

## **Kennisdeling: zoeken en vinden – context en begrip**

### **Inhoudsopgave**

1. Inleiding
2. Zoeken en vinden
3. Oplossing

### **1. Inleiding**

Kennismanagement heeft alles te maken met kennis delen. En kennisdeling heeft weer alles te maken met het vastleggen en overdragen van kennis. Maar evenzeer met het zoeken en (kunnen) vinden van die vastgelegde kennis. Als er goede mogelijkheden zijn om te vinden (en dus ook om te zoeken) kunnen we pas echt aan kennisdeling toekomen. Gelukkig zijn er nu ook middelen beschikbaar om te helpen bij het zoeken en vinden, rekening houdend met context en begrip.

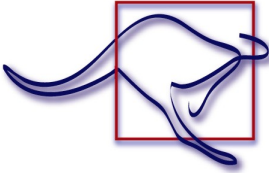
### **2. Zoeken en vinden – Context en begrip**

U zoekt informatie. Waarover? Waarom? Wat precies? U weet dat wat u zoekt te vinden is, maar niet hoe of waar. Wat doet u dan? U vraagt het aan iemand. Wat doet die persoon dan? Hij vertelt u wat u weten wilt. Of anders verwijst hij u naar iets of iemand anders.

Wat is het voordeel hiervan? Een mens begrijpt wat u bedoelt. U krijgt geen antwoord op uw letterlijke vraag, maar op wat u bedoelt. Ook wanneer u niet de juiste termen of formulering gebruikt. U krijgt alternatieven te horen. Dingen, termen, concepten waar u zelf niet aan had gedacht. Het antwoord dat u krijgt gaat dus niet alleen over de letterlijke vraag, maar ook over de vraag achter de vraag. Soms is daarvoor wat interactie nodig. Door er over te praten wordt een vraag pas echt duidelijk. Zo wordt het antwoord beter.

Computers kunnen dit niet zomaar. Ze zijn te letterlijk. Ze hebben geen idee dat een fiets en een rijwiel ongeveer hetzelfde zijn. Of dat je met fiest ook fiets bedoelt. Eén klein typefoutje en je krijgt een resultaat dat nergens op slaat. Computers kunnen over het algemeen weinig met "ongeveer".

Informatie zit in bronnen, informatiebronnen. In de normale wereld zijn dat mensen, websites, boeken, tijdschriften en/of bibliotheken. Als je iemand iets vraagt, kan het best zijn dat je als antwoord krijgt dat je een of ander boek even moet doorbladeren. Of een artikel in een bepaald tijdschrift moet lezen. Of met die-en-die moet praten. Bij een voldoende heldere vraag is vaak niet de hele bron relevant. Een klein stukje is vaak genoeg, een fragment. Van een tijdschrift. Of een hoofdstuk uit een boek.



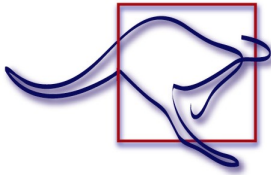
Een fragment bestaat weer uit begrippen of termen. Neem bijvoorbeeld de term *bedrijf*. Soms bestaan termen uit een verzameling andere termen. Zoals *ICT-bedrijf* of *ICT bedrijf*. Voor mensen is het geen probleem om te begrijpen dat *ICT bedrijf* een bepaald soort bedrijf is. Of dat *ICT-bedrijf* en *ICT bedrijf* hetzelfde zijn. Wanneer je iemand een vraag stelt waar de term "ICT bedrijf" inzit krijg je geen antwoord dat gaat over bedrijven als loodgieters of accountants. Bij computers heb je geluk wanneer het antwoord gaat over een ICT-bedrijf. Computers kunnen gewoon het verschil niet zien.

Om echt effectief informatie te kunnen vinden, zouden computers veel meer op mensen moeten lijken. Computers zouden dus moeten kunnen begrijpen dat je met "fiest", "fiets" en "rijwiel" dezelfde dingen bedoelt. Of wanneer je naar "ICT bedrijf" vraagt, dat je niet geïnteresseerd bent in "grootwinkelbedrijf". Of dat als je vraagt naar "winkel" je wel geïnteresseerd bent in "detailhandel", maar niet in "groothandel".

Om dat te kunnen, moeten computers, meer dan nu, begrip hebben van de context van informatie. Maar ook van synoniemen. En van homoniemen. Misschien ook van verschillende talen (bike, Fahrrad, bicyclette). Computers moeten dus iets begrijpen van de betekenis van woorden en teksten. Dat lukt nog niet echt, maar we komen langzamerhand wel een beetje in die buurt.

In de literatuur wordt regelmatig gerefereerd aan '*het semantische web*'. Dat begrip houdt in dat alle informatie die op het internet staat is voorzien van standaard betekenissen, dat vergelijkbare informatie in context kan worden gevonden. Om dat te kunnen moeten alle documenten die op het internet staan voorzien worden van metadata. Wanneer we met ons allen daar snel mee beginnen, zou het mogelijk zijn om over een paar jaar de vruchten daarvan te plukken. Echter, de standaardisatie is nog niet zo ver, dus kunnen we daar nog niet mee beginnen. Ook is het op korte termijn niet te verwachten dat dit automatisch kan. Voorlopig is het dus nog handwerk want helaas is het semantische web er nog niet.

Wat kan dan wel? Langzamerhand wordt softwaretechnologie ontwikkeld die iets begrijpt van tekst. Veel van die technologie gaat uit van interpretatie van zinnen. Op basis van de zinsstructuur, de syntactische opbouw, wordt een analyse van de betekenis gemaakt. Een probleem daarbij is vaak dat correcte zinsbouw een randvoorwaarde is voor de syntactische analyse en dat zonder correcte syntactische opbouw van een zin ook de betekenis (semantiek) niet bepaald kan worden. Dat is jammer want mensen spreken en schrijven vaak geen 'perfecte' (of correcte) zinnen. Jonge kinderen, twee of drie jaar oud, maken over het algemeen geen zinnen die syntactisch correct zijn. Toch begrijpen we ze.



Eigenlijk zoeken we dus naar een vorm van softwaretechnologie die wel kan begrijpen waar tekst over gaat, maar niet beperkt wordt door de eis dat de zinnen syntactisch correct zijn. Softwaretechnologie die in staat is om flexibel om te gaan met 'brokken betekenis'. Softwaretechnologie die misschien wel gebruik maakt van syntactische analyse, maar daar niet van afhankelijk is.

### **3. Een oplossing**

Als een dergelijke oplossing voor zoeken en vinden beschikbaar is zou het wel leuk zijn als die technologie ook nog betaalbaar is en zonder al te veel problemen met allerlei informatiebronnen geïntegreerd kan worden. Kan dat?

Het lijkt er op van wel... Er zijn al geruime tijd veel praktische en werkende oplossingen mogelijk. Welke voor uw organisatie interessant kan zijn is niet op voorhand te zeggen. Wanneer u daar in een korte tijdspanne echt iets concreets van wilt zien, en dan ook nog toegepast op uw eigen organisatie, is het mogelijk om om daar met iets als een Quick Scan achter te komen.