

SECURITATEA UTILIZĂRII TEHNOLOGIEI IDN PENTRU DOMENIILE ÎNREGISTRATE DE ROTLD

Ionuț Sandu

ionut@rotld.ro

Institutul Național de Cercetare-Dezvoltare în Informatică – ICI București

Rezumat: Resursele în Internet utilizează caracterele ASCII (American Standard Code for Information Interchange). În anul 1996, Martin Dürst a propus pentru prima oară introducerea altor seturi de caractere decât ASCII, în asociere cu resursele din Internet, iar prima implementare a fost posibilă în anul 1998 de către Tan Juay Kwang și Leong Kok Yong. Necesitatea introducerii altor seturi de caractere pornește de la extinderea utilizării calculatoarelor în zone precum China, Japonia, Rusia ș.a., zone care nu folosesc caracterele ASCII.

Limba română conține cinci diacritice (ă, â, î, ș și ț), multe din cuvinte au un înțeles total diferit în comparație cu cuvinte cu ortografie similară dar fără diacritice ”paturi” față de ”pături”. Prin aceasta diferențiere se pot identifica corect resursele atribuite acestor cuvinte, fără a introduce erori.

Registrul Național de Administrare al Domeniilor .RO RoTLD a introdus aceasta facilitare doar pentru domeniile care conțin cele cinci caractere menționate mai sus în data de 02 Iunie 2015.

Cuvinte cheie: IDN, domenii .ro, ROTLD, securitate cibernetică, domenii cu diacritice.

Abstract: The Internet resources use the ASCII (American Standard Code for Information Interchange) characters. In 1996, Martin Dürst proposed for the first time the introduction of characters other than ASCII in association with Internet resources, and the first implementation was made possible in 1998 by Tan Juay Kwang and Leong Kok Yong; the necessity for introducing other character sets came from spreading of computer utilization in areas such as China, Japan, Russia, areas that don't use ASCII characters.

Romanian, as a language, has five diacritics (ă, â, î, ș, and ț), and words with diacritics have a totally different meaning compared to words with similar spelling but without diacritics. For instance: “paturi” (beds) and “pături” (blankets). Using diacritics helps identifying correctly the resources allocated to the such words, and reduces the risk of errors.

The National Registry for .ro Domains Administration – RoTLD – has introduced this facility, only for domain names that contain the above mentioned diacritics, since the 2nd of June 2015.

Key words: IDN, .ro domains, ROTLD, cybersecurity, internationalized domain name.

1. Introducere

Denumirea de Internet provine de la prescurtarea a două cuvinte Interconnected și network. Acesta înseamnă o rețea interconectată de calculatoare în care accesul de la un calculator la altul se face prin numele de domenii.

Numele de domenii reprezintă șiruri de caractere asociate adreselor IP aferente calculatoarelor interconectate în rețeaua Internet.

Un nume de domeniu este unic și IP-ul asociat calculatorului este de asemenea unic, iar asocierea domeniu - IP este unică.

Internetul de la apariția sa și până în prezent are un rol dominant în:

- activitățile sociale prin oferirea posibilității de comunicare și socializare între utilizatori aflați la distanțe foarte mari; astfel a apărut sintaxa: „la un click distanță”; prin interconectarea mai multor calculatoare personale, posesorii acestor calculatoare au devenit la rândul lor interconectați, iar distribuirea de informații a devenit facilă indiferent de locație sau timp;
- activitățile economice prin oferirea posibilității de realizarea comerțului la distanță; astfel prin intermediul internetului, o organizație are posibilitatea să fie vizibilă fără granițe; fără internet o organizație poate să își desfășoare activitatea într-o zonă restrânsă ca spațiu, iar prin intermediul internetului aceste limitări teritoriale sunt eliminate și astfel aceasta își poate desfășura activitatea la nivel global; prin eliminarea limitărilor de spațiu în activitățile economice pătrunde și ideea de comparație la nivel global; astfel produsele sau

serviciile achiziționate de către clienți sunt comparate cu alte produse sau servicii furnizate de către alte companii; în acest mod calitatea produselor și serviciilor contează foarte mult, iar companiile furnizoare îmbunătățesc tot mai mult calitatea acestora din cauza concurenței la nivel global.

Dezvoltarea internetului s-a realizat cu precădere în limba engleză. Aceasta presupune bariere de limbă pentru non-vorbitorii de limba engleză.

Numele de domenii au fost create de asemenea în limba engleză și conțin caractere ale alfabetului englez simplu, 26 de litere latine, neaccentuate, cifrele arabe de la 0 la 9, cratimă și punct.

Numele de domenii nu sunt cazenitive astfel nu se face diferența între majuscule și minuscule. Domeniile INTERNET.ro, InTeRneT.Ro și internet.ro sunt echivalente.

Pentru promovarea Internet-ului, este necesar susținerea limbilor și alfabetelor locale în denumirea numelor de domenii.

Procesul de susținere a scrierii multilingve și altor necesități lingvistice și culturale în Internet este cunoscut, în general, ca Internaționalizare.

2. Tehnologia IDN

ICANN (Internet Corporation for Assigned Names and Numbers) a decis introducerea numelor de domenii IDN din categoria ccTLD (country code Top-Level Domains) încă din anul 2009, testarea a durat aproximativ un an, urmând ca din 2010 să poată intra în producție.

Numele de Domenii Internaționalizate (IDN - Internationalized Domain Names) sunt nume de domenii sau adrese web reprezentate cu caractere locale.

Internaționalizarea Numelor de Domenii în Aplicații (IDNA - Internationalizing Domain Names in Applications) - sunt aplicații capabile a recunoaște numele de domenii scrise în (aproape) tot spectrul simbolurilor cunoscute.

Introducerea IDN necesită eforturi de colaborare și coordonare între:

- ICANN;
- Guverne;
- Experți în lingvistică;
- Registratori (ccTLD și gTLD);
- WIPO.

Componentele principale ale implementării IDN implică dezvoltarea tabelor limbilor și a variantelor sale pentru diferite limbi, precum și elaborarea politicii publice privind limba domeniului. Implementarea va implica următoarele trei proceduri:

- standardele tehnice de implementare și protocol IDN;
- politicile privind limba pentru înregistrarea IDN;
- implementarea conștientizării IDN în sistemul de înregistrare.

În cazul unui domeniu de internet clasic browser-ul vede denumirea unui host - expediază o solicitare la serviciul DNS de transformare a adreselor, care la rândul său expediază solicitarea la un server al numelor de domeniu pentru a întoarce o adresă IP ce corespunde denumirii respective a host-ului. Atunci când adresa IP este expeditată înapoi, este efectuată o conectare cu Web server-ul potrivit.

În cazul domeniilor de internet IDN, un browser compatibil primește semnal pentru IDN, el convertește IDN în Punycode – standardul oficial care a fost aprobat pentru convertirea domeniilor IDN în domenii ASCII rezolvabile. Acest proces convertește denumirea într-o secvență ASCII cu prefixul "xn--", care poate fi căutată în denumirea server-ului TLD pentru a determina amplasarea website-ului.

Când un utilizator scrie un nume de domeniu, care folosește caractere ce nu fac parte din cele 37 de simboluri de bază (un Nume de Domeniu Internaționalizat, sau IDN cum ar fi <românia.ro>), aplicația va converti acest IDN într-un nume de domeniu, care folosește numai cele 37 de simboluri de bază (în cazul acesta <xn--romnia-ya.ro>). Modul în care aplicația efectuează această conversie este fixat în regulile elaborate de comunitatea tehnică internațională.

3. Implementarea tehnică

Pentru implementarea tehnică a utilizării domeniilor IDN sunt expuse două metode generice de manipulare a unui șir de coduri Unicode și anume: **ToAscii()** și **ToUnicode()**.

ToAscii() convertește un șir de coduri Unicode într-un șir de coduri ASCII la care este adăugat prefixul ACE "xn--". Această funcție poate eșua în procesul de preparare a stringului pentru conversie.

Algoritmul funcției ToAscii parcurge următorii pași:

1. secvența de intrare este pregătită pentru conversie aplicând procedura NAMEPREP;
2. dacă flag-ul UseSTD3ASCIIRules := TRUE atunci se fac verificări adiționale;
3. asupra secvenței astfel obținute se aplică algoritmul PUNYCODE;
4. se concatenează prefixul ACE "xn--" la cantitatea obținută mai sus;
5. rezultatul reprezintă o secvență ASCII pregătită pentru a fi folosită în așa-zisele slot-uri IDN.

ToUnicode acceptă ca input o secvență de caractere Unicode și produce o secvență de caractere Unicode. Dacă secvența de intrare începe cu prefixul ACE "xn--" atunci funcția produce o etichetă internaționalizată echivalentă care nu este format ACE, altfel secvența de intrare reprezintă secvența de ieșire nealterată.

Algoritmul funcției ToUnicode parcurge următorii pași:

1. secvența de intrare este scanată folosind procedura NAMEPREP;
2. se verifică dacă secvența începe cu prefixul ACE "xn--";
3. se îndepărtează prefixul ACE;
4. se aplică procedura de decodificare PUNYCODE asupra secvenței rămase;
5. se aplică procedura inversă ToAscii asupra rezultatului de la pct. 4;
6. se compară rezultatul de la pct. 5 cu secvența de intrare originală;
7. se întoarce rezultatul calculat la pct. 4;

Aceste metode sunt apelate automat de către browser-ele compatibile IDN, fără ca utilizatorul să conștientizeze acest lucru.

Introducerea numelor de domenii cu diacritice, deși exista de multă vreme, încă din anul 1998, nu toate aplicațiile au fost dezvoltate pentru a le putea folosi, astfel, de multe ori este nevoie să se folosească forma ASCII a respectivului nume de domeniu. Serviciul de email este un exemplu care la această dată nu poate "traduce" în întregime o adresă de email care conține diacritice.

Din punct de vedere al completării fișierelor de zonă ale nameserver-elor cu nume de domenii care conțin diacritice, acestea trebuie completate **numai** în format standard ASCII.

Totodată doresc să atrag atenția asupra faptului că standardele DNS-ului nu permit ca un nume de domeniu să conțină mai mult de 63 de caractere și în acest context, domeniile cu diacritice trebuie să se supună acestor standarde în formatul ASCII (domeniul bucurești.ro se translatează în xn--bucureti-m9c.ro și astfel conține 19 caractere).

Diacriticele românești permise la înregistrarea numelor de domenii .ro sunt prezentate în tabelul de mai jos și fac parte din setul de caractere Unicode numit Romanian Standard:

Caracter Unicode	Code Point	Reprezentare hexazecimală UTF-8	Reprezentare binară UTF-8 (BAZA#INTREG_BINAR)	Descriere
Ț	538	0x21A	2#001000011010.	latin capital letter t with comma below
ț	539	0x21B	2#001000011011.	latin small letter t with comma below
Ș	536	0x218	2#001000011000.	latin capital letter s with comma below
ș	537	0x219	2#001000011001.	latin small letter s with comma below
ă	259	0x103	2#000100000011.	latin small letter a with breve
Ă	258	0x102	2#000100000010.	latin capital letter a with breve
î	238	0xEE	2#11101110.	latin small letter i with circumflex
Î	206	0xCE	2#11001110.	latin capital letter i with circumflex
â	226	0xE2	2#11100010.	latin small letter a with circumflex
Â	194	0xC2	2#11000010.	latin capital letter a with circumflex

Întârzierea implementării sistemului IDN la Registrul de administrare a numelor de domenii .RO se datorează faptului că setul de caractere Unicode din standardul Romanian Legacy conținea caracterele ș și ț cu sedilă (caractere turcești), caractere care erau asemănătoare cu cele din tabelul de mai sus, dar aveau o altă reprezentare față de cea din tabel.

4. Securitatea tehnologiei IDN

În cazul utilizării domeniilor IDN există posibilitatea realizării atacurilor de tip omograf. Astfel un domeniu înregistrat original și oficial poate fi clonat prin utilizarea caracterelor speciale, dar care însă seamănă cu caracterele originale. Figura 1 prezintă un astfel de exemplu.

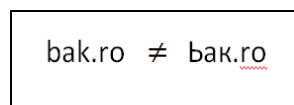


Figura 1. Atac homograf

Exemplul este dat cu caractere rusești. Pentru exemplu domeniilor cu caractere românești există posibilitatea utilizării diacriticelor iar utilizatorii să nu observe diferența dintre diacritice și caracterele de bază din alfabet, figura 2.

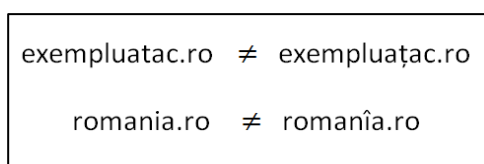


Figura 2. Exemple de atac homograf cu caractere românești

Pericolul este prezent în cazul domeniilor de internet care au în componență caractere asemănătoare cu diacriticele românești:

- a - ă;
- a - â;
- i - î;
- s - ș;
- t - ț.

Evitarea acestor atacuri se face prin înregistrarea tuturor posibilităților de combinații de caractere IDN.

5. Concluzii

În data de 23 februarie 2013 a fost pus în funcțiune un nou sistem de înregistrare nume de domenii .ro, care permite înregistrarea în timp real, atât prin protocolul EPP cât și prin REST. De asemenea, au fost adăugate unele caracteristici noi care să crească nivelul de securitate al serviciilor de administrare domenii .ro, atât pe panoul de control a registrarilor cât și pe panoul de control al deținătorilor de domenii. În aceeași măsură a fost actualizată și lista cuvintelor interzise.

La lansarea acestui sistem, s-a constituit o listă care conține numele de domenii ce reprezintă administrația publică centrală și locală, nume de localități și instituții. Acestea se pot identifica prin furnizarea următoarei expresii "Rezervat pentru Administrația Publică Centrală și Locală" la interogarea sistemului WHOIS. Domeniile au starea "DeleteProhibited" nefiind funcționale. Pentru deblocare, instituția publică care dorește înregistrarea va contacta Registrul de domenii .ro, iar Registrul va analiza și înregistra ulterior domeniul.

Astfel, noul sistem de înregistrare nume de domenii .ro aduce în plus următoarele caracteristici față de vechiul sistem:

- noul sistem de înregistrare și administrare nume de domenii .ro prevede, în plus față de versiunea veche, comenzi EPP și REST de rezervare și activare a unui domeniu, precum și comenzi de transfer domeniu de la un registrar la altul (TRANSFER) și transfer drept de folosință de la un deținător la altul (TRADE);
- cererile de înregistrare pentru domenii noi trimise prin e-mail efectuate de către parteneri (registrar-i) vor continua să fie acceptate, dar domeniile nu se mai înregistrează automat, ci doar se rezervă. Domeniile rezervate vor putea să fie activate de către registrar din panoul de control. În plus, registrar-ii vor putea să activeze domeniile rezervate și prin comenzi EPP sau REST. Recomandăm ca registrar-ii să activeze un domeniu numai după ce au primit plata de la client și sunt siguri că domeniul trebuie activat. La activarea unui domeniu, prețul de înregistrare al acestuia se scade din balanța disponibilă a registrar-ului la acel moment de timp. După activarea unui domeniu, plata este nerambursabilă;
- partenerii (registrar-ii) care folosesc în prezent protocolul EPP pentru înregistrare și administrare domenii .ro, trebuie să folosească noul format al comenzilor EPP după 23 februarie 2013, vechiul format nu va mai fi funcțional;
- noul sistem de înregistrare al unui domeniu .RO are asociat doar date de contact aparținând deținătorului dreptului de folosință. Pentru a nu mai exista confuzie cu privire la cine are dreptul de a utiliza sau administra un domeniu .ro, un nume de domeniu nu mai are asociat date de contact tehnic;
- la înregistrarea unui nume de domeniu .ro, partenerul (registrar-ul) trebuie să specifice tipul de persoană juridică sau fizică a deținătorului dreptului de utilizare a domeniului.

Totodată, în vederea creșterii nivelului de protecție a deținătorilor de nume de domenii .ro, au fost adăugate noi funcționalități pe panoul de administrare domenii .ro, astfel:

- orice operație de actualizare date de contact, creare de nameserver-e, adăugare sau ștergere de nameserver-e asociate unui domeniu, se face numai cu acordul deținătorului domeniului prin transmiterea unui e-mail cu o cheie de autorizare la adresa de e-mail de contact, iar operația se va finaliza numai după ce deținătorul domeniului revine pe pagina de administrare și introduce cheia de autorizare pe panoul de control. În cazul în care adresa de e-mail de contact nu mai funcționează, deținătorul domeniului trebuie să-și actualizeze adresa de email de contact conform procedurii afișată pe www.rotld.ro;

- au fost impuse noi reguli pentru noile parole de administrare domenii, după cum urmează:
 - conțin între 6 și 40 caractere;
 - nu conțin spațiu;
 - conțin cel puțin o literă majusculă A-Z;
 - conțin cel puțin o literă minusculă a-z;
 - conțin cel puțin o cifră 0-9;
 - conțin un caracter special din setul: - . , : ; [] { } _ + = @ # \$ ^ * ? ! | ~

BIBLIOGRAFIE

1. ***: Portalul oficial al Registrului Român de Domenii .RO, <http://www.rotld.ro>
2. ***: Site-ul oficial al ICANN, <http://icannwiki.com/IDN> și <http://icann.org>