

TAGGING: EEN NIEUWE MANIER VAN ZOEKEN?

Anne Helmond



INLEIDING

Zestig jaar geleden beschreef Vannevar Bush het probleem van de toegankelijkheid van de toenemende hoeveelheid informatie. Hij bedacht hiervoor een apparaat, de memex: “Consider a future device for individual use, which is a sort of mechanized private file and library”¹. In de memex zou een ongekeerde hoeveelheid aan onder andere beeld- en tekstmateriaal opgeslagen kunnen worden. Een belangrijke eigenschap was de mogelijkheid van raadpleging, connotatie en uitbreiding.

Anno 2005 hebben we de mogelijkheid tot het opslaan van een gigantische hoeveelheid informatie op onze pc. Bovendien maakte het Internet mogelijk onze informatie ook buiten onze “mechanized private file and library” op te slaan en te delen met anderen.

Het probleem dat Bush aankaartte, hoe komen we bij de relevante informatie, is echter nog steeds relevant vraagstuk. Waar Bush voornamelijk spreekt over de wetenschapper als gebruiker van de memex, is de doelgroep van de pc en Internet vele malen groter. In tegenstelling tot de memex, waar we voornamelijk wetenschappelijke artikelen en beeldmateriaal vinden, wordt op het World Wide Web een enorme variëteit aan items gepubliceerd: van vakantie foto's tot dagboeken tot CDs.

De zoekmachine Google is tegenwoordig de meest gebruikte manier om informatie te zoeken op het World Wide Web.² Een bekend probleem is echter dat een zoektocht honderden, zo niet miljoenen resultaten kan opleveren, waarbij lang niet alle resultaten even relevant zijn. Tevens weet Google ook niet waar een webpagina daadwerkelijk over gaat. Hier ligt een mogelijkheid voor tagging.

Tagging is voor veel internetgebruikers nog een onbekend woord/onbekende toepassing, maar het wordt steeds vaker toegepast in allerlei online websites en applicaties zoals Amazon.com, Flickr en del.icio.us. Tonie van Ringelestijn geeft in het artikel 'De belofte van Web 2.0' de volgende definitie van tags:

TAGS: Trefwoorden die je aan berichten, foto's en links kan toevoegen (tagging), zodat je er content op kan organiseren of doorzoeken. Zoals een bibliotheek trefwoorden meegeeft aan boeken.³

Het idee van tagging is dat gebruikers zelf deze trefwoorden bepalen en koppelen aan een item om zo zelf structuur/zingeving te geven aan de informatie.

1 Bush, Vannevar. 'As We May Think' *Atlantic Monthly*, 176 (1) (juli 1945)

2 Sullivan, Danny 'Nielsen NetRatings Search Engine Ratings' <http://searchenginewatch.com/reports/article.php/2156451> 16 december 2005

3 Ringelestijn, Tonie 'De belofte van Web 2.0' *Bright* (december 2005)

De social bookmarking website del.icio.us is een populaire website om je relevante resultaten van een zoektocht op te slaan. Het is hierbij mogelijk om persoonlijke notities en tags toe te voegen.

REFERENTIEKADER

Vannevar Bush zag in de Memex ook een dergelijke rol voor de gebruiker, waarbij hij of zij “can add marginal notes and comments” aan een item.⁴ De kern van het systeem lag echter in de manier waarop de gerelateerde inhoud aan elkaar gelinkt was d.m.v. associatieve links en zo een *trail* vormde. “...he builds a trail of his interest through the maze of materials available to him.”⁵

Het zou dan dus mogelijk zijn om door een onderwerp te bladeren (ook wel *social navigation*⁶ genoemd), waarbij telkens relevante gerelateerde stukken worden gepresenteerd. Iets dergelijks zien we nu terug op de social bookmarking website del.icio.us. Elke gebruiker voegt zijn eigen tags toe aan zijn bookmarks. Zodra je als gebruiker bijvoorbeeld de website met een onderzoek naar [Feedback, Motivation and Collectivity in a Social Bookmarking System](#) opslaat, kun je deze de volgende tags meegeven “[bookmarking del.icio.us folksonomy research tagging tags delicious academicpaper pdf](#)”. Het interessante is dat direct getoond wordt welke andere gebruikers deze website ook hebben opgeslagen en welke tags zij hebben meegegeven aan deze website. Wellicht dat deze andere gebruikers ook interesse hebben in gerelateerde literatuur en deze al gevonden en opgeslagen hebben.

In tegenstelling tot de Memex dat een privé apparaat is, maar waarvan gerelateerde documenten wel uitgewisseld kunnen worden, is del.icio.us een sociaal en open systeem. Hierdoor stuit je bijvoorbeeld op nieuwe documenten die jezelf over het hoofd had gezien of simpelweg niet gevonden had, maar eventueel wel relevant of interessant zijn.

Beiden methoden (search engines en tagging) hebben hun voor- en nadelen en het is interessant om te bekijken wat een combinatie van de mogelijkheden van beiden voor ons zou kunnen betekenen voor de manier waarop we met het organiseren en zoeken van onze informatie omgaan. Hieruit volgt dan ook de volgende vraag: **Wat is de mogelijke impact van tagging op onze manier van zoeken naar informatie?**

4 Bush, Vannevar. 'As We May Think' *Atlantic Monthly*, 176 (1) (juli 1945)

5 Bush, Vannevar. 'As We May Think' *Atlantic Monthly*, 176 (1) (juli 1945)

6 Dourish and Chalmers [1] define social navigation as provoking a user to move through an information space guided by the activity of others in that space: uit Jarett, A.C. en B.M. Dennis 'NusEye: Designing for Social Navigation in Syndicated Content'

DEELVRAGEN

- Wat zijn *social software* tools?
- Wat verstaan we onder het begrip folksonomy? Hoe verhoudt dit zich tot de taxonomy/classificatie die gebruikt wordt door professionals, bijvoorbeeld bibliothecarissen?
- Moet er toch niet een vorm van hiërarchie worden ingevoerd? Als je bijvoorbeeld zoekt op de tag “Bush” krijg je resultaten van personen, flora en nog veel meer. Of bijvoorbeeld een stem systeem, waarbij gebruikers kunnen stemmen op relevante resultaten.
- Heeft Yahoo een public statement gemaakt met wat zij met Flickr en del.icio.us willen? Dat zijn toekomstplannen van Yahoo met deze twee sociale software sites?
- Wat is het Semantic Web en Web 2.0? Wat zijn de doelstellingen en welke verbeteringen voor het zoeken van informatie stellen zij voor.

BEPERKINGEN

Niet te technologisch, meer filosofisch. Tagging wordt nu vooral gedaan door “geeks”, maar zal de massa het ook gaan doen? Slechts zestien items op de populaire site Amazon.com dragen nu de tag “web” in tegenstelling tot 257003 op de “geek” site del.icio.us. Daar valt tegenin te brengen dat Amazon.com pas enkele maandende mogelijkheid biedt om tags te gebruiken in tegenstelling tot del.icio.us dat al sinds 2003 actief is. En zijn op dit moment dus veel meer actieve taggers te vinden op del.icio.us dan Amazon.com.

Daarentegen is Amazon.com een veel drukbezochtere website met een groter publiek en het is dus daarom interessant om te kijken hoe het taggen zich daar zal gaan ontwikkelen.

DISCUSSIE VAN DE LITERATUUR

Belangrijkste uitgangspunt is het idee van de memex van Vannevar Bush en zijn idee van het indexeren van informatie d.m.v. associatie. Ook tagging op websites als del.icio.us maakt gebruik van deze associatieve indexering:

Tagging allows for the kind of multiple, overlapping associations that the brain itself uses, rather than rigid categories. In the canonical example, a Flickr photo of a puppy might be tagged both "puppy" and "cute"--allowing for retrieval along natural axes generated user activity.⁷

7 O'Reilly, Tim. 'What Is Web 2.0: Design Patterns and Business Models for the Next Generation of Software' <http://www.oreillynet.com/lpt/a/6228> 17 december 2005

Deze associatieve indexering leidt tot het gebruik van een eigen manier van indexering, niet langer door de auteur of distributeur, maar door de gebruiker van de informatie:

Flickr infers its knowledge from the tags entered by every other user in the system—creating a so-called “folksonomy,” a group intelligence derived by association.

De correctheid van de term *folksonomy* is omstreven, maar wordt gebruikt om aan te duiden dat het gaat om een taxonomy door 'het volk' in tegenstelling tot de taxonomy van de professionals.⁸

"We are joining forces to build my vision of creating a way for people to remember things together," Schachter told Reuters in a phone interview. "It is a shared-memory site."

Dit is een fragment uit de overname statement van de oprichter van del.icio.us, Schachter. Hij spreekt hier over een collectief, gedeeld geheugen online. Dit in tegenstelling tot Bush' Memex wat in principe een privé apparaat was, waarvan wel de trails uitgewisseld konden worden, maar niet altijd voor iedereen toegankelijk waren.

“Consider a future device for individual use, which is a sort of mechanized private file and library”⁹

Dit idee van een collectieve, gedeelde plek voor informatie is wat tagging potentieel zo krachtig maakt. Adam Weiss spreekt dan ook over hoe gefragmenteerde delen tezamen een meerwaarde kunnen brengen:

Both through collected photos and community-defined tags, Flickr demonstrates how aggregate sharing of fragments can create value, ultimately becoming more useful than the sum of its parts.

Dr. Licklider, hoofd van het Advanced Research Project Agency, beschreef in 1968 al hoe mensen met elkaar in contact zouden moeten worden gebracht om tezamen hun ideeën te delen, waardoor er een intellectuele meerwaarde zou ontstaan. Hij beschrijft vervolgens de werking van de groep psychologie en hoe *collaborative communication systems* (die op dat moment nog niet bestonden) de samenwerking zouden kunnen verbeteren.¹⁰

“...One must consider the dynamics of ‘critical mass,’ as it applies to cooperation in creative endeavor. Take any problem worthy of the name, and you find only a few people who can contribute effectively to its solution. Those people must be brought into close intellectual partnership so that their ideas can come into contact with one another.”

Een dergelijke groepsdynamica zien we nu ook aan de gang op sites zoals Flickr en del.icio.us. Gebruikers delen hier de resultaten van een zoektocht en kunnen andere

8 Mathes, Adam. 'Folksonomies: Cooperative Classification and Communication Through Shared Metadata' (december 2004), <http://www.adammathes.com/academic/computer-mediated-communication/folksonomies.pdf>, 9 december 2005.

9 Bush, Vannevar. 'As We May Think' *Atlantic Monthly*, 176 (1) (juli 1945)

10 Weiss, Aaron. 'The Power of Collective Intelligence' *Collective Intelligence* (september 2005): p.18

gebruikers wijzen op potentiële andere interessante websites. Tevens kunnen gebruikers door de resultaten van andere gebruikers browsen.

Een probleem waar echter tegenaan gelopen wordt is dat het opslaan van websites en het taggen daarvan op del.icio.us op dit moment voornamelijk door de zogenaamde *geeks* gedaan wordt. De vraag hierbij ontstaat is of het ook geschikt is voor de massa of ook geschikt voor de massa.

Yahoo ziet blijkbaar potentie in tags voor de massa, want het kocht de twee populairste websites die gebruikmaken van taggen op: del.icio.us en Flickr. Dit is een zeer interessante overname, want ongeveer een kwart van de gebruikers van zoekmachine maakt gebruik van de zoekmachine van Yahoo.¹¹ Yahoo heeft dus nu zowel de techniek van een redelijk goed functionerende zoekmachine in huis evenals de kracht van de collectieve intelligentie van de gebruikers ervan.

METHODE

Om tot een antwoord te komen op de onderzoeksvraag zal ik gebruik maken van een literatuurstudie in combinatie met enkele case studies. De volgende websites komen hiervoor in aanmerking:

1. Yahoo's Myweb2.0: alhoewel deze website zich nog in een testfase bevindt, vinden we hier wellicht tot nu toe de ultieme bundeling van zoeken en taggen. Met name nu Yahoo ook over de techniek van del.icio.us en Flickr beschikt. del.icio.us: de website die het social bookmarking op de kaart heeft gezet. Gebruikers van de website slaan voor hun belangrijke website op, becommentariëren en organiseren deze. (Wellicht ook interessant: Connotea: deze website lijkt op del.icio.us, maar is bedoeld voor wetenschappelijk onderzoek). Flickr: deze website heeft een andere inslag dan de voorgaande, omdat hier eigen materiaal getagged en georganiseerd wordt.
2. Google Persoonlijk: op diverse weblogs wordt in oktober 2005 vermeld dat ook Google Persoonlijk ook mogelijk maakt om bookmarks uit de zoekgeschiedenis op te slaan en te taggen. In december is deze mogelijkheid echter niet meer terug te vinden, hetgeen interessant is. Hoe ziet Google de toekomst van zijn zoekmachine? Zien zij niet langer heil in tags?

¹¹ Sullivan, Danny 'Nielsen NetRatings Search Engine Ratings'
<http://searchenginewatch.com/reports/article.php/2156451> 16 december 2005.

HYPOTHESE

Het is lastig om te voorspellen of tagging zal gaan leiden of bijdragen tot/aan betere zoekresultaten. Wat in ieder geval wel zichtbaar is, is dat erop websites als del.icio.us op een andere manier met informatie wordt omgegaan. Het lijkt erop dat op dergelijke websites weer meer browsen in plaats van zoeken met queries. Met het idee van collectieve intelligentie in het achterhoofd lijkt het dat de combinatie van de kracht van goede zoekmachine en de kracht van de collectieve, menselijke associatieve betekenisgeving zou kunnen leiden tot betere resultaten.

WAAROM

Op 20 maart van dit jaar kocht Yahoo Flickr en op 9 december kocht Yahoo del.icio.us, waardoor Yahoo nu twee van de meest vooraanstaande social software tools van dit moment in handen heeft. Wat betekent dit voor de toekomst? Zien we een nieuwe generatie zoekmachines tegemoet die het taggen zal integreren? Zal het taggen de betaversie overleven?

BIBLIOGRAFIE

- Bush, Vannevar. 'As We May Think' *Atlantic Monthly*, 176 (1) (juli 1945)
- Jarett, A.C. en B.M. Dennis 'NusEye: Designing for Social Navigation in Syndicated Content' *Proceedings of the 2005 conference on Diversity in computing TAPLA '05* (oktober 2005)
- Mathes, Adam. 'Folksonomies: Cooperative Classification and Communication Through Shared Metadata' (december 2004), <http://www.adammathes.com/academic/computer-mediated-communication/folksonomies.pdf>, 9 december 2005.
- O'Reilly, Tim. 'What Is Web 2.0: Design Patterns and Business Models for the Next Generation of Software' <http://www.oreillynet.com/lpt/a/6228> 17 december 2005
- Proceedings of the sixteenth ACM conference on Hypertext and hypermedia 2005, <http://portal.acm.org/toc.cfm?id=1083356&type=proceeding>, 17 december 2005.
- Rouch, Wade 'Tagging is it: Social bookmarking emerges as a wiser way to navigate the web' *Technology Review* (juni 2005)
- Sullivan, Danny 'Nielsen NetRatings Search Engine Ratings' <http://searchenginewatch.com/reports/article.php/2156451> 16 december 2005.
- Udell, Jon. 'Group Knowledge Gardening' *Infoworld.com* (augustus 2004)
- Walker, Gill 'Feral Hypertext: When Hypertext Literature Escapes Control'

Weiss, Aaron. 'The Power of Collective Intelligence' *Collective Intelligence* (september 2005)
Schraefel, M.C., ea. 'The Evolving mSpace Platform: Leveraging the Semantic Web on the
Trail of the Memex' *Proceedings of the sixteenth ACM conference on Hypertext and hypermedia*
(september 2005)

<http://del.icio.us/silvertje>