

“Toon me uw data en ik vertel u of u succesvol bent”

Internationale data vergaren op je eigen website; zet een contactformulier online en klaar is kees. Of is het niet zo simpel als het lijkt?

Heerlijk, dat internet. Nog nooit was het zo makkelijk om internationaal prospects en nieuwe klanten binnen te halen. Gewoon een eenvoudig formuliertje op je website zetten waar mensen hun contactgegevens kunnen invullen, en de database vult zich bijna als vanzelf. Zou je denken.

Ga willekeurig wat sites af, op zoek naar de registratieformulieren, en je ziet al snel hoe weinig aandacht bedrijven geven aan correcte invoer van data. Gevolg: het aloude garbage in, garbage out.

Onbegrijpelijk dat organisaties zichzelf in de voet schieten, vindt Graham Rhind, van GRC Database Information. Rhind staat bekend als de expert op het gebied van internationale datakwaliteit. De Engelsman woont al geruime tijd in Nederland. Hij heeft drie boeken over databases en datakwaliteit op zijn naam staan. Over het belang van datakwaliteit zegt Rhind onomwonden: “Ik kan je op basis van de datakwaliteit vertellen of een bedrijf wel of niet succesvol is.”

Op de Dag van de Datakwaliteit, georganiseerd door de commissie Datakwaliteit van dm-brancheorganisatie DDMA, liet hij zien hoe hij met een simpele zoektocht op internet bij het ene na het andere bedrijf op onvolkomenheden was gestuit. En dat terwijl je echt niet op al te veel welwillendheid van de andere partij hoeft te rekenen, stelt Rhind. Want: bezoekers van je website zijn zich niet bewust van ‘correcte’ methodes van adressering bij data-entry, of ze staan er onverschillig en soms zelfs vijandig tegenover. Dan helpt het niet als je ze ook nog ‘ns gaat tegenwerken met onduidelijke, onvolledige en mogelijk schofferende keuzes.

Verwarrend

Maar ja. Als de wereld je speelveld is, dan heb je te maken met circa 244 landen (dat kan nogal eens verschillen), meer dan 6.000 talen, grofweg 122 adresformaten en meer dan 35 manieren om persoonsnamen te schrijven, vertelde Rhind. Correcte data, zei u? Met een beetje kennis van zaken en begrip voor culturen kom je al een heel eind.

Zo kun je mensen wel een invulveld voor ‘postal code’ voorschotelen, maar de meerderheid van de landen heeft helemaal geen postcode-systeem. Dat werkt voor veel mensen dus verwarrend. Wat moeten ze daar

invullen? De kans dat ze hier afhaken, is groot. Je kunt iemand ook schofferen – in Nederland mogen we dan zo vooruitstrevend zijn dat we genoeg nemen met een indeling in ‘mr’ en ‘ms’, maar in andere culturen kunnen mensen het als een belediging zien als ze niet hun titulatuur kunnen opvoeren. En daar verlaten weer de nodige potentiële klanten uw site.

Rhind kon zo nog wel een poosje doorgaan. Laat bijvoorbeeld mensen zelf hun land invullen, en je krijgt ontelbare manieren om een landsnaam te schrijven (Germany, Allemagne, Deutschland, BRD, Ger, et cetera). Maar wie – lekker makkelijk – uit een voorgekookt lijstje met landsnamen laat kiezen, moet zich ook de moeite gunnen om er eerst zelf een paar tellen langer naar te kijken. Want in een lijst met internationale benamingen alleen

Data Quality Award: opgeruimd staat netjes



Het werd ‘ns tijd. De bedrijven, en de mensen bij die bedrijven, die zich onvermoeibaar inzetten om hun databases op orde te brengen en te houden, verdienen ook een prijs. Vandaar de Data Quality Award, een initiatief van brancheorganisatie DDMA, Human Inference en Experian. Ook het CRM Genootschap doet mee. Op 29 november wordt bekend welke organisatie als eerste met de eer gaat strijken. De inschrijving voor de eerste editie van de Data Quality Award is geopend. Bedrijven kunnen zich nog online aanmelden op www.dataqualityaward.nl.

Het idee voor de prijs komt uit de koker van directeur Hugo Verwijs van Human Inference, leverancier van datakwaliteitsoplossingen. “Al vele jaren zijn mensen binnen organisaties hard aan de slag om de datakwaliteit te verbeteren”, zegt Verwijs. “Vaak zijn dit de visionairs met een dosis volharding om projecten tot een goed einde te brengen en de kwaliteit op peil te houden. Tegenwoordig staat datakwaliteit meer in de belangstelling door de vele compliance- en customer data integration-projecten. Om die reden leek het mij een goed idee om het harde werken van de afgelopen jaren en het huidige besef dat datakwaliteit onontbeerlijk is voor een goede bedrijfsvoering te onderstrepen met een award.”

Doe geen enkele aanname, behalve dat je aannames fout zijn

de keuze 'België' geven, en niet 'Belgique', zal ook de helft van onze zuiderburen op z'n minst de wenkbrauwen doen fronsen. Een site voor hotelboekingen maakt zich hier schuldig aan, zo liet Rhind zien.

Heel Nederlands

Maar ook ons ministerie van Buitenlandse Zaken – "je zou toch zeggen dat ze daar wel iets afweten van andere landen" – slaagt erin om het bezoekers van de site lastig te maken. In een responsformulier zijn netjes zeven karakters ingeruimd voor een postcode-veld; heel Nederlands. Maar er zijn landen waar mensen toch echt tien karakters kwijt moeten kunnen.

"In het verleden kon ik mensen aan het eind van een presentatie geruuststellen door te verwijzen naar een organisatie die het wél begrepen heeft", zei Rhind. "De KLM." De luchtvaartmaatschappij vroeg voorheen bezoekers naar hun land van herkomst, en schotelde ze vervolgens een site voor in de eigen landstaal, én met de toepasselijke invulvelden. Maar dat is voorbij. Ieder bezoeker die zich moet registreren, krijgt dezelfde eenheidsworst voorgeschoteld. Met weer allerlei internationale misvattingen en kans op wegklikken.

Risicovol

Conclusies trekken en analyses maken op basis van de gegevens in een vervuilde database, is risicovol. Zelfs als je alleen Nederlandse persoonsgegevens verzamelt. Zo vertelde Rhind dat bij een onderzoek op bijna één miljoen 'schone' adressen de naam 's-Gravenhage op



Graham Rhind (GRC Database Information): correcte invoer van data, het laat bezoekers van je internetsite meestal koud.

57 verschillende manieren geschreven werd, waarvan 53 fout. Rhind: "Ga je dan beslissingen nemen op basis van aantallen klanten in Den Haag die volgens de database een bepaald gedrag vertonen of aan een bepaald profiel voldoen, dan zit je dus altijd fout."

Rhind formuleerde twaalf regels voor internationale datakwaliteit:

- benader databeheer vanuit de data, niet vanuit de beperkingen van applicaties, hardware, software, standaarden of iets anders
- verminder de afstand tussen datagebruikers en data – geïntegreerde systemen houden data verborgen
- datakwaliteit moet altijd de norm zijn.

Bezuinig nooit op datakwaliteit

- verzamel liever goede data dan dat je vervuilde data probeert te verbeteren. Valideer data bij invoer
- sta open voor diversiteit
- doe geen enkele aanname, behalve dat je aannames fout zijn
- bestrijd onverschilligheid. Je vindt geen oplossingen als je niet herkent dat er problemen zijn
- culturele en taalkundige verschillen zijn persoonsgebonden, niet plaatsgebonden
- pas je dataprocesen aan op culturele en linguïstische normen, niet andersom
- hou rekening met context
- vind niet het wiel opnieuw uit
- gebruik je gezond verstand! #

(Advertentie)

Uw database écht op orde brengen?

Bel 070 334 42 00

CENDRIS
a  TPGPOST company