

Capitolul 3

Continutul Digital Date Externe în Content

Conectarea la autostrada informațională



punctual

- **Avem nevoie de date externe că să avansăm**
- **Există o planificare a schemei de date / SODD**
- **Datele externe presupun mentenanță continuu**
- **Fluxurile de date externe cer Unelte specifice**

- **Suntem doar verigi în lanțul informațional, la rândul nostru suntem surse de date externe**

Prescurtările de două și trei litere sunt “mama limbajului tehnologic”. Încă din (pre)istoria internetului se pare că aceste prescurtări au o funcționalitate deosebită atât în comunicare cât și în influența deosebită asupra creierului uman – dovadă este platforma mIRC (strămoșul rețelelor de socializare) unde toate discuțiile aveau impresia că sunt astfel: “Sal, e ok?-Yes, asl pls-[Name]-hahaha, lol ! gtg, sry ...” (inteligibil pentru cei care iubesc spațiul virtual).

Recomand înainte de a ne apuca de afaceri online să învățăm terminologia, o simplă căutare cu “dictionar+informatica” ne pune într-o înțelegere mai bună a articolelor care le citim sau a discuțiilor care purtăm. Schema de Date despre care urmează rândurile de mai jos conține mulți termeni prescurtați: RSS, CSV, DOC, XLS, ERP, CRM , etc.

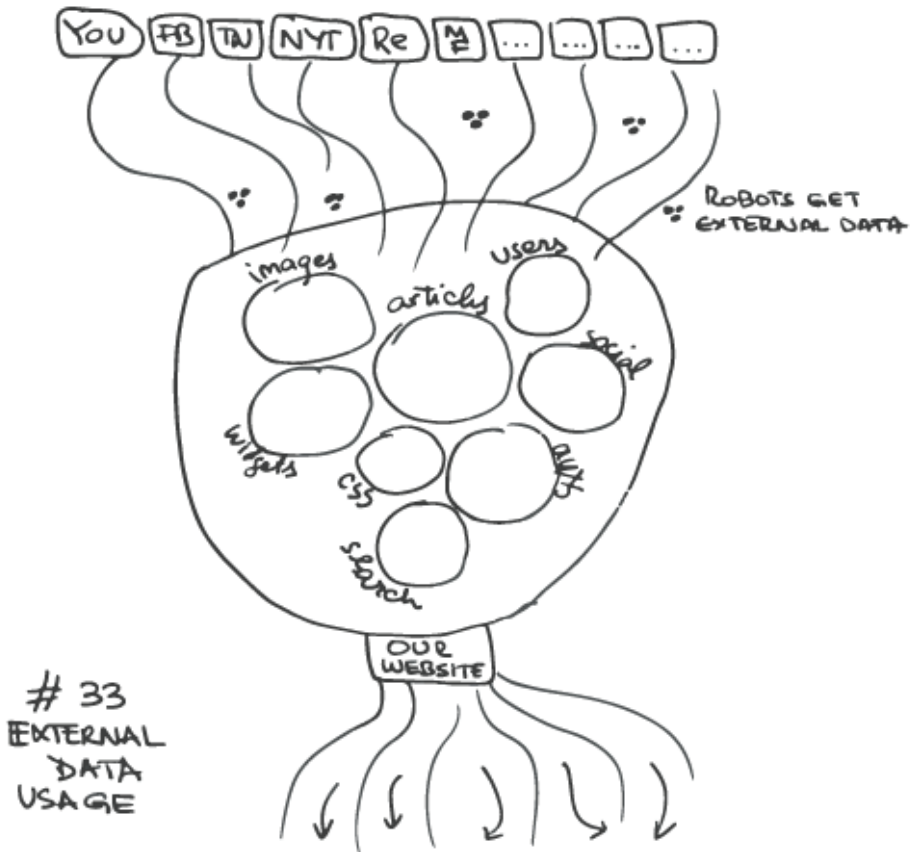
Schema de Date cuprinde a)Sursele de Date, b)Organizarea Datelor, b) Datele în sine, d)Destinația Datelor-o putem prescurta **SODD** (Surse; Organizare; Date; Destinații). Datele externe sunt o categorie aparte în Conținutul unei afaceri online pentru că permite folosirea unor surse externe ca și furnizori care îmbogățesc rapid, cultural și funcțional situl nostru web. Sub forma datelor externe putem accesa orice tip de conținut: articole de presă, titluri, poze, filme, date care ulterior să fie clasificate în baze de date, reacții, scoruri, date meteo, etc. Aici discutăm de orice conținut care nu presupune OPT-IN, adică exprimarea consimțământului expres pentru a face o acțiune. Datele pot fi colectate direct, prin intermediul roboților (în același mod în care Google **strânge** datele din toate siturile mondiale pentru motorul sau de căutare) și **indirect**, prin bunăvoința (și marketingul) celor care acceptă să transmită aceste date.

Surse externe de date

De unde iau date? **SURSE sunt toate siturile și serviciile web care oferă conținut**. Gratuit sau Plătit, Direct sau Indirect. Culese de noi manual sau Importate automat. Ca să identifiți sursele trebuie să știți CE ai nevoie și CÂT ai nevoie. Dacă întrețin un blog de tehnologie și sunt talentat, public propriile materiale și mă rezum la ce creez. Dacă vreau

În plus să comentez actualitatea, o să citesc/citez alte surse web, care devin Surse. La aceste surse pot să le apelez automat printr-un serviciu care să le organizeze și care să mă ajute să găsesc și altele (eu folosesc Feedly). Pot de asemenea să încerc să fac situl mai interesant apelând la widgets (blocuri grafice-text în cadrul sitului, care rulează date externe) care îmi furnizează diverse date externe (exemple: meteo, un motor de căutare personalizat și poate chiar un clasament al echipelor de fotbal).

WEBSITES & SERVICES



Dacă am nevoie de date externe folosesc programe automate (un alt exemplu pozele proprii care sunt făcute publice prin intermediul serviciului Flickr). Se poate introduce și un formular de înregistrare într-un serviciu exterior dar care îl readuc în propriul website sub forma de widget (exemplu: serviciul de newsletter automat prin Mailchimp). Dacă sursele externe sunt foarte importante în afacerea ta, **sunt doar câțiva pași în a-l imita pe Google, să ai proprii tăi roboți, atât de puternici încât să poți accesa orice informație din orice website din lume ...** asumând corect că ai resurse tehnologice porportionale cu nevoile tale.

Tratarea datelor externe

Tratez special datele externe automate? Da, pentru că dacă vreau să le introduc în sistemul meu am nevoie de funcționalitățile adecvate. Ele vin în majoritatea timpului sub forma de text sau așezate sub formă unui text specific, care urmează să fie adaptat propriului site din punct de vedere vizual și funcțional – baze de date. Cea mai mare bătaie de cap care poate apărea este dacă se schimbă în timp formatul datelor în automatizările speciale (în special cele obținute din Data Mining). Și nu uit că nu sunt “ultima verigă” a datelor, la rândul meu voi organiza datele pentru export, gata să fie preluate de alții.



Utilizatorii online

Daca ne este foarte clar de unde luam date externe si unde le dam, nu strica sa avem o imagine asupra

utilizatorilor de internet din lume. Nu de alta, dar poate preluam datele in engleza si le oferim in chineza ! **Asia - 49.5%, Europa - 17.2%, America Latina 10.5%, Africa 9.4%, America de Nord 9%, Orientul Mijlociu 3.6%, Oceania 0.8%.**

vezi <http://www.internetworldstats.com/stats.htm>

Tipuri de date automate



Ce tipologii produc Datele automatizate? Dacă ne referim la articole publicate, avem siturile de tip “agregator” care centralizează automat datele din mai multe surse. Dacă ne referim la magazine online, fluxurile de date automate produc sistemele de comparații de prețuri și marketplace-uri. Automatizarea conținutului de pe bloguri în Social Media face ca

orice articol publicat pe blog să apăra direct în contul tău pe Facebook sau Twitter provocând vizibilitate. Roboții clasificatori de informație produc motoare de căutare sau rapoarte de date agregate sub formă



cutiuța cu idei reține “preluare automata”



“parsing” sau “scanare” sau “scrapping” sau “extragere de date”. Acesta desemnează procesul automat prin care comandăm un program automat, numit BOT sau Robot, să execute o anumită sarcină și să ne întoarcă datele. De exemplu pentru SNEC.ro folosim un robot care strânge date despre toate siturile care tratează produse care pot fi subiectul unor vânzări online, date care ulterior sunt supuse unui al doilea robot dotat cu algoritmi de recunoaștere, pentru a identifica magazinele online. Al treilea robot face scanarea unui potențial magazin online ca al patrulea să îi facă descompunerea datelor în texte din care se extrage conținut.

de studii, alerte, business intelligence, infografice. Sistemele CRM leagă datele vânzărilor cu conturile și activitatea utilizatorilor internet. **Practic orice website se poate conecta la Autostradă Informațională că și Consumator sau Furnizor, în funcție de Strategia sa.**

Datele externe scutesc adesea eforturi inutile și sunt o bună pistă de lansare pentru creatorii de conținut. Bine cadrate, ele devin o resursă pentru afacerea online. Am încercat să tratez aici doar procesul, și nu “aspectele sensibile” unde există o serie de discuții generate de legislatorii din Statele Unite și Uniunea Europeană, la care adaugăm bunele norme deontologice ale practicii cotidiene. De foarte multe ori fluxurile de date îți sunt impuse (vezi magazinele online care trebuie să livreze datele de o anumită manieră și într-un anumit format pentru comparațiile de prețuri sau vânzătorii marketplace).



Sursele de Date Externe sunt foarte importante pentru accelerarea și îmbogățirea Conținutului, și o sursă foarte bună pentru a deveni o referință pentru alții;

**pontul
#33**

verificare



Stiu ce date externe am nevoie pentru site



Stiu unde pot deveni furnizor de date externe



Stiu datele care le stochez, restul pleaca



Datele stocate se salveaza automat



Am sistemul de verificare a datelor stocate



SODD (Surse; Organizare; Date; Destinații).

Checklist: Planificarea schemei informativale SODD in 6 pasi

1. Am nevoie de date externe? Dacă da, ce fel de conținut mă aranjează ? În primul rând, determin tipologia de date externe care îmi vor ajuta afacerea online să apară mai repede pe piață, să vândă mai bine, să dețină Conținut de bună calitate, actualizat și ne-perimabil și le trec prin filtrele CEDAC și 5S (vezi capitole anterioare). Lista cuprinde : poze, descrieri, articole, citate, actualizări, conținut întreg, cioturi de informații, anumite stringuri de date, comportamente, etc.

2. Pot fi EU un furnizor de date? Dacă da, cu ce scop fac asta. Facem parte din sistem și în aceeași măsură în care noi vrem date de la alții, la rândul nostru putem fi un furnizor de date, rezultate din activitatea noastră cotidiană. Punem la dispoziție fluxuri, produse, poze, conținut, RSS, Creative Commons, grafice, infografice, comunicate, etc.

3. Partajare între ce fel de date vreau să rețin și ce vreau doar să tranzacționez informațional. Listele “de luat și de dat” nu sunt foarte lungi, însă fiecare operațiune presupune resurse tehnologice (software, procesare, stocare, bandă) plus resurse de timp și umane pentru Organizare. Chiar dacă avem acces la multe posibilități, limitarea resurselor ne pune în postura să le alegem DOAR pe cele importante și care generează plus-valoare atât în afacere cât și în ecosistemul online.

4. Generare flux cum introduc datele care le rețin în propria bază de date. Noi avem deja o organizare internă a bazei de date și datele externe nu ar trebui să aibă o incidență prea mare în schimbarea structurii acestuia, pentru ca dacă o are, propriul nostru sistem va trebui reproiectat. De aceea este important desenul mișcării datelor - cum intră ceea ce nu controlăm în administrarea noastră – de multe ori ele trebuie editate și verificate, activități care presupun resurse. **Desenul datelor le dă Destinația**, locul în care se manifestă și produc efecte. (de

exemplul numerele de telefon sunt destinate să alimenteze o bază de date de SMS-uri automate)

5. Unelte de care am nevoie că să automatizez procesele (unele sunt deja prevăzute în CMS). Automatizarea proceselor presupune metodele de aducere a datelor și stocarea lor în fișiere sau baze de date. Ca datele să devină vizibile în situl nostru, trebuie să operăm cu unelte specifice, care să le prevedem de la început. Pote fi vorba de calculul tabelar excel pentru prelucrarea fișierelor sau CSV (fișiere comma separated value). Pot fi editoare de text, simple sau combinate cu calcul tabelar online. Pot fi module în cadrul CMS-ului, construite anume pentru automatizare. Sau pot fi plug-ins de Wordpress care sunt deja automatizate pentru prelucrarea de Content. **Toate sunt Unelte.**

6. Prevedem verificarea și mentenanța sistemelor de date. Periodic și prin sondaj. Este destul de complicat de făcut o logică infailibilă a datelor, a căror sursă sunt oameni, supuși deseori greșelii. **Plecăm de la premiza că sistemele nu sunt perfecte și că putem doar spera într-un grad mai mic de eroare.** În planul de management prevedem: frecvența sondajelor, cum validez dacă sunt bune sau nu, cine e responsabil de asta, dacă există recurență în alimentarea externă cu date din același surse. Despre Mentenanță, revin la **principiul de Santatate al Conținutului, ceea ce este relevant și corect este menținut online, ce nu mai este de actualitate pleacă.**