**Site-urile web sub lupa proprietății intelectuale (I): Web Scraping**

Av. Alexandru-Laurențiu MIHAI

1. **Introducere**

Protecția drepturilor de proprietate intelectuală în mediul online este un subiect a cărui complexitate derivă tocmai din numeroasele aspecte tehnice relativ inaccesibile și ininteligibile pentru un destinatar nespecializat. „*Universul”* online se caracterizează printr-o dinamică incredibilă, instrumentele juridice clasice fiind adeseori neadaptate (și chiar inadaptabile, în pofida eforturilor doctrinei sau a jurisprudenței) realităților tehnologice contemporane.

Trendul mondial de digitalizare și de *trecere* în online a multor aspecte din viața cotidiană (tranziția de la plata la ghișeu, la plata la Ghișeul.ro – tranziția de la școala fizică la școală online și multe altele asemenea) implică necesitatea unei mai bune cunoașteri a instrumentelor digitale din perspectivă juridică.

Legătura dintre om și *online*, accesul unei persoane la vasta dimensiune a *internetului* trece printr-un prim portal: **pagina web.** Când căutăm informații pe *internet*, în realitate ne lovim de pagini web – începând cu pagina motorului de căutare și terminând cu pagina pe care vom găsi informația sau conținutul dorit. Pagina web este **rama** prin care “*vedem”* Internetul și prin care îl percepem. Totodată, pagina web este o componentă cheie, indispensabilă oricărei activități comerciale. O veritabilă carte de vizită, aptă să convingă și să atragă clientela – un element ce contribuie la conturarea unei imagini a comerciantului.

Având în vedere necesitatea unei mai bune înțelegeri a aspectelor tehnologice ce ne înconjoară și în speranța de a oferi un instrument juridic util atât pentru practicieni ai dreptului, precum și pentru specialiști în domeniul tehnologiei informației, autorul propune ca, printr-o serie de articole relevante, să antameze subiectul **drepturilor de proprietate intelectuală în materia site-urilor web,** precum și a practicilor și mecanismelor care pot constitui o încălcare a dispozițiilor legale în materie de proprietate intelectuală.

Astfel, acest prim articol va antama subiectul ***Web Scraping-ului****,* stabilind o serie de coordonate conceptuale și clarificând o serie de situații, în care acest mecanism se poate circumscrie unor manifestări licite sau unor veritabile încălcări ale drepturilor de proprietate intelectuală.

1. **Concept. Preluarea conținutului altor site-uri și afișarea pe site-ului propriu**

*Scraping-ul* (alternativ denumit și *harvesting* sau *data extraction)* reprezintă o **tehnică constând în colectarea unor informații de pe pagini web, de omanieră automată și sistematică prin intermediul unui program de calculator special realizat în acest sens.**

Întrucât *informația* are o valoare intrinsecă deosebită, nu este dificil a observa multiplele utilizări ale *web scraping-ului*: în sfera comerțului electronic, *scraping-ul* asigură o monitorizare a prețului afișat pe paginile web ale competitorilor; în marketing, feedback-ul clienților despre un anumit produs sau serviciu este colectat în vederea ajustării propriei strategii de branding.

Totodată, *scraping-ul* are multiple valențe și în sfera politică – companii specializate în consultanță politică colectează informații (de pe Facebook– spre exemplu, de câte ori și în ce context a apărut numele unui politician în comentarii și postări) și oferă analize cu privire la *sentimentele*, *dorințele și comportamentul* alegătorilor. De asemenea, activitățile de cercetare științifică în diverse domenii (*neuropsihologie*) presupun colectarea unor informații de pe internet în vederea modelării unor tipare ale comportamentului uman.

**Reprezintă aceasta o activitate ilicită?** *Per se*, *scraping*-ul, adică simpla activitate de colectare și extragere a informațiilor de pe pagini web (de altfel publice), nu se circumscrie unei fapte ilicite. Totuși, în anumite circumstanțe, în funcție de **scopul și natura informațiilor colectate**, *scraping-ul* poate reprezenta o gravă încălcare a drepturilor de proprietate intelectuală.

În primul rând, o problemă particulară vizează **colectarea unor informații de pe alte site-uri și afișarea conținutului acestora pe propriul site.**

Învederăm că există numeroase pagini web al căror unic scop este tocmai de a face *scraping.* Un exemplu interesant este „*Compari.ro”*, realizat exclusiv pentru a colecta informații despre prețurile unui anumit produs, afișând ulterior o listă comparativă a acestora. Apreciem, totuși, că această practică este corectă, fiind în beneficiul consumatorului final, care este mai bine informat.

Totuși, există și site-uri care utilizează tehnica *scraping-ului* de o manieră parazitară. Spre pildă, pe site-ul *„Booking.com”* se listează diferite proprietăți, în vederea închirierii acestora în regim hotelier. Există însă o serie de site-uri parazit care preiau informațiile, pozele și conținutul de pe pagina *„Booking.com”* afișându-le pe site-ul propriu. Totuși, nu se pot face rezervări și nu se poate contacta titularul apartamentului, site-ul parazit trimițând înapoi la *„Booking.com”.* Se poate afirma că o astfel de pagină este lipsită de funcționalitate, nerealizând niciun beneficiu. Totuși, este o practică constantă ce poate fi întâlnită în spațiul virtual. Motivul este unul simplu: pentru fiecare vizualizare se pot încasa sume de bani ca urmare a reclamelor afișate pe website-ul parazit, prin intermediul *Google Ads* sau al altor servicii de publicitate similare.

În cele ce urmează, pentru a putea delimita situațiile în care *scraping-ul* constituie o activitate parazitară și/sau o încălcare a drepturilor de proprietate intelectuală vom prezenta viziunea Curții de Justiție a Uniunii Europene (“CJUE”).

1. ***Web scraping-ul* sub lupa considerentelor CJUE**

O jurisprudență importantă a CJUE antamează problema *scraping-ului*. Cauza C-30/14 (**Ryanair v. PR Aviation, 2015)** a gravitat în jurul următoarei probleme: **Agenția de turism *„PR Aviation”* afișa pe site-ul propriu o listă comparativă a zborurilor low-cost, a rutelor și a ofertelor companiilor diferitelor companii aeriene, aceste informații fiind colectate prin *web scraping* de pe pagina Ryanair.**

PR Aviation a fost chemat în judecată de Ryanair, pentru motivul încălcării drepturilor de proprietate intelectuală cu privire la ***bazele de date*** pe care Ryanair le administra (*rutele de zbor, prețurile practicate de companie*), baze de date protejate accesibile prin intermediul site-ului Ryanair și aflate sub protecția Directivei 96/6/CE.

Preliminar, este important de a aminti următoarele aspecte: la nivelul Uniunii Europene, Directiva 96/9/CE a Parlamentului European și a Consiliuluireprezintă un **instrument referitor la protecția juridică a bazelor de date.** În sensul Directivei, *„bază de date”* înseamnă o culegere de opere, de date sau de alte elemente independente, dispuse sistematic sau metodic și accesibile în mod individual prin mijloace electronice sau de altă natură.Totodată, învederăm că există dispoziții relevante și la nivel național: Legea 8/1996 conține prevederi (articolul 8, precum și articolul 140 și următoarele) care oferă o protecție ***bazelor de date***.

Amintim că Directiva 96/6/CE (transpusă de o manieră fidelă și în conținutul Legii nr. 8/1996, fiind aceleași raționamente și nivel de protecție la nivel național) prevede:

* În principal, ***structura*** unei baze de date poate fi protejată prin **drept de autor**, dacă implică o **activitate creativă cu privire la modul de organizare și aranjare.** Spre exemplu, o persoană alcătuiește o listă a imnurilor statelor membre ale Uniunii Europene, ordonate în funcție de numărul de cuvinte din imn, de la cel mai lung la cel mai scurt.
* În subsidiar, ***conținutul*** unei baze de date poate fi protejat în baza unui ***drept sui generis***, cu condiția să se fi făcut ***o investiție substanțială în compilarea datelor*** (financiară, materială, de timp). Este important de menționat că nu sunt protejate elementele privite individual, ci baza de date ca o ***colecție***. Spre exemplu, dacă o persoană identifică, descoperă și compilează o listă cu toate clădirile în stil Art Nouveau din București, ca urmare a efortului considerabil depus, această colecție poate fi protejată în ansamblul ei deoarece a implicat o ***investiție substanțială***.

În fine, instanțele naționale au considerat că ***informațiile*** afișate pe site-ul Ryanair *(rute de zbor, prețuri practicate de companie)* nu se află sub protecția Directivei 96/6/CE, nereprezentând veritabile *baze de date*.

Totuși, prin întrebarea adresată CJUE, s-a pus problema dacă ***baze de date care nu se află sub protecția Directivei 96/6/CE (adică informațiile precum cele în cauză) pot fi protejate printr-un contract***. Acest lucru deoarece în momentul în care PR Aviation accesa site-ul Ryanair pentru a face *scraping*, se considera a fi acceptat *Termenii și Condițiile* impuse de Ryanair care prevedeau interdicția preluării informațiilor, fără autorizația titularului, de pe site-ul companiei de zbor.

În fine, CJUE a statuat că nimic nu se opune ca o bază de date care nu se află sub protecția Directivei 96/6/CE (adică o simplă colecție de informații) să fie protejată pe temei contractual, angajându-se răspunderea contractuală a celui care reproduce sau afișază public fără consimțământul titularului paginii de pe care au fost preluate.

Soluția CJUE a fost adeseori înțeleasă greșit, concluzionându-se că site-urile care fac *scraping* întotdeauna ar încălca drepturi de proprietate intelectuală. **CJUE nu a constatat ca PR Aviation ar fi încălcat vreun drept al companiei Ryanair**. CJUE a oferit doar un răspuns coerent și logic. Dacă o colecție de informații nu poate fi protejată de Directivă, această din urmă Directivă împiedică titularul să limiteze contractual posibilitatea terților de a o folosi? Răspunsul, evident, a fost nu.

Drept urmare, PR Aviation ar fi putut răspunde numai contractual în baza *Termenilor și Condițiilor* (dacă ar fi fost considerate ca acceptate la momentul accesării site-ului Ryanair), iar nu pe temeiul încălcării unui drept de proprietate intelectuală.

La nivelul statelor membre, însă, a existat o reticență consistentă în privința posibilității de a angaja răspunderea unor site-uri OTA (Online travelling agencies) precum PR Aviation. Aceasta întrucât ***site-urile de scraping nu acceptă ab initio Termenii și Condițiile,*** accesând conținutul site-ului Ryanair fără a fi de acord cu termenii prevăzuți.

Totodată, din punct de vedere tehnic, site-urile care folosesc tehnica *scraping*-ului **nu accesează propriu-zis** site-ul de pe care preiau informații, ci utilizează programe pentru colectarea informațiilor din motorul de căutare sau alte asemenea metode prin care obțin conținutul unei anumite pagini.

De o manieră interesantă, practica judiciară la nivelul altor state ale Uniunii Europene a relevat că încercările de a limita compararea de prețuri și rute aviatice prin intermediul site-urilor precum PR Aviation este o atingere adusă consumatorilor, limitându-le accesul la informație. Drept urmare, considerente de ordine publică ce țin de interesul consumatorului ar împiedica interdicțiile companiilor aeriene vizând *scraping*-ul.

1. **Considerente finale privind *web scraping-ul***. **Delimitarea dintre licit și ilicit**

*Scraping*-ul va constitui o activitate ilicită atunci când informațiile colectate, preluate și afișate pe site-ul propriu se bucură de o protecție specifică. Drept urmare, un site care colectează **simple informații** referitoare la prețurile practicate de alți comercianți, afișandu-le în vederea realizării unei liste comparative nu se poate aprecia ca desfășurând o activitate ilicită.

În schimb, în momentul în care informațiile preluate pot fi clasificate drept **baze de date** (în sensul Directive 96/9/CE sau a legislației naționale), *scraping*-ul devine o activitate ilicită. Cu atât mai mult, preluarea sistematică a unor texte științifice, literare sau altele asemenea care se bucură de protecție a drepturilor de autor (**copyright**) reprezintă o activitate ilicită. În acest sens, redăm cu titlu de exemplu activitatea parazitară desfășurată de anumite persoane care colectează și publică pe site-ul propriu conținutul unor știri sau a unor texte preluate automat de pe alte pagini web.

Subliniem că este posibil ca **simplele informații** să fie protejate pe temei contractual, prin acceptarea *Terms and Conditions*. În acest sens, titularul paginii, de pe care se colectează conținutul, poate interzice *scraping-ul*. Totuși, **este discutabil dacă în ipoteza *scraping-ului***, **unde nu există nicio activitate propriu-zisă de a *accesa* site-ul** **și de a se confirma *Terms and Conditions***, **totul fiind realizat de o manieră sistematică prin intermediul unor programe de calculator, o astfel de protecție contractuală se poate considera ca fiind efectivă.**

**În concluzie, tehnica scraping-ului nu este *per se* ilicită, ci implică o analiză temeinică a obiectului *„derivat”* – ce anume este preluat de pe un site pe altul?**

Astfel, *scraping-ul* rămâne un instrument des întâlnit în spațiul online, având o dublă valență: fie reprezintă o activitate permisă, licită, uneori cu un beneficiu pentru consumatori (*e.g.* site-urile Vola.ro sau Skyscanner.net care permit compararea rutelor, prețurilor și a altor oferte din partea companiilor de zbor), fie o activitate parazitară, ilicită, utilizată pentru obținerea unor venituri din fraude intelectuale.

*Un articol semnat de Av. Alexandru-Laurențiu MIHAI, Junior Lawyer, -* *lmihai@stoica-asociatii.ro* *-* [*STOICA & Asociații*](https://www.stoica-asociatii.ro/)

**Bibliografie**

1. Directiva 96/9/CE a Parlamentului European și a Consiliului referitoare la protecția bazelor de date.
2. Legea 8/1996 privind dreptul de autor și drepturile conexe.
3. Directiva 2019/790 a Parlamentului European și a Consiliului privind dreptul de autor și drepturile conexe pe piața unică digitală și de modificare a Directivelor 96/9/CE și 2001/29/CE.
4. Decizia Curții de Justiție a Uniunii Europene în cazul C-30/14 (Ryanair Ltd v. PR Aviation BV).
5. Ciprian Raul Romițan,*Protecția juridică a bazelor de date prin drept de autor sau printr-un drept sui-generis,* Revista Română de Dreptul Proprietății Intelectuale, Nr. 2/2015, București, 2015
6. Felicity Reeve, Will Deller, *Website database owners may use terms and conditions to prohibit screen-scraping*, 2015, disponibil la: <https://www.twobirds.com/en/news/articles/2015/global/website-database-owners-may-use-terms-and-conditions-to-prohibit-screen-scraping>
7. Ion Țigănaș, *Măsuri juridice împotriva încălcării drepturilor de proprietate intelectuale în mediul online (experiența Republicii Moldova)*, Revista Română de Dreptul Proprietății Intelectuale, Nr. 3/2020, București, 2015.
8. Patrycja Szwed, *Is web scraping legal? A short guide on scraping under EU Law,* 2021, disponibil la: <https://discoverdigitallaw.com/is-web-scraping-legal-short-guide-on-scraping-under-the-eu-jurisdiction/>