
Taxonomieën

Globaal op het programma

- > **Wat is een taxonomie?**
- > **Waarom een taxonomie?**
- > **Wat is een goede taxonomie?**
- > **Werken met taxonomieën**

- > **Semantiek**

Algemene opmerkingen en termen

Wat is taxonomie?

- > De wetenschap van het indelen (in groepen)

Zowel

- > classificatie van ...
- > als: methode van classificatie

- > Voor alles is een taxonomie te maken

Meer definities

- > het ordenen, rangschikken en classificeren van de beschikbare kennis in een bepaalde wetenschap (www.mijnwoordenboek.nl)
- > een structuur die het mogelijk maakt om content over personen, organisaties, gebeurtenissen en dingen te clusteren in (hiërarchische) groepen zodat ze gemakkelijk te identificeren, te bestuderen en terug te vinden zijn.

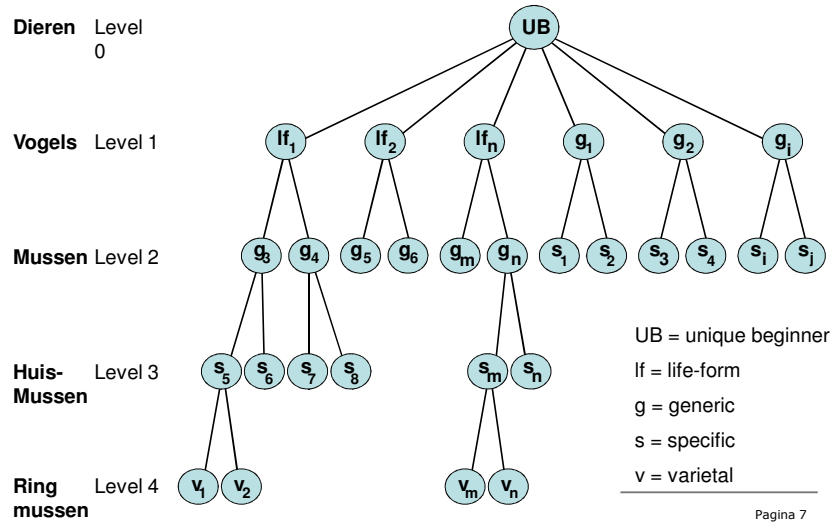
Een taxonomie definieert objectklassen en de onderlinge relaties.

(<http://labyrinth.rienkjonker.nl/glossary>)

Of anders....

- > Leer van de systematische indeling van planten en dieren.
- > Ordening en naamgeving van soorten.
Vooral de Zweed Linnaeus heeft bijgedragen aan het opzetten van een systematisch systeem voor de naamgeving.
- > Biologie die zich bezighoudt met leer en studie van de biologische systematiek.
- > Studie of leer van de onderliggende verwantschap van organismen. Het classificeren van planten- en diersoorten.

Voorbeeld biologische taxonomie



Taxonomie

> Indexsysteem

Indexeren

proces van

het vaststellen van zoekkenmerken

om het terugzoeken van informatie te vergemakkelijken

- > Indextaal heeft:
 - eigen vocabulaire
 - met eigen grammatica
(verbinden van indextermen tot grotere complexiteiten)

- > Ontsluitingssysteem

2 kanten van ontsluitingssystemen

- > Ordenen

- > Terugzoeken(/vinden)

Objecten en sferen

- > Sfeergericht (deductieve) indexsystemen
 - classificaties
- > Objectgericht (inductieve) indexsystemen
 - trefwoordsystemen, thesauri
- > Onderwerp wel/niet in context

Enkelvoudig vs. complex

- > Onderwerpen bestaan uit één object of zijn te (ver)vangen door actie/proces of kenmerk
- > Onderwerpen worden gevormd door meerdere elementen

- > Fysieke omgeving / geautomatiseerde omgeving!

Pré- en postcoördinatie

> Précoördinatieve systemen:

- conventionele / 'handmatige' systemen
- onderwerp moet zoveel mogelijk in 1 (complexe) indexterm worden uitgedrukt
- dit gebeurt vooraf aan de zoekfase
- combinatievolgorde?

> Postcoördinatieve systemen:

- aantal te gebruiken indextermen wordt vastgesteld in de zoekfase
- indextermen bestaan uit één begripseenheid
- in zoekprofiel zijn verschillende metadata gelijkwaardig (niet bij voorbaat 'geavanceerd' zoeken)

Indexeerdiepte en specificiteit

- > Indexeerdiepte: mate van inhoudsanalyse
- > Specificiteit: mate van exacte omschrijving van de inhoud

Verbonden met:

- > Vangst (recall)
- > Relevantie (precision)

hier tussen: ruis!

Semantiek en syntaxis

Semantiek

- > Betekenisleer, inhoud

semantische relaties:

- gelijkwaardig , synoniemen (zie)
- hiërarchisch (genus-species, geheel-deel, bt, nt, zie ook)
- associatief

Syntaxis

- > Grammatica, zinsbouw, vorm
 - structuur/reglement van relaties tussen indextermen
 - maar ook: te bedenken door de gebruiker (bijv. bij formuleren zoekvraag)

Voorbeelden?

- > Bedenk voorbeelden van classificaties of taxonomieën uit de eigen omgeving

En/of

- > Wat groepeer je zo allemaal in het dagelijks leven?

Wat is een taxonomie?

- > Trefwoordenlijst?
- > Classificatie?
- > Thesaurus?
- > Ontologie?
- > Topic map?
- > ...

Trefwoordenlijsten

Een alfabetische lijst van trefwoorden, evt. met onderlinge relaties

Standaardtrefwoordenlijst: lijst van bij voorkeur te gebruiken trefwoorden, met vermelding van onderlinge relaties

Trefwoordenlijsten

- > Gecontroleerd
vaak afkomstig uit thesaurus
- > Ongecontroleerd
vrije termen

Woordenlijsten NEN-ISO / NPR-ISO/tr 15489

- > Lijst van toegelaten onderwerpsaanduidingen
- > Afgeleid van termen in het classificatieschema
- > Zonder betekenissen en zonder relaties
- > Bedoeld voor controle van naamgeving (acceptabel, gebruikelijk)
- > Regelt gebruik synoniemen, afkortingen e.d.

Voorbeeld trefwoordenlijst

arbeidsovereenkomsten
arbeidsvoorwaarden
archeologie
archieven
architectuur
architectuurwetenschap
artotheken
 zie: kunstuitleen
asielzoekers
 zie: vluchtelingen
assessment
 zie: toetsing
ateliers

Thesaurus

- > Gevalideerde trefwoordenlijst
(te bouwen volgens ISO 2788)
- > Gecontroleerde lijst van termen
- > Met semantische, hiërarchische, gelijkwaardige of
associatieve relaties (dus: met betekenis en relaties)
- > Hulpmiddel bij toekennen classificatietermen

Paardrijden

UFA

> [ruiterkampen](#)

BT

> [sporten](#)

RT

> [maneges](#)

RUB

> [S23.3 sport](#)

Uitleg

- > UF *Used for* ('gebruikt voor') wordt toegevoegd aan de term waarnaar USE heeft verwezen.

ouderenwerk

UF GOW

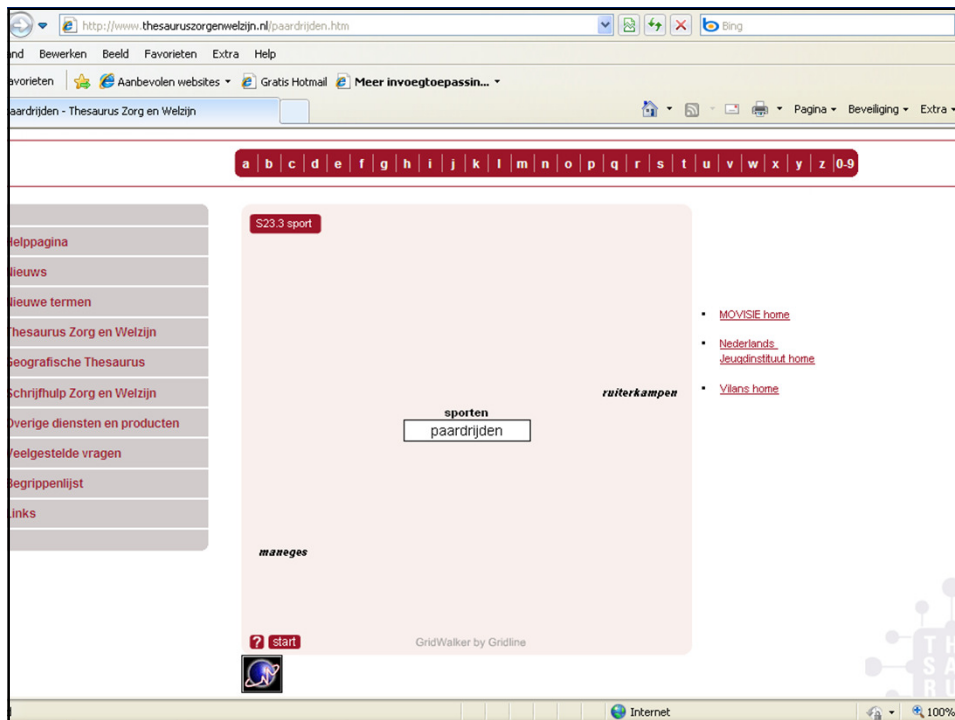
(GOW → USE ouderenwerk)

- > UFA *Used for and* ('in combinatie gebruikt voor') geeft aan dat deze voorkeursterm in combinatie met een andere voorkeursterm gebruikt wordt voor een niet-voorkeursterm.

arbeidsongeschikten

UFA WAO-platforms

- > BT De *broader term* is een bredere en meer algemene term, hiërarchisch hoger gelegen; er kunnen meer BT's bij een term voorkomen.
- > NT De *narrower term* is een nauwere, meer specifieke term, hiërarchisch lager gelegen; er kunnen meer NT's bij een term voorkomen.
- > RT De *related term* is een zie ook-verwijzing of associatieve term die de gebruiker attendeert op een inhoudelijk verwante term; er kunnen meer RT's bij een term voorkomen.



The screenshot shows a web browser window with the URL <http://www.thesauruszorgenwelzijn.nl/paardrijden.htm>. The browser's address bar and menu bar are visible. The page content includes a navigation menu on the left with items like 'Helppagina', 'Nieuws', 'Nieuwe termen', 'thesaurus Zorg en Welzijn', 'Geografische Thesaurus', 'Schrijfhelp Zorg en Welzijn', 'Overige diensten en producten', 'Geelgestelde vragen', 'Begrippenlijst', and 'Links'. A central search area displays 'S23 3 sport' and a list of related terms: 'sporten', 'paardrijden', and 'ruiterkampen'. A sidebar on the right contains links to 'MOVISIE home', 'Nederlands Jeugdinstituut home', and 'Vilans home'. The page footer includes a 'start' button and a 'GridWalker by Gridline' logo. The browser's status bar at the bottom shows 'Internet' and '100%' zoom level.

Classificaties

- > Enumeratief
- > Analytisch synthetisch

Voorbeeld enumeratief, SISO

- > 340-380 Economie
 - > 360 Bedrijfseconomie
 - > 366-367 Leer van de organisatie
 - > 367 Mens en arbeid in de onderneming
 - > 367.1 Bedrijfpsychologie en -sociologie
 - > ...
- > 390-398 Rechtswetenschap

Voorbeeld analytisch-synthetisch, UDC

159.9

Psychology

- 159.91 Psychophysiology (physiological psychology).
Mental physiology
- 159.92 Mental development and capacity. Comparative
psychology
- 159.93 Sensation. Sensory perception
- 159.94 Executive functions
- 159.95 Higher mental processes
- 159.96 Special mental states and processes

Deelclassificaties

Co-ordination. Addition

Relation. Subgrouping. Order-fixing

[Common auxiliaries of language](#)

[Common auxiliaries of form](#)

[Common auxiliaries of place](#)

[Common auxiliaries of ethnic grouping and nationality](#)

[Common auxiliaries of time](#)

Specification by non-UDC notation (*, A/Z)

Common auxiliaries of [properties](#), [materials relations](#), [processes and operations](#), [persons and personal characteristics](#)

Zo kun je gemakkelijk maken:

> 159.95 (07)=7

Studieboek

over 'Higher mental processes'

in een taal uit het gebied 'Indo-Pacific.Australian'

De Basis Archiefcode

> Enumeratief of analytisch synthetisch?

Classificatiesystemen NEN-ISO 15489

- > Weerspiegelen de activiteiten van de organisatie
- > Gebaseerd op een analyse van de activiteiten van de organisatie
- > Gebruikt om informatie- en archiefbeheerprocessen te ondersteunen
→ krachtig hulpmiddel!

Krachtig hulpmiddel:

- > Leggen van koppelingen tussen individuele archiefbescheiden
- > Bevorderen/veilig stellen consistentie in naamgeving van archiefbescheiden (door de tijd)
- > Terugvinden van archiefbescheiden

Krachtig hulpmiddel:

- > Passende beveiliging en toegang archiefbescheiden
- > Toekennen rechten van
 - toegang aan gebruikers tot, of
 - uitvoeren van activiteiten op specifieke groepen van archiefbescheiden
- > Verdelen van verantwoordelijkheden voor beheer

Krachtig hulpmiddel:

- > Distributie van archiefbescheiden om er iets mee te doen
- > Bepalen bewaartermijnen
- > Bepalen overdrachts- of vernietigingsacties

Bouw een dergelijke classificatie, die

- > Koppelingen legt tussen archiefbescheiden
- > Consistente naamgeving van archiefbescheiden veilig stelt
- > Terugvinden van archiefbescheiden ondersteunt met betrekking tot functie/activiteit
- > Passende beveiliging en toegang bepaalt
- > Rechten toekent mbt toegang tot en uitvoering van activiteiten met archiefbescheiden
- > Verantwoordelijkheid voor beheer verdeelt
- > Distributie van archiefbescheiden regelt
- > Passende bewaartermijnen bepaalt en acties mbt overbrenging en vernietiging

NPR-ISO/TR 15489-2:2001

- > Beschrijft een aanpak

Topic map

Semantische laag bovenop informatie

- > Identificeert de onderwerpen, of concepten, die in de informatie ligt opgeslagen
- > Legt relaties tussen deze concepten vast en omschrijft ze
- > Legt verbinding tussen de concepten en de onderliggende informatie

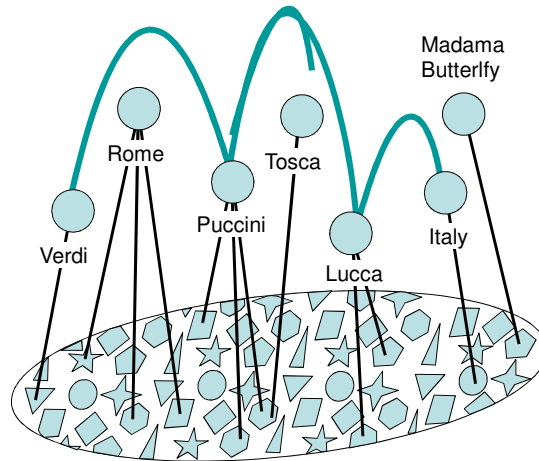
Voorbeeld (www.itadvanced.nl)

- > U zoekt naar Columbus, dat is een topic.
- > Columbus ontdekt Amerika.
- > Amerika is ook een topic.
- > Tussen deze twee topics ligt een relatie: ontdekken.
- > 'ontdekken' is hier een relatie, maar kan zelf ook een topic zijn.

Zo kan vanuit de topic 'ontdekken' weer een relatie gelegd naar 'ontdekkingen in de 15e eeuw' of 'ontdekkingen door Europeanen'.
Dus tijd en herkomst zijn weer relaties van 'ontdekking'.

Enz.

Voorbeeld topic map



Wat zou je hieruit kunnen afleiden?

- > Puccini
- > Geboren in Lucca
- > Studeerde aan conservatoto Verdi
- > Opera's Tosca en Madama Butterfly
- > Uitgevoerd in o.a. Rome
- > In Italië

Wat doet Topic Map nog meer?

- > Verwijzingen / doorverwijzingen naar andere bronnen met vermelding onderwerp van die bronnen:

Bij de geloofsovertuiging van Columbus:

'Zie ook: Katholicisme'.

- > Vermelding van vindplaatsen

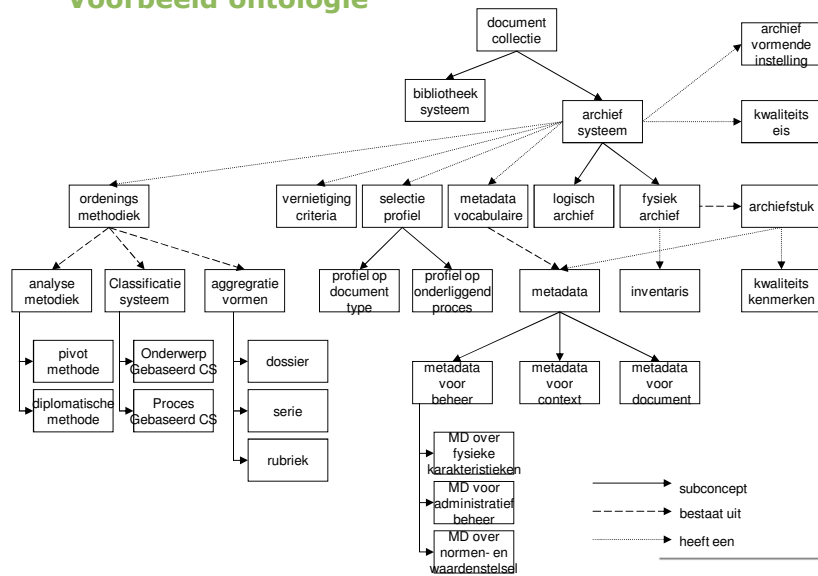
Bij Columbus verwijzingen naar:

- een standbeeld in Genua;
- een ISBN-nummer van een boek;
- een biografie die verbonden is aan een universiteit;
- een locatienummer van de biografie in een bibliotheek;
- een URL van een website over Columbus;
- het museum waar zijn scheepslogboeken liggen;
- enz.

Ontologie

- > Conceptualisering van een domein: een gestructureerd datamodel met concepten en mogelijke relaties tussen concepten die courant en belangrijk zijn in een bepaalde discipline of werkgebied.
- > Een vocabularium van alle woorden waarvan de onderlinge relaties vast te leggen zijn samen met een aantal regels voor gevolgtrekkingen en logica.
- > Leer van het zijnde, het zijn van het geheel van de dingen (filosofie, metafysica)

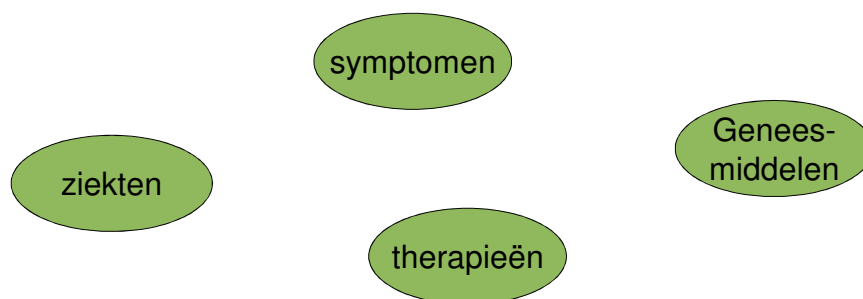
Voorbeeld ontologie



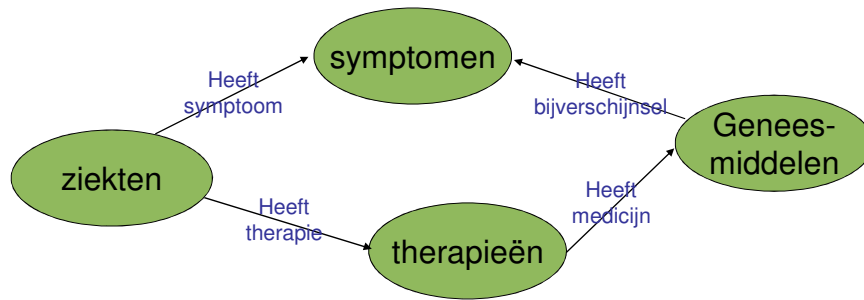
Formele ontologieën

- > Beschrijven een domein door 'concepten' EN hun relaties te definiëren:
- > Eigenschappen van concepten
 - > Naam, subs en supers, kenmerken + toegestane waarden
- > Eigenschappen van relaties
 - > Naam, domein, bereik, e.d.

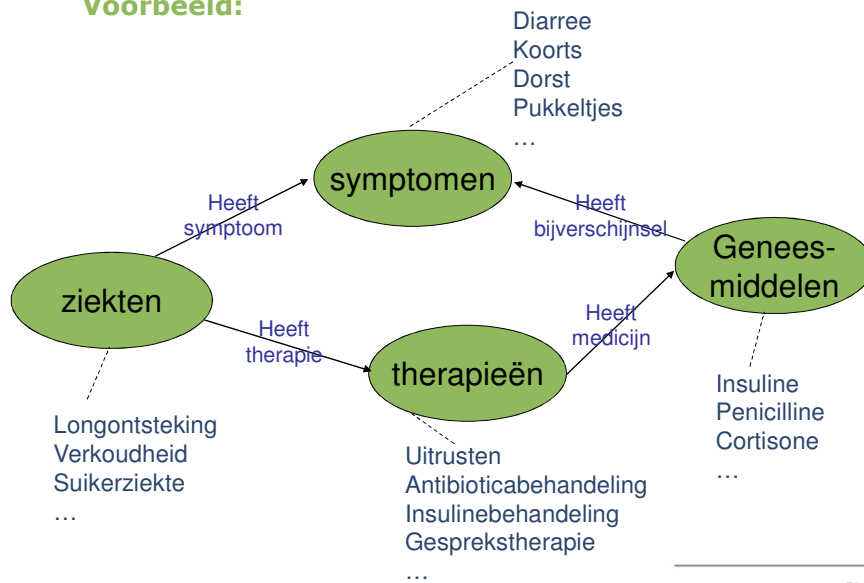
Voorbeeld:



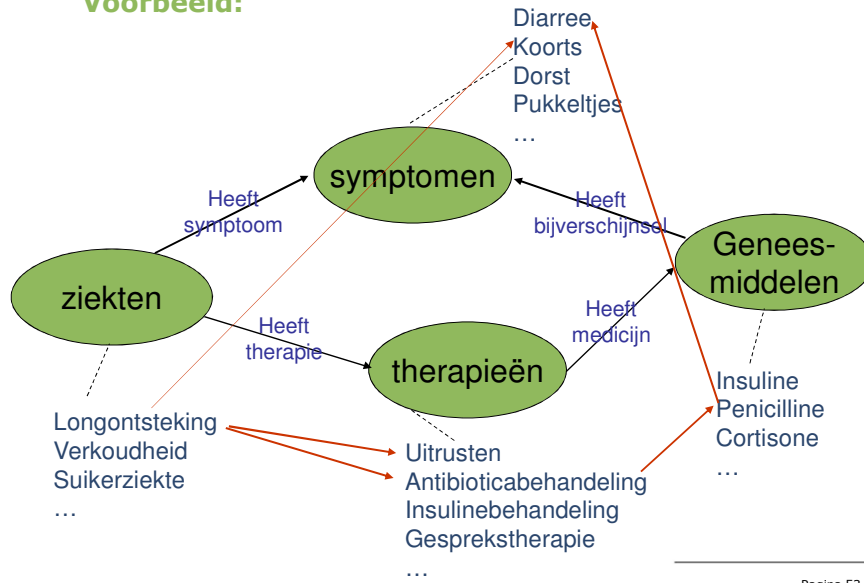
Voorbeeld:



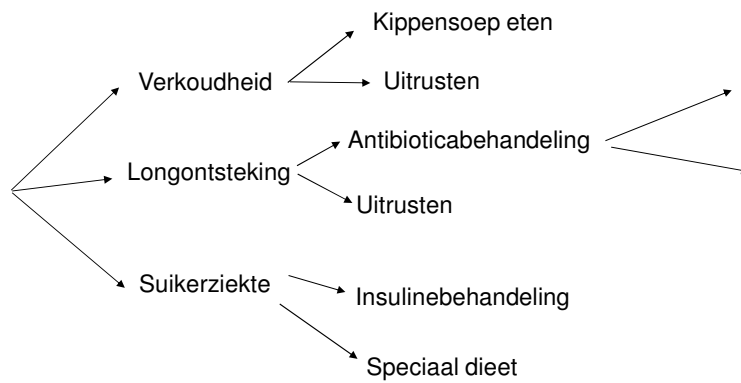
Voorbeeld:



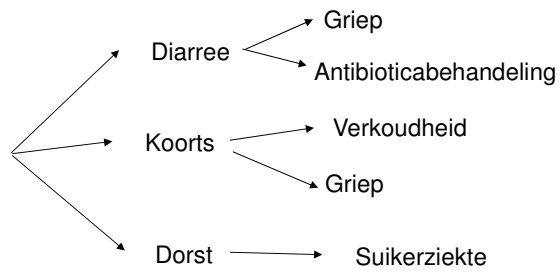
Voorbeeld:



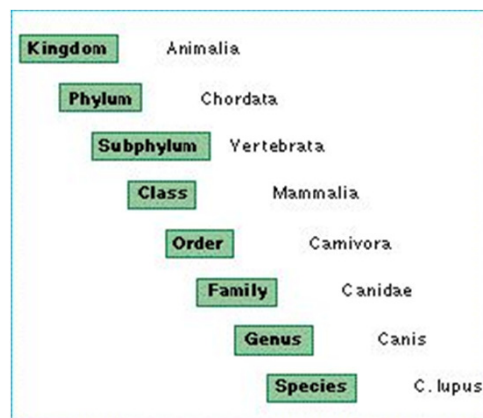
Ontologieën en taxonomieën



Ontologieën en taxonomieën



Toepassingsgebied: wetenschap



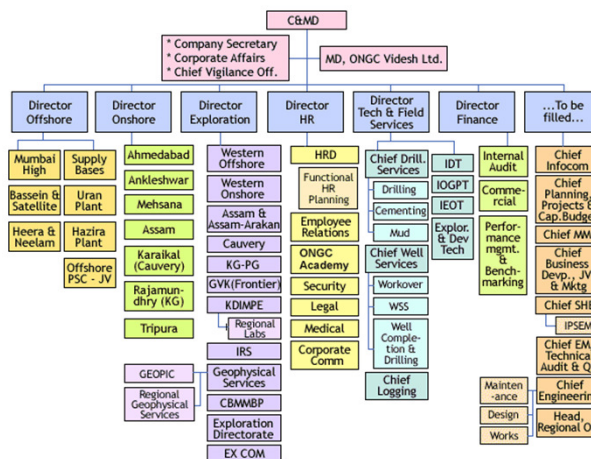
Wat wordt hier geclassificeerd?

Toepassingsgebied: wetenschap

1																	2
H																	He
3	4											5	6	7	8	9	10
Li	Be											B	C	N	O	F	Ne
11	12											13	14	15	16	17	18
Na	Mg											Al	Si	P	S	Cl	Ar
19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36
K	Ca	Sc	Ti	V	Cr	Mn	Fe	Co	Ni	Cu	Zn	Ga	Ge	As	Se	Br	Kr
37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54
Rb	Sr	Y	Zr	Nb	Mo	Tc	Ru	Rh	Pd	Ag	Cd	In	Sn	Sb	Te	I	Xe
55	56	57	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86
Cs	Ba	La	Hf	Ta	W	Re	Os	Ir	Pt	Au	Hg	Tl	Pb	Bi	Po	At	Rn
87	88	89	104	105	106	107	108	109	110								
Fr	Ra	Ac	Unq	Unp	Unh	Uns	Uno	Une	Unn								
58	59	60	61	62	63	64	65	66	67	68	69	70	71				
Ce	Pr	Nd	Pm	Sm	Eu	Gd	Tb	Dy	Ho	Er	Tm	Yb	Lu				
90	91	92	93	94	95	96	97	98	99	100	101	102	103				
Th	Pa	U	Np	Pu	Am	Cm	Bk	Cf	Es	Fm	Md	No	Lr				

Wat wordt hier geassocieerd?

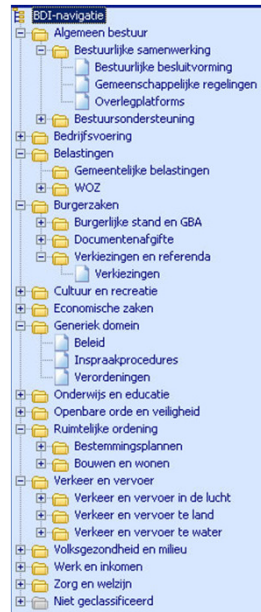
Toepassingsgebied: organisatie



Wat wordt hier geassocieerd?

Toepassingsgebied: organisatie

Wat wordt hier geclassificeerd?



Toepassingsgebied: informatie

- > Basisarchiecode
- > Universal Decimal Classification
- > Dewey Decimal Classification
- > AAT
- > ...

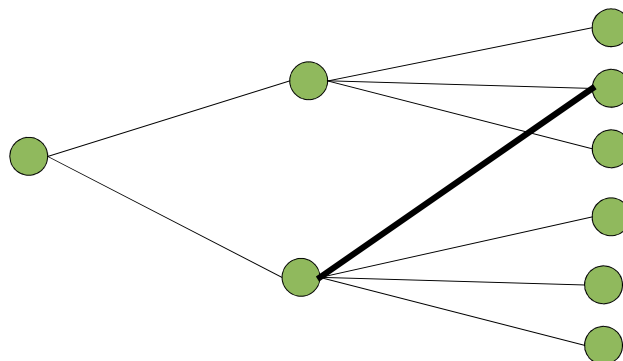
Wat wordt hier geclassificeerd?

Kenmerken van een taxonomie

- > Strikt hiërarchisch
- > Maar één soort relatie
- > Classificatie
- > Niveaus met betekenis
- > Weerspiegelt 'natuurlijke ordening'

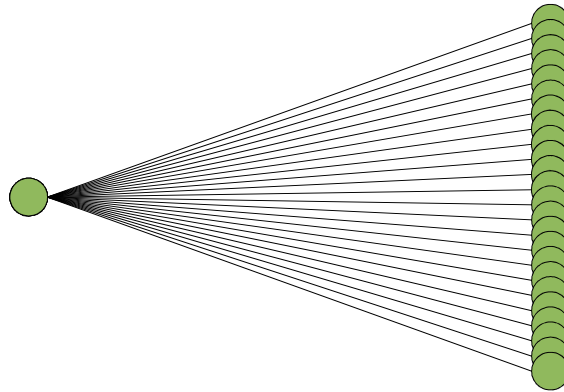
Syntactische criteria (1)

- > Heeft de taxonomie een zuivere boomstructuur?



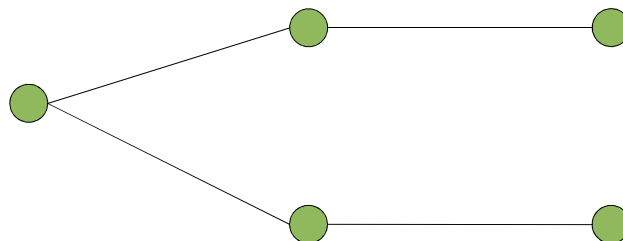
Syntactische criteria (2)

- > Is de omvang van categorieën binnen de grenzen?



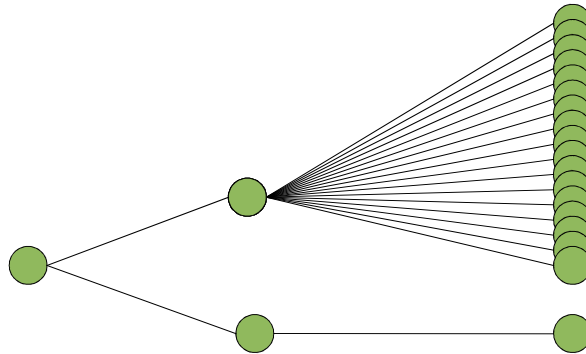
Syntactische criteria (3)

- > Past de diepte bij de omvang?



Syntactische criteria (4)

> Is de taxonomie in balans?



Terminologische criteria (1)

> Zijn telbare zelfstandige naamwoorden in meervoudsvorm?

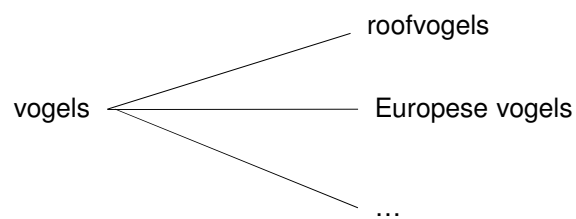
- > Bossen i.p.v. bos
- > Woods i.p.v. wood
- > Wood i.p.v. woods

Terminologische criteria (2)

- > Is er een consistente benaming van samengestelde termen?
 - > 'Onderwijsorganisatie' of 'Organisatie van het onderwijs'?
 - > 'Verkeer en vervoer' versus 'burger en overheid'

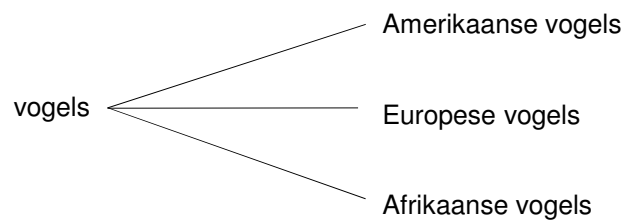
Semantische criteria (1)

- > Zijn de subcategorieën uitsluitend?



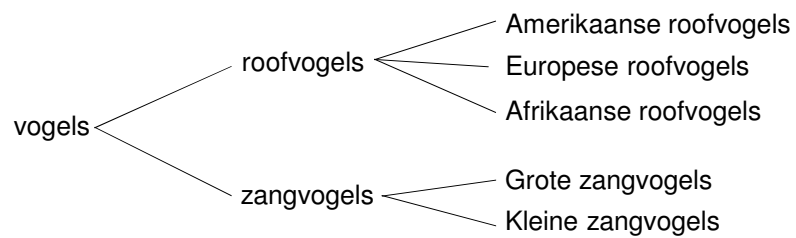
Semantische criteria (2)

> Zijn de subcategorieën uitputtend?



Semantische criteria (3)

> Is er systematiek in de subcategorisatie?



Onderhoudscriteria

- > Wat is de dynamiek van de taxonomie?
- > Wat zijn de consequenties van veranderingen in de taxonomie?

Pragmatische criteria

- > Door wie en hoe wordt taxonomie gebruikt?
- > In hoeverre is de gebruikte taxonomie begrijpelijk voor de doelgroep?

Opdracht:

- > Tijd: 30 minuten
- > Ontwikkel een taxonomie van tenminste 3 niveaus om documenten te classificeren die betrekking hebben op (een aspect van) een gekozen onderwerp.

SEMANTIEK

Classificatiesystemen voorzien een organisatie van een hulpmiddel om:

- Documenten te rangschikken, te beschrijven en te relateren
- Interdisciplinaire documenten met elkaar te verbinden en te delen, **binnen en buiten de organisatie**
- Verbeterde toegang, zoekmogelijkheden, **gebruik en verspreiding** van documenten te leveren, op de gewenste wijze

NORA en semantiek

- > Semantiek basiskeuze van NORA
 - bevorderen samenhang
- of: semantische samenhang / semantische interoperabiliteit

- > Informatiestromen hebben
 - bron
 - bestemming

Begrijpen beiden elkaar / interpreteren zij de informatie hetzelfde?

- betekenis van velden?
- bedoeling van bericht?
- symbolen op een beeldscherm?

Semantiek is inhoudelijke kern

Noodzakelijk voor functioneren van de e-overheid

Overschaduwd door:

- welke functie wordt waar belegd?
- aandacht voor bedrijfsprocessen door managers (los van de inhoud van de processen)

Semantiek: semantische modellen

- > En dat zijn: vocabulaires, taxonomie, ontologie, enz.!

Verschillen:

- hoeveelheid structuur die uitgedrukt wordt
- onderscheid statische en dynamische begrippen en mate van verbindingen daartussen

vb: er zijn burgers, er zijn huwelijken, voor huwelijk zijn 2 (ongehuwde) burgers nodig, status burgers verandert bij huwelijk

Overheidsambities

- > Meervoudig gebruik van gegevens
 - eenmalige aanlevering
 - hergebruik
- > Belangrijk:
 - intensieve afstemming
 - 100% duidelijkheid over de betekenis

Semantiek op stelselschaal

- > Betekent 'onderneming' hetzelfde voor
 - Kamer van Koophandel
 - Belastingdienst
 - ondernemingsraad?
- > Is uitwisseling van gegevens zinvol als betekenissen van gegevens niet bekend zijn en/of gedeeld worden?

- > Er mag verschil zijn in betekenis
- > Dit moet duidelijk zijn
- > Overeenkomsten in betekenis benutten

Mogelijkheid: metapattern

- > Conceptueel modelleren van informatie
- > Begrippen en samenhang tussen begrippen
- > Hoofdrol voor context: de situatie waarin informatie wordt geïnterpreteerd bepaalt de betekenis van die informatie

Boeiende tijden!!