

Veliki podatki

Študija primera Target

Žiga Drašček
17. marec 2020

Pred leti je veliko razburjena povzročil članek Duhigga: "Kako podjetja izvedo za vaše skrivnosti". V članku je opisano, kako naj bi podjetje povzročilo razburjenje, ker je pred njim izvedelo, da je njegova hčerka noseča, vendar veliko ljudi dvomi v resničnost te anekdote, kar lahko preberemo v članku, ki ga je napisal Piatetsky: "Je Target res predvidel najstniško nosečnost? Interna zgodba ", saj je sumljivo to, da bi le njegova hčerka dobila katalog, lahko bi šlo enostavno le za promocijo nosečniških izdelkov.

Target podatke o svojih strankah pridobi iz večih različnih virov in sicer stranki pripiše edinstveno številko stranke (ang. Guest ID number), ki jo pripišejo stranki, ko uporabi kupon, izpolni anketo, zahteva povračilo, pokliče pomoč strankam, odpre e-sporočilo ali pa obišče spletno stran. Ker je Target veliko podjetje, lahko sklepamo, da stranki pripišejo številko zelo pogosto, saj se s podjetjem opravi veliko transakcij, kjer se uporabi kupon, prav tako verjetno pošljejo veliko e-sporočil, kar zadošča, da je izpolnjen eden od pogojev za big data in sicer pogostost generiranja podatkov (ang. velocity), prav tako pa bi to lahko zadoščalo za enega izmed ostalih pogojev in sicer količina podatkov (ang. volume), čeprav je ta pogoj nekoliko nejasen, saj se velikost podatkov, ki so potrebni za izpolnitev tega pogoja z leti večja, pri čemer različni avtorji navajajo različne začetne velikosti in različne hitrosti povečevanja tega pogoja.

Iz podatkov, ki jih zberejo pridobijo veliko različnih podatkov o osebi, kot so na primer spol, etnografski podatki, ali si starš, ali si poročen, približno plačo, katere revije bereš in podobno. Že sam dostop do teh informacij je lahko sporen, saj so morda stranke površno prebrale pogoje poslovanja, so si premislile in podobno, drži pa, da bi se lahko temu izognile v primeru, da bi poslovale izključno z gotovino in se odpovedale vsem ugodnostim (niti ne bi smele reklamirati blaga). Še več razburjenja pa je povzročilo to, da je statistik in ekonomist Pole, ki je zaposlen v Targetu razvil model, ki napoveduje nosečnost (z zelo natančnim rokom). Cilj tega modela je ujeti obdobje, ko se ljudem najbolj spremenijo navade. Ljudje (in živali) zelo težko spreminjamo navade, saj obstaja nek sprožilec, ki sproži avtomatizirano obnašanje, ki smo ga razvili skozi učenje, na koncu pa nas čaka neka nagrada. Primer tega je lahko sladica (sprožilec), ki jo pojemo, ko jo vidimo ter nam dvigne krvni sladkor in se zato bolje počutimo (nagrada), podobno je pri kajenju, kjer je

lahku kadilcu dolgčas, si želi družbo ali gre kaditi njegova družba (različni sprožilci), kadilec si prižge cigareto in kadi (avtomatizirano vedenja), na koncu pa se ponavadi počuti bolje, saj je zadovoljil svojo potrebo po nikotinu. Psihologi so dokazali, da se lahko pa se tovrstne navade spremeni, kar skušajo spodbuditi Target in njegovi konkurenti. Ljudje največ nagrad spremenijo ob prelomnih dogodkih v življenju, kot so prva zaposlitev, poroka, ločitev, veliko povišanje ali znižanje plače, verjetno najbolj pomemben dogodek pa je rojstvo otroka, ki pa je ponavadi objavljeno javno, zato takrat mater dobiva veliko ponudb različnih trgovcev, Target pa jih je želel prehiteti in navade vzpostaviti že v času nosečnosti. To je kontroveržno, saj veliko ljudem ni všeč, da bi z velikim podjetjem delili tako intimne podrobnosti svojega življenja, tega pa se je zavedel tudi Target, zato poskušajo tovrstno oglaševanje prikriti, saj bi se stranke zaradi tega lahko odločile za konkurenco in so sedaj tovrstne oglase postavljajo zraven oglasov za artikle, ki jih noseče ženske ponavadi ne kupujejo, kot so na primer kosilnice.

Sedaj je očitno tudi, da je izpolnjen še zadnji pogoj za to, da so podatki big data in sicer raznolikost podatkov (ang. variety). Poleg že prej omenjenih podatkov Target kombinira te podatke tudi z odkritji znansvenih raziskav, ki so bile omenjene v prejšnem odstavku (pidoobivanje in spreminjane navad).

Menim, da so podatki, ki jih od strank pridobi Target točni, saj so primarni in izhajajo neposredno iz transakcij, čeprav zagotovo v modelu prihaja do določenih napak, kot je na primer ne odkritje določenih nosečnic, menjava nenosečih oseb, ki pa zaradi slučaja kupujejo podobne artikle kot nosečnice v podobnem časovnem obdobju. Nekoliko pa dvomim v to, da se je zgodba s hčerko res zgodila, niti ne vem, zakaj bi se želel kdo postavljati s to zgodbo, saj je precej jasno, da bo odziv javnosti negativen.

Branje člankov me je spomnilo (AHA-moment), da je za ljudi, ki jim je zasebnost pomembna, najboljši način, da jo ohranijo vsaj malo ta, da kljub prednostim brezgotovinskega plačevanje za kupce in prodajalce (sledljivost transakcij, enostavnejše zaključevanje blagajne, ni napak pri štetju denarja) vse transakcije opravijo z gotovino, pa tudi to ne zadošča več, saj lahko podjetja sledijo tudi preko e-maila, v določenih državah pa je večji del velikih

mest (primeri so London, Beograd in kitajska mesta) pod kamero, oblasti tam pa tudi niso omejene z strogo zakonodajo o varovanju podatkov.

Drugi AHA-moment je bil to, kako oblikujemo navade, saj sem opazil, da je tako dejansko sestavljena večina našega življenja, upam tudi, da mi bo zavestno spreminjanje navad pomagalo zaobiti stare slabe navade in vzpostaviti boljše nove. Članek v New York Timesu me je spomnil, tudi na to, da je zelo pomembno tudi to, da se za določena stvar ujame primeren trenutek, v primeru, ki smo ga analizirali je to oglaševanje artiklov, velja pa to tudi za zaposlitev, potovanja, saj na primer po rojstvu otroka ne bi želeli na novo nastopiti v zahtevni službi, ampak se posvetili družini, podobno pa bi lahko prenesli tudi na druge dogodke v našem življenju.