Tema: Big data

När årets vinskörd är klar i Bordeaux står alla vinkännare och luktar och smakar på vinet. De har tränat sina näsor i åratal, vissa äter inte kryddor för att förbättra sina sinnen. När de smuttat på vinet börjar de fundera på om det är en lyckad årgång eller ej. Men människor är inte de enda som kan bedöma kvaliteten på ett vin – det kan även datorer. Genom att analysera historiska mätserier för nederbörd och temperatur och korrelera dem mot priserna för årgångsviner så kan man finna tydliga samband. Vissa väderkombinationer visar sig statistiskt sett vara mer gynnsamma för att vinet ska bli bra och därmed betinga ett högre pris. I själva verket håller systemet så hög klass att dess prognoser för framtida prisnivåer år efter år har högre träffsäkerhet än de framstående vinexperterna.

Mamma Mia, datorer börjar tala vin.

Exemplet må vara trivialt men allt fler organisationer analyserar enorma databaser för att finna oupptäckta mönster och komplexa samband. Statistisk analys i sig är inget nytt, men snabbare datorer, större datamängder och komplexare algoritmer gör att man kan vaska fram raffinerade insikter i en helt annan omfattning än tidigare. Informationsfiltrering och statistiska modeller kan användas inom allt från att marknadsföra på intelligentare sätt till att mäta hur politiska åtgärder faller ut. Det kan användas för skapa bättre filmer eller för att bedöma ifall en viss kund är lämplig att få ett banklån. Man kan studera om klassisk musik i butiken ökar försäljningen eller om någon försöker använda stulna kreditkort.

De senaste månaderna har Futurewise genomfört fem workshops på temat big data. Vi har således samlat på oss en del material och även reaktioner från omkring 150 chefer och deras inställning till fenomenet. Detta nyhetsbrev sammanfattar en del av våra tankar.

Det största sedan internet

Konsultbolaget McKinsey & Co. hävdar att big data är den största förändringen för försäljning och marknadsföring sedan internet slog igenom för drygt 20 år sedan. Det är ett rätt kraftfullt uttalande. De menar att vi just nu står vid en brytpunkt där organisationer som lär sig dra nytta av big data kommer att överträffa dem som inte gör det. Studier visar att de som använder big data på ett effektivt sätt har vinstmarginaler som är 5-6 procent högre än de som inte gör det. Visionären Tim O'Reilly menar att vi går en ny värld till mötes där data kan bli viktigare än mjukvara.

Gravida unga kvinnor och upprörda fäder

För att förstå hur avancerade dagens metoder är kan vi ta ett exempel från detaljhandeln. Amerikanska butikskedjan Target analyserar precis allt som deras kunder köper. Via deras kundkort registreras alla köp som en viss kund gör, vilket skapar underlag för anpassade erbjudanden. Detta kompletteras med andra demografiska data som de samlar in eller köper externt, som ålder, kön och bostadsort.

Deras raffinerade dataalgoritmer låter dem finna de mest intrikata samband. Deras analytiker insåg att kvinnor som var gravida gick över till att köpa parfymfri hudkräm. Dessutom började de köpa kosttillskott som kalcium, zink och magnesium. Efter fördjupade studier vaskade analytikerna fram en lista på 25 produkter som gravida tenderade att köpa. Utifrån denna kunskap kunde de inte bara få fram sannolikheten för att en viss kvinna var gravid, utan till och med göra en ganska god uppskattning av vilken månad hon var i.

Target vet om en kvinna är gravid, snudd på före henne själv.

Hur välfungerande systemet är kan följande anekdot illustrera. En mycket upprörd man gick in i Targets butik i Minneapolis och krävde att få tala med chefen. ”Min dotter fick den här reklamen via mejl där ni ger henne rabattkuponger på babykläder och blöjor. Hon går fortfarande på gymnasiet. Försöker ni uppmuntra henne till att bli gravid?” sa han argt. Chefen hade ingen aning om vad mannen pratade om, men bad om ursäkt å Targets vägnar och lovade att återkomma via telefon. När chefen ett par dagar senare ringde upp mannen var denne mycket förlägen. Han hade talat med sin dotter och det visade sig att hon i själva verket *var* gravid. Den blivande morfadern var nu den som fick be butikschefen om ursäkt.

För näringslivet – och för det offentliga

Att big data kommer att bli stort inom försäljning och marknadsföring är uppenbart. Vd:n på PR- och reklambyrån WPP med 162 000 medarbetare, menar att hälften av deras anställda håller på med något vetenskapligt eller saker som relaterar till vetenskap. Det är lätt att tro att en reklambyrå har ett gäng kreativa personer som med sin fantastiska intuition kan sätta nästa smarta kampanj, men världen förändras. Reklam håller på att bli en vetenskap.

Polisman i Los Angeles letar bovar innan de begått brott.

Det är nu inte bara till reklam och försäljning big data kan användas. Sannolikt kommer det bli väldigt intressant för det offentliga att dra nytta av big data för att förbättra verksamheter och servicen till medborgarna. Inom polisväsendet använder man sig av big data för att förutspå brott. Det kanske inte fungerar lika bra som i sci-fi-filmen Minority Report där man ser exakt när och av vem ett brott ska begås, men systemen är ändå rätt kvalificerade. Genom att studera historiken och se var brott tidigare begåtts, på vilka veckodagar, vid vilka tidpunkter på dagen, vid vilken väderlek, etc. så kan man skapa en ganska god uppfattning var brotten – statistiskt sett – härnäst kommer att ske. När en zon börjar bli röd på översiktskartan så skickar man dit en polisbil som skrämmer bort buset och det lugnar ner sig en aning. Tillräckligt för att brottsligheten påtagligt ska minska.

Det är naturligtvis inte svårt att komma på hundratals andra spännande områden där det offentliga kan dra nytta av big data. Sannolikt kommer vi att få se mycket mer av detta i framtiden även inom offentlig sektor.

Ingen bransch undgår big data

Även inom HR kommer system som dessa att få ett stort genomslag. Om man kan ana att någon är gravid borde man väl också kunna ana om någon är på väg att säga upp sig? På Hewlett-Packard tar man datakraft till hjälp för att försöka förutspå medarbetarnas beteenden. Genom att tilldela alla deras 330 000 anställda ett så kallat ”flight risk score” kan de bilda sig en uppfattning om vilka som är på väg att lämna företaget. Denna siffra grundar sig på prediktiv analys av faktorer som lön, hur länge man har varit anställd, vilken position man har, avstånd till arbetet, etc. Genom att de högre cheferna har fått tillgång till analyserna har de snabbt kunnat erbjuda högre lön eller en bättre tjänst till de personer de har velat ha kvar. I en inledande pilotgrupp har man minskat personalomsättningen med 25 procent.

Vi synas ständigt i sömmarna, av företag som vill sälja saker och också av våra arbetsgivare.

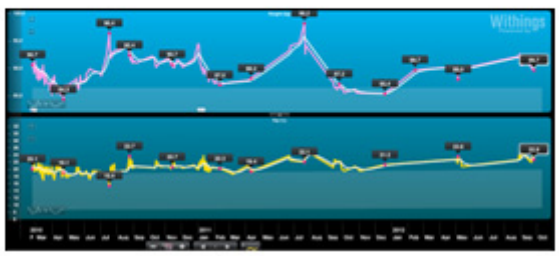
System som dessa har till och med kunnat indikera hur man kan motivera anställda genom att studera vilka kvaliteter som är viktiga för olika jobbpositioner. Man kan också mäta de anställdas motivation, hur väl nyrekryterade motsvarar förväntningarna eller vilka som är lämpliga att befordra. Elektronikåterförsäljaren Best Buy kan ganska exakt uppskatta värdet av en ökning i medarbetarnas engagemang. Varje ökning med 0,1 procent i personalens engagemang motsvarar en intäktsökning om 600 000 kronor för en viss butik. Datauppgifterna kan också användas för att förutse sjukskrivningar, som – föga förvånande – har visat sig sammanfalla med en viss typ av väderlek eller vissa idrottsevenemang.

En viktig insikt är att denna typ av kompetens inte självklart är något som personalavdelningar har idag. Det kan komma att krävas nya resurser som har den analytiska förmågan att tolka data, och chefer som korrekt kan agera på resultaten. I USA använder omkring fem procent av de största företagens personalavdelningar någon form av prediktiv analys av de anställdas beteenden. Det är förstås lätt att se en integritetsproblematik i detta, där vissa orättfärdigt kan bedömas vara illojala statistiskt sett. Att syftet är att arbetsgivaren helst vill ha dem kvar, och till och med ger högre lön eller en bättre position till dem som är på väg att sluta, är förstås en förmildrande omständighet. Även utan dataanalys gör cheferna troligen kontinuerligt en intuitiv bedömning av de anställdas kvaliteter.

Trestegsraketen

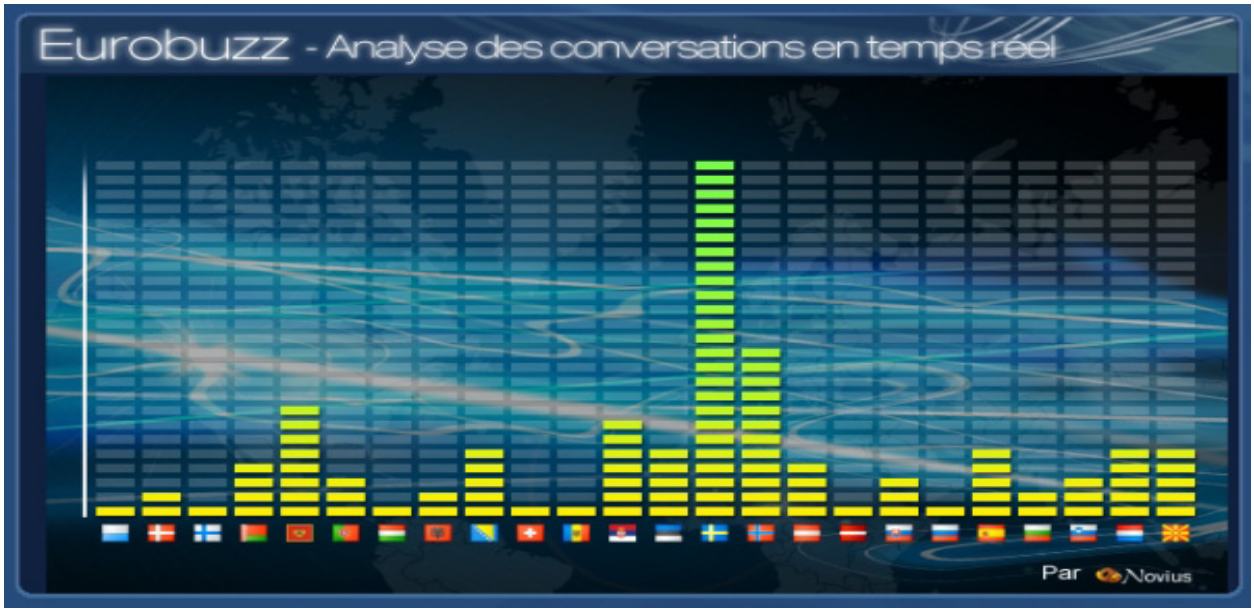
Om vi ser till själva uttrycket ”big data”, kan vi konstatera att det inte finns något motsvarande vedertaget svensk begrepp (det närmaste man kommer är att säga big daaata, istället för big dejta). Big data säger rätt tydligt vad det handlar om: analys av stora datamängder. Det är i princip en trestegsraket: 1) samla in data 2) analysera och utforska data och 3) agera på informationen.

Om vi studerar den första punkten så genereras det en enorm mängd data. Varenda köptransaktion, telefonsamtal, webbsidebesök, googling, facebookuppdatering, annonsklick, e-postmeddelande, Twitter-inlägg, sms, etc. genererar mängder med ny data. IBM påstår att 90 procent av all data har tillkommit de senaste två åren. Tidigare har organisationer också haft den del data, huvudsaklig väl strukturerad. I ekonomisystem, HR-system eller CRM-system har man lagrat information om anställda, kunder, försäljning med mera. Detta kompletteras nu alltså med en hel räcka andra källor som input, många av dessa är *ostrukturerade*. Med detta menas att det inte alltid är trivialt att utläsa meningen med vad som står. Materialet är inte alltid etiketterat och semantiskt beskrivet. Texten i ett Twittermeddelande är ett exempel på ostrukturerad information. Vad vi har sett på senare tid är framväxten av särskilda ”data brokers”. Företag som Axciom, Datalogix, Epsilon med flera har som affärsidé att sälja rådata till sina kunder. Datalogix uppger exempelvis att de har transaktionsdata som avser närmare 1 000 miljarder dollar i konsumentköp i USA. Lägg till detta det snabbt växande antalet sensorer i världen, IBM uppskattar antalet till 30 miljarder. Vi blir också rätt duktiga på att samla in data med appar, aktivitetsarmband, twittrande vågar, med mera.



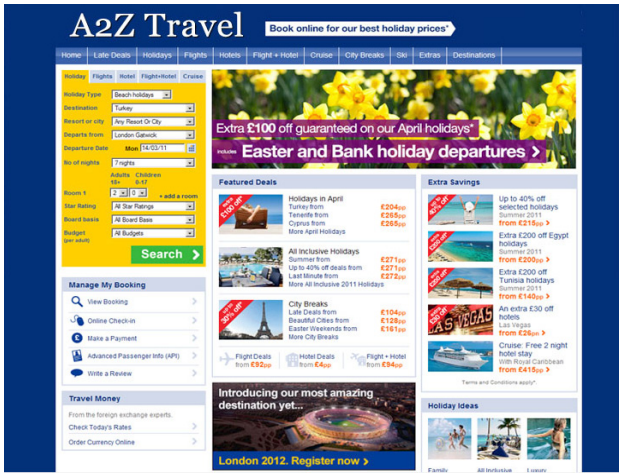
Aktivitetsarmband, twittrande vågar och en sensor som meddelar oss när det är dags att vattna.

Nästa steg är att analysera de stora datamängderna och här finns en uppsjö aktörer. Allt från grundtekniker som databaser för lagring till avancerade system för grafisk visualisering av komplex information. Vi har med dessa system själva analyserat Futurewise försäljningsdata (även om detta knappast kan kallas big data), och nått nya insikter i vilka kunder och affärer som är mest lönsamma. Utifrån detta har vi också förändrat förpackningen av våra trend-föredrag och också anpassat marknadsföringen av dem. Även den ostrukturerade datan går förstås att analysera genom så kallad semantisk analys. Exempelvis analyserade man hur snacket gick inför Eurovision Song Contest härom året och kunde tre dygn innan finalen säga att Loreen skulle vinna.



Att Loreen skulle vinna var uppenbart flera dagar före finalen.

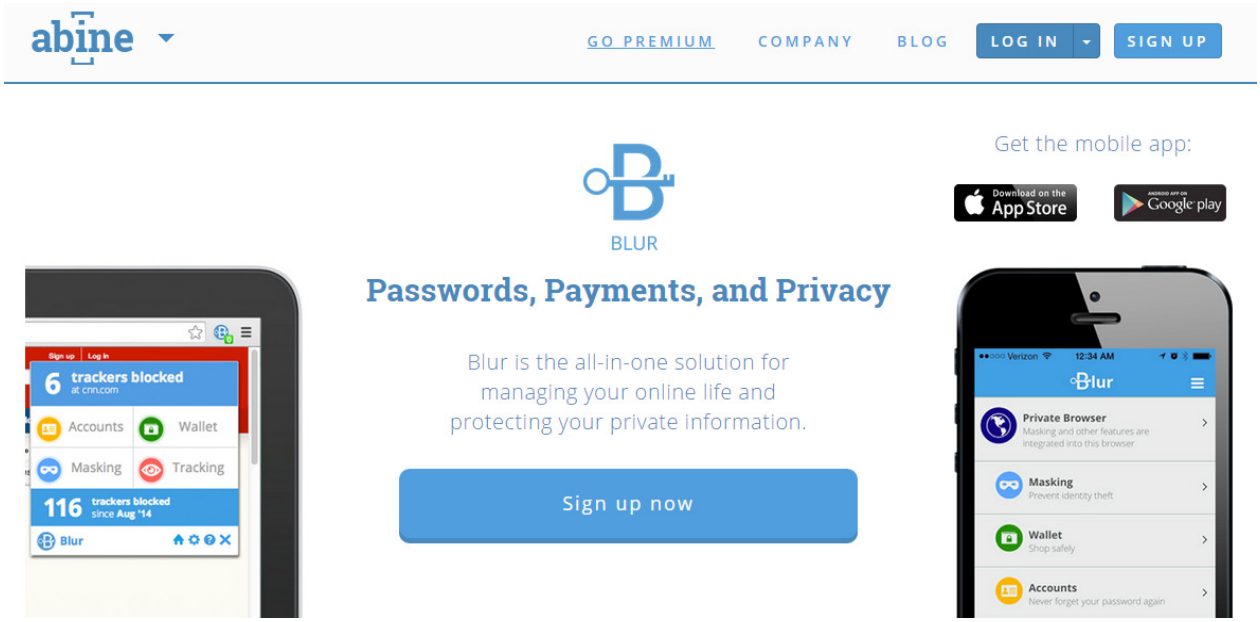
Slutligen måste man agera på informationen. Det är ju en form av beslutsstödssystem det handlar om. Det intressanta är att man kan göra detta mycket snabb, ja till och med i realtid. Medan någon är inne på en webbsida kan innehållet förändras beroende på var klicken sker. Innehållet kan till och med ändras beroende på hur muspilen rörs!



När det är uppenbart att en besökare är på jakt efter skidsemester visas bara sådan reklam.

Det är alltid gratis ost i råttfällan

Det går förstås att framhålla kritik till överdrivet siffertroende. Till att börja med är resultaten helt beroende av kvaliteten i indatan. Skit in, ger som vanligt skit ut. När det är medborgares och konsumenters beteende som analyseras så blir även integritetsaspekten av vikt. Det kan förstås upplevas lite obehagligt med företag och myndigheter som i detalj vet allt om ens preferenser här i livet. Just denna aspekt har framförts som en stor problematik av de 150 cheferna vi mött. I USA pratar man om ”the freaky line”. Om man kliver över den har man gjort ett övertramp och tagit sig lite för långt in i folks privata sfär. Target är ett bra exempel på detta. Vad de gjorde för att det inte skulle vara så uppenbart att visste om att någon är gravid, var att kombinera reklam för blöjor och välling med annat som ketchup, vin.



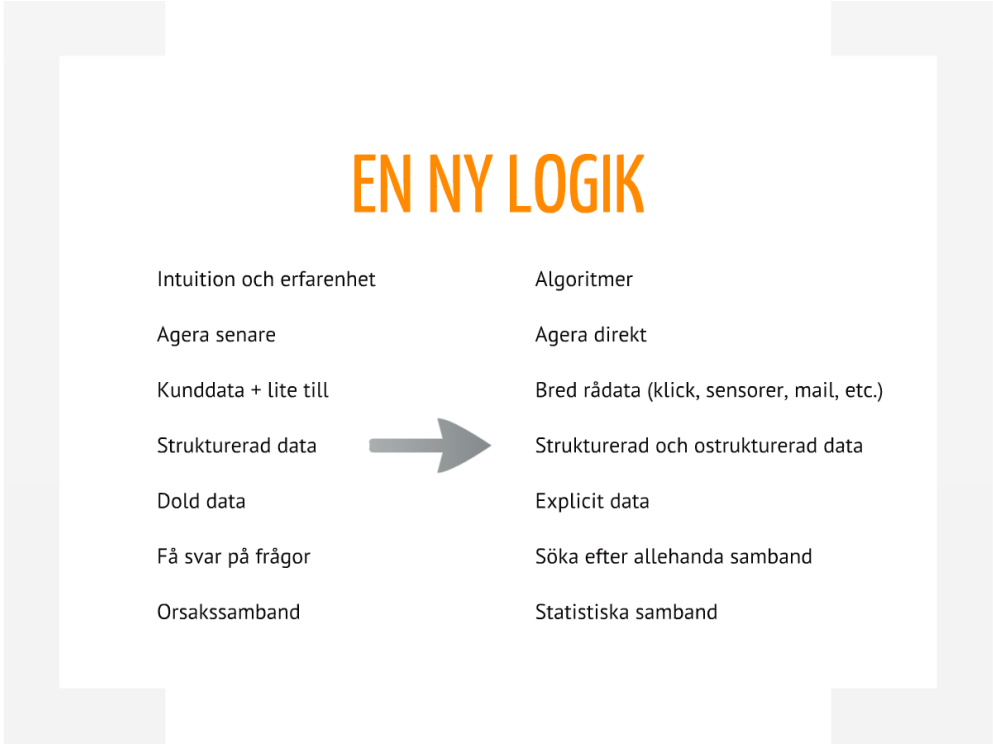
Blur, en av alla dessa mjukvaror som ska skydda vår personliga integritet.

En svensk man som företrädde nykterhetsföreningen IOGT-NTO skrev fler inlägg på Facebook om alkoholrelaterade frågor. Vad han inte vara beredd på var att det på hans Facebook-sida började dyka upp reklam om vin! Systemen är inte perfekta och många upplever det besvärande med personlig reklam. Men, som man säger, det är alltid gratis ost i råttfällan. Att använda Gmail, Facebook, m.m. är helt frivilligt. Det är gratis ur ett monetärt perspektiv, men vi betalar med vår uppmärksamhet. Inställningen till integritet är högst personlig, men sannolikt nöts många ner och resignerar över tiden. ”You have zero privacy anyway? Get over it!” som Scott McNealy sa.

Vinreklam till nykterister pekar på att systemen inte är perfekta.

En ny logik

Vi kan sammantaget se en ny logik växa fram. Algoritmiseringen av världen fortsätter på bekostnad av intuition och erfarenhet. Samtidigt måste förstås analyserna tolkas med sunt förnuft. Big data ger också möjlighet att agera direkt istället för att lägga in lite siffror i ett Excel-ark som plockas fram på varje veckomöte. Traditionell kunddata kompletteras med bredare rådata, som ofta är ostrukturerad. Vad vi ser hända är att dold data håller på att bli explicit. Detta är egentligen den övergripande och helt avgörande trenden. Det kommer inte att finnas så förfärligt många hemligheter i framtiden. Big data handlar inte bara om beslutsstödssystem som ger svar på frågor, utan det handlar lika mycket om att explorativt själv gräva i datan för att finna intressanta mönster och affärsmöjligheter. Orsakssamban-den blir också ointressanta så länge vi hittar de statistiska sambanden. Det spelar ingen roll varför något sker, vi är bara intresserade av att veta att det gör det när en annan variabel förändras. Kausalitet ersätts av korrelation, vilket något hårddraget leder till teoriernas död.



Detta nyhetsbrev är författat av Peter Siljerud som är trendspanare på Futurewise. Peter är en uppskattad föredragshållare som har inspirerat mer än 25 000 personer. Under våren 2015 kommer hans nya bok *Bryt normerna - Allt du lärt dig om arbetslivet är fel*.

Intresserad av inspirerande trendföredrag?
Maila på **info@futurewise.se**

För mer info och prenumeration på detta nyhetsbrev,
kolla in **www.futurewise.se**

FUTUREWISE