dinilai—tidak jelas tidak ataupun pendek. Bisa saja menurut orang yang lebih pendek memandang seseorang dengan tinggi sedang, mungkin mereka mengatakan bahwa orang itu tinggi.

* Studi Floyd Lounsbury

Mengenai system kekerabatan penduduk asli Amerika juga menyinggung teori klasik. Dia menemukan bahwa diantara berbagai kelompok nama yang sama (kategori) digunakan untuk mengekspresikan hubungan kekerabatan dari beberapa jenis kerabat yang berbeda.

* Brent Berlin and Paul Kay

Dalam karyanya yang menemukan dua sampai sebelas ekspresi untuk warna dasar dalam berbagai bahasa. Hingga pada saatnya, telah mengidentifikasi warna dasar ke-12 dalam berbagai bahasa.

* Roger Brown

Mempelajari kategori tingkat dasar. Menurutnya tingkat pertama itu tempat anak-anak mempelajari kategori, hal ini dianggap Brown sebagai level alami. Hingga ke sejumlah nama yang lebih spesifik maupun yang lebih umum yang dianggap sebagai level pencapaian imajinasi.

Teori Prototipe

* Keretakan utama dalam teori klasik kategori muncul ketika Eleanor Rosch mengembangkan *theory prototype* dengan karyanya antara 1973 dan 1981. Dia berteori bahwa jika, seperti teori klasik menyatakan, kategori didefinisikan hanya dengan properti yang dibagikan oleh semua anggota, maka tidak ada anggota yang harus menjadi contoh kategori yang lebih baik daripada anggota lainnya. Lebih lanjut ia berteori bahwa jika kategori hanya ditentukan oleh properti yang dimiliki semua anggota berbagi, maka kategori harus independen dari manusia yang melakukan kategorisasi.
* Orang yang berbeda akan memasukkan hal yang berbeda ke dalam kategori seperti perlengkapan berkemah tergantung pada pengalaman mereka, ke mana mereka pergi, bagaimana mereka akan berkemah, dan sebagainya.

Karena skema klasifikasi yang paling banyak digunakan di Amerika Serikat didasarkan pada teori kategori klasik, penggunaan klasifikasinya kadang-kadang cukup frustrasi untuk menemukan bahwa mereka memiliki konsep subjek yang tidak pas masuk ke dalam salah ssatu kategori. Misalnya, multifaset, lintas disiplinsumber daya informasi tidak dapat dengan mudah ditampung oleh struktur dan kekakuan kategori klasik.Penggolong ini sering lega mengetahui bahwa teori klasik memiliki celah besar; Jadi bukan hal yang aneh untuk dijalankan menjadi non-prototipe dan *set fuzzy.*

1. **KLASIFIKASI BIBLIOGRAFI**

Salah satu jenis klasifikasi adalah klasifikasi bibliografi, yang bertujuan untuk mengatur dan menemukan kembali sumber informasi lalu digunakan untuk mengatur rekor metadata dalam katalog perpustakaan dan alat temu kembali informasi lainnya.

Contoh-contoh klasifikasi bibliografi utama adalah DDC (Dewey Decimal Classification), LCC (Library of Congress Classification), serta UDC (Universal Decimal Classification). Masing-masing memiliki skema klasifikasi yang berbeda-beda.

**Skema Klasifikasi Bibliografi Utama**

Skema klasifikasi bibliografi utama (mis., DDC, LCC, UDC) memiliki komponen-komponen tertentu yang umum. Ini telah diuraikan oleh Hope Olson dan John Boll sebagai berikut:

* Deskripsi verbal, topik demi topik, tentang hal-hal dan konsep yang dapat direpresentasikan dalam atau oleh skema.
* Susunan uraian verbal ini dalam urutan kelas atau logis yang dimaksudkan untuk memungkinkan pengaturan topik yang berarti dan akan nyaman bagi pengguna.
* Sebuah notasi yang muncul di samping setiap deskripsi verbal, yang digunakan untuk mewakilinya dan yang mana menunjukkan urutannya. Seluruh kelompok deskripsi dan notasi verbal membentuk jadwal.
* Referensi dalam jadwal untuk memandu classifier dan pencari ke berbagai aspek topik yang diinginkan atau ke topik terkait lainnya (seperti referensi istilah terkait dalam daftar alfabet).
* Indeks alfabet istilah yang digunakan dalam jadwal, dan sinonim dari istilah-istilah itu, yang mengarah ke notasi.
* Instruksi untuk penggunaan. Instruksi umum (dengan contoh) biasanya dapat ditemukan di awal skema, dan instruksi yang berkaitan dengan bagian-bagian tertentu dari jadwal diberikan, atau harus, diberikan di bagian yang mereka hubungkan.
* Organisasi yang akan memastikan bahwa skema klasifikasi dipertahankan, yaitu, direvisi dan diterbitkan ulang. Ini di luar skema, tetapi merupakan faktor penting dalam mengevaluasi komparatifnya kegunaan.

Selain jadwal yang disebutkan dalam poin ketiga di atas, skema termasuk tabel, yaitu daftar independen notasi untuk konsep yang dapat ditambahkan atau ditambahkan ke notasi dari jadwal. Tabel tersebut berisi daftar konsep seperti bentuk / genre, wilayah geografis, kronologis periode, dan bahasa. Selain instruksi dalam skema itu sendiri, ini daftar pustaka utama klasifikasi memiliki manual terpisah yang memberikan instruksi dan contoh lebih lanjut. Manual dapat ditulis oleh organisasi administrasi yang bertanggung jawab atas pemeliharaan skema, oleh pihak independen, atau oleh keduanya.

Traugott Koch telah menyarankan bahwa ada empat jenis skema klasifikasi:

* Skema universal — skema yang mencakup semesta pengetahuan (mis., DDC, UDC, LCC, CC, BC2);
* Skema umum nasional — mirip dengan skema universal, tetapi dirancang untuk digunakan di satu negara (mis., Nederlandse Basisclassificatie di Belanda dan Sveriges Allmänna Biblioteks mencari di Swedia);
* Skema khusus subjek — skema yang mencakup area subjek tertentu atau domain pengetahuan (mis., Iconclass untuk sumber daya seni; Perpustakaan Nasional Klasifikasi Kedokteran, dan Subjek Matematika Klasifikasi untuk materi pelajaran mereka yang jelas);
* Skema yang dikembangkan sendiri — skema yang dibuat untuk memenuhi kebutuhan layanan atau institusi tertentu (mis., Metis: Klasifikasi Perpustakaan untuk Anak-Anak, seperangkat 26 kategori luas berlabel A hingga Z didirikan oleh empat pustakawan yang bekerja di Ethical Culture School di New York City) .

1. **KONSEP-KONSEP KLASIFIKASI**

Sejumlah konsep dan masalah klasifikasi memengaruhi penggunaan skema klasifikasi. Beberapa dari ini berlaku untuk skema terlepas dari bagaimana mereka akan digunakan (mis., Sebagai cara mengatur sumber daya fisik, sebagai cara mengidentifikasi konten subjek dalam metadata, sebagai cara mengatur dan / atau menyajikan sumber daya online). Lainnya adalah masalah khususnya dalam penggunaan klasifikasi sebagai perangkat untuk mengatur sumber daya fisik. Bagian berikut ini diatur dari umum ke spesifik. (Taylor, 2018)

* ***Broad versus Close Classification***

***Broad Classification*** adalah klasifikasi yang hanya menggunakan kelas utama dan divisi skema dan mungkin hanya satu atau dua tingkat subdivisi. ***Close Classification*** adalah klasifikasi yang menggunakan semua subdivisi kecil yang tersedia untuk mata pelajaran yang sangat spesifik.

* ***Classification of Knowledge versus Classification of a Particular Collection***

***Classification of Knowledge*** adalah konsep bahwa sistem klasifikasi dapat dibuat yang akan mencakup semua pengetahuan yang ada. DDC dimulai sebagai klasifikasi pengetahuan — setidaknya pengetahuan Barat sebagaimana dipahami oleh Melvil Dewey. ***Classification of a Particular Collection*** adalah konsep bahwa sistem klasifikasi hanya boleh dirancang untuk sumber daya informasi yang ditambahkan ke koleksi, menggunakan konsep ***literary warrant***. ***literary warrant*** adalah nama yang diterapkan pada konsep yang notasi baru dibuat untuk skema klasifikasi dan istilah baru ditambahkan ke kosakata terkontrol hanya ketika sumber daya informasi benar-benar ada yang tentang konsep baru. LCC dimulai sebagai klasifikasi koleksi tertentu.

* ***Integrity of Numbers versus Keeping Pace with Knowledge***

***Integrity of Numbers*** adalah konsep bahwa dalam pembuatan dan pemeliharaan skema klasifikasi, notasi, setelah ditetapkan, harus selalu mempertahankan makna yang sama dan tidak boleh digunakan dengan makna lain. ***Keeping Pace with Knowledge*** adalah konsep bahwa dalam penciptaan dan pemeliharaan skema klasifikasi, diakui bahwa pengetahuan berubah, dan oleh karena itu perlu untuk dapat memindahkan konsep, memasukkan konsep baru, dan mengubah makna angka.

* ***Closed versus Open Stacks***

***Closed Stacks*** adalah nama yang diberikan pada situasi di mana area penyimpanan sumber daya informasi hanya dapat diakses oleh staf perpustakaan, arsip, atau tempat lain yang menampung sumber daya informasi. ***Open Stacks*** adalah nama yang diberikan untuk situasi yang berlawanan, di mana pelanggan fasilitas memiliki hak untuk pergi ke area penyimpanan itu sendiri.

* ***Fixed Location versus Relative Location***

Lokasi tetap berartikan sebuah lokasi sebagai tempat untuk menyimpan sumber informasi dalam bentuk fisik akan **selalu ditemukan** atau selalu kembali ke lokasi semula sesudah dipakai. Pengidentifikasian lokasi tetap dapat berupa nomor akses; atau desain yang terbuat dari nomor ruang, nomor tumpukan, nomor rak, posisi rak; ataupun desain lainnya. Sebaliknya, lokasi relative berartikan bahwa sumber informasi akan berada di tempat yang berbeda setiap kali mereka kembali di rak. Sumber informasi akan diperlakukan seperti ini jika ada informasi baru yang diterima, dikeluarkan, dikembalikan, dan seterusnya setiap mereka kehilangan guna.

* ***Location Device versus Collocation Device***

Kontoversi lain yang menyelimuti klasifikasi adalah berupa klasifikasi yang digunakan sebagai alat kolokasi atau hanya berupa alat lokasi. Sebuah alat lokasi adalah nomor atau desain lain untuk memberitahu dimana fisik sumber informasi itu berada. Hal itu bisa berada disekitar desain-desain atau nomor akses, nomor akses fisik, atau nomor panggilan. Alat kolokasi adalah nomor atau desain lain disebuah sumber informasi yang digunakan untuk menempatkannya setelah sumber lainnya yang mirip. Biasanya ini adalah sebuah klasifikasi notasi.

* ***Klasifikasi Serial versus Klasifikasi Alfabet.***

Serial kadang disebut jurnal dan majalah. Sebuah serial dapat diartikan sebagai publikasi yang ditanggungkan untuk selalu berlanjut. Klasifikasi serial berarti notasi klasifikasi dipakai oleh serial, dan notasi klasifikasi ini dipasangkan pada setiap volume dan atau setiap bagian dari serial tersebut. Serial secara alphabet berarti ditempatkan secara teratur di rak, atau disusun menurut abjad dari judul serial. Masalah mengenai serial klasifikasi dan alphabet ini hanya diterapkan pada serial tercetak, lain halnya jika serial yang ada hanya berbentuk online, akan timbul banyak masalah dan pertanyaan. Tetapi klasifikasi tidak menjadi bagian dari masalah tersebut.

* ***Klasifikasi monografik ( Klasifikasi terpisah versus Klasifikasi sepaket)***

Monograf adalah unit pustaka sumber informasi yang lengkap. Terkadang hanya ada satu kali berkarya, tapi terkadang ada banyak lagi karya-karya yang dikerjakan, perbedaannya lebih jelas dengan serial pada bagian kelanjutan karya, monograf bermaksud untuk menyelesaikan karyanya secara langsung tidak seperti serial. Seri monografik adalah gabungan antara serial dan monograf.

1. **PENGGUNAAN KATEGORISASI DAN TAKSONOMI ONLINE**

Penggunaan kategori atau taksonomi dilakukan secara online. Di situs belanja komersial seperti Amazon.com atau situs web untuk pengecer skala besar, seperti toko grosir yang dikenal sebagai Wegmans , taksonomi sering ditemukan langsung di beranda atau di menu tarik-turun (atau keduanya). Taksonomi ini digunakan untuk memberikan kemudahan dalam pembelian dengan cepat, set tautan yang ramah pengguna ke divisi utama dalam rangka meningkatkan navigasi dan pengalaman keseluruhan situs, serta kemampuan untuk menemukan informasi dan produk dengan cepat dan efisien.

Berikut ini adalah kategori tingkat atas yang disediakan di taksonomi Wegman, antara lain :

* Toko Roti
* Minuman
* Makanan
* Kedai Kopi
* Kebutuhan Rumah Tangga

Tingkat hierarki kedua berisi kumpulan judul tambahan. Disini mengambil contoh roti sebagai tingkat kedua menawarkan kategori berikut untuk dijelajahi, antara lain :

* Roti Panggang
* Roti Dikemas
* Makanan Penutup

Tingkat hierarki ketiga meliputi sub-kategori yang lebih detail lagi. Disini mengambil contoh dari kategori makanan penutup, antara lain :

* Brownies
* Kue kering
* Tart

Taksonomi juga digunakan oleh situs-situs non-komersial, seperti pusat data publik dan Pusat Informasi Bioteknologi. Contoh yang paling terkenal dari menggunakan kategorisasi online adalah Yahoo! Direktori dimana situs web ditempatkan ke dalam kategori yang dibuat oleh Yahoo! pengindeks, dan kategorinya ditelusuri secara hierarkis. Selain Yahoo! Direktori terdapat contoh lain yang masih menggunakan pendekatan ini untuk mengatur daftar sumber daya internet yaitu Best of the Web directory31 dan WWW Virtual Library. Seperti yang terlihat di Wegmans , sub-kategori dalam “Best of the Web” yang secara hierarki terdapat di bawah ini level atas kategori, antara lain :

* Seni
* Bisnis
* Komputer

Namun, hierarki bukanlah hierarki yang benar dalam arti bahwa mereka mewakili hubungan genus-spesies yang sebenarnya. Dengan kata lain, kategori-kategori pada level hierarki kedua tidak selalu merupakan tipe atau jenis konsep yang lebih luas pada level pertama dan tidak ada sub-kategori pada level kedua yang harus pada level yang setara dengan spesifisitas satu sama lain. Sebagai contoh, terdapat sub-kategori perkumpulan, antara lain :

* Ekonomi
* Pendidikan
* Bahasa dan linguistik
* Ilmu Sosial
* Sosiologi
* Laki-laki
* Perempuan

Beberapa sub-kategori jelas terkait dengan yang lain di tingkat atas hierarki. Misalnya, ekonomi dan bahasa dan linguistik ditempatkan di bawah ilmu sosial, yang terdaftar sebagai sub-kategori di bawah sains dan masyarakat. Sedangkan pria dan wanita ditempatkan di bawah orang, yang terdaftar sebagai sub-kategori di bawah people.

1. **KATEGORISASI HASIL PENCARIAN**

Beberapa peneliti mempercayai bahwa pencarian yang memuaskan dan efektif dapat berimprovisasi jika teknik klasifikasi diaplikasikan ke mesin pencarian. Tentunya hal tersebut bukanlah ide yang baru karena pada akhir tahun 1990-an dn awal 2000-an, Northern Light, sebuah mesin pencarian yang gratis dan tersedia secara umum, menggunakan teknik pengelompokan dokumen untuk membagi hasil pencarian ke dalam kategori topik yang luas.

Meskipun tidak bekerja langsung dgn sistem klasifikasi formal, northern light menggunakan pengelompokan dokumen untuk meningkatkan kemampuan pengguna dalam untuk mengemudikannya melewati kelompok hasil. Saat itu, hanya northern light yang menggunakan cara ini untuk menampilkan hasil hasil. Tahun 2001, Northern Light berhenti memberikan web mesin pencarian yang umun dan berfokus kepada perusahaan komersial.

Sejak saat itu, penelitian terus berlanjut pada teknik klasifikasi secara otomatis, dan pengelompokan lainnya dalam mesin pencarian pun telah dikembangkan. Contohnya, Carrot menawarkan beberapa pendekatan untuk hasil pengelompokan, terbagi dalam bentuk gambar dan berdasar pada teks. Yippy menyediakan hasil dengan halaman individu bersamaan dengan sebuah daftar kategori untuk menolong para pengguna mempertajam atau memfokuskan kasil pencarian mereka.

Semua proyek ini menggunakan teknik pengelompokan dokumen, apa definisinya? Menurut Srividhya Subramanian dan Keith Safer, pengelompokan adalah pendekatan ilmu komputer untuk mengklasifikasi dokumen di sebuah koleksi, berdasarkan konten atau isi dari seluruh koleksi.