

Digitální ARCHEOLOGIE

Ve srovnání s tím, na co lze narazit v archivech z dob internetového pravěku, mohou historické vykopávky jen blednout závistí. Pojdme se společně podívat na počátky internetu.

PETR KRATOCHVÍL, CHRISTOPH SACKMANN

Sedmnácté století může být pro milovníky historie dobou Ludvíka XIV. a směšných paruk, devatenácté století zase většinou páry. Jak by ale historici označili přelom tisíciletí? Možná jako dobu internetovou.

Každé historické období má své vykopávky a artefakty, nad kterými se vědci rozplývají a obyčejné lidé je obdivují (nebo se jim smějí) v muzeích. Přelom tisíciletí to ale tak snadné mít nebude. Základní otázkou zde totiž je, co z teh-

dejší (internetové) kultury zanecháme pro příští generaci. Obecně sice panuje povědomí, že internet nezapomíná, pro věci starší deseti let ale platí spíše opak. Počítačový vědec Scott Ainsworth zjistil, že 35 až 90 procent všech webových stránek má jen jednu oficiální digitální kopii. Odhady jsou obtížné, protože konkrétní údaje nejsou vždy k dispozici a zejména u jednotlivých kopií může velmi rychle dojít ke ztrátě. Bezpečná archivace, kdy je pořízeno více než šest ko-

TECHNOLOGIE: Jazyky internetu

HTML

Pomocí jazyka HTML (Hyper-Text Markup Language) můžete definovat základní strukturu obsahu stránek. Jazyk byl vyvinut ve švýcarském CERN, v současnosti se o jeho vývoj a nové standardy stará World Wide Web Consortium (W3C).

JavaScript

Specializovaný jazyk umožňující spuštění aplikací přímo na webových stránkách. Nejčastěji je používán například pro práci s formuláři nebo pro měnící se reklamní bannery. Jazyk mají na svědomí Netscape a Sun Microsystems.

CSS

Kaskádové styly (Cascading Style Sheets) doplňují HTML definováním vzhledu jednotlivých prvků stránky. Konstrukce a obsah stránky tak zůstávají odděleny. To je důležité například z toho důvodu, aby projektant a programátor mohli společně pracovat na jednom projektu.

Flash

Komerční software od firmy Adobe, sloužící na webu ke snadnějšímu vytváření her či přehrávání videa. Bohužel také hodně často obsahující zranitelnosti. Předpokládá se, že ho brzy nahradí nový standard HTML5.



pií, nepokrývá ani osm procent webu. Tohoto žalostného trendu si je vědoma celá řada odborníků, a proto vznikají projekty zabývající se archivací internetu. Vzhledem k celkovému objemu dat uložených na webu jde o nekonečnou práci, a tak příslušní lidé bojují alespoň za záchranu stránek typických pro dané období.

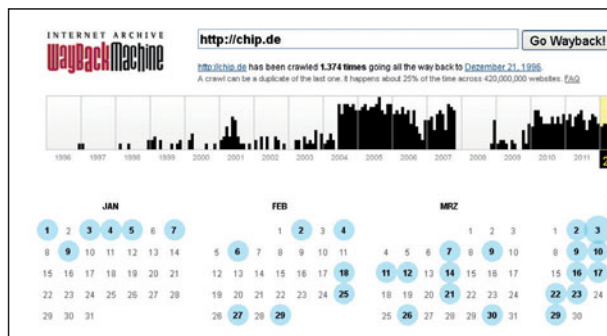
První kroky internetu

Když se v současné době řekne internet, většina nezkušených uživatelů si vybaví webové stránky Seznamu. Na přelomu tisíciletí byly fenoménem osobní weby, a tak byly pro celou řadu surfařů internetem webové stránky GeoCities. Pokud si například na konci tisíciletí chtěl český uživatel založit osobní stránku a prozradit světu, co má rád a kdy se narodil, na výběr měl buď domácí webhosting **mujweb.cz**, nebo zahraniční, ale komfortnější **www.geocities.com**. Není tedy divu, že zde celá řada domácích uživatelů přispěla k vytvoření nezanedbatelného zlomku internetové historie. Pokud i vy patříte mezi vlastníky deníčku na GeoCities, možná vás potěší zpráva, že zaměstnanci projektu dokázali před jeho ukončením uložit do archivů přes dva miliony webových stránek. Od té doby jsou zaneprázdnění tříděním spamových a malwareových stránek a nelegálního obsahu. Tyto záležitosti totiž v posledních letech provozu GeoCities představovaly dost velký problém. Všechny ostatní stránky jsou uloženy na serveru OoCities a v současné době jsou volně přístupné – celkem zabírají téměř 640 GB. Jsou klasifikovány podle svých starých „obcí“, od Area 51 (sci-fi & fantasy) po Yosemite (outdoor a sporty). Lze mezi nimi najít např. předchůdce stránek fanoušků Facebooku (Stránka Denise), sbírku receptů na Che-

ARCHIVNÍ SLUŽBY: Pohled zpět

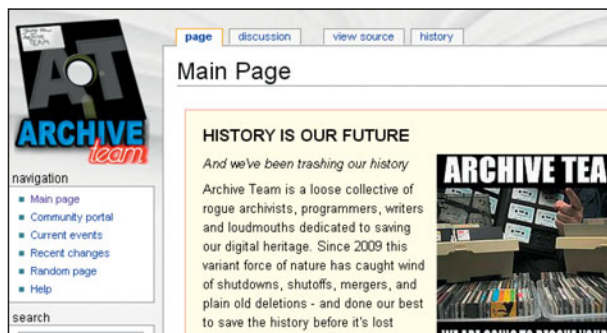
Internet Archive (archive.org)

Služba Internet Archive existuje od roku 1996 ve formě digitální knihovny. I to je důvod, proč zde najdete nejen archivy s webovými stránkami, ale i texty, audio- a videosoubory nebo software. Většinu uživatelů však bude zajímat sekce Wayback machine, kde si lze prohlédnout, jak dříve vypadaly zajímavé weby.



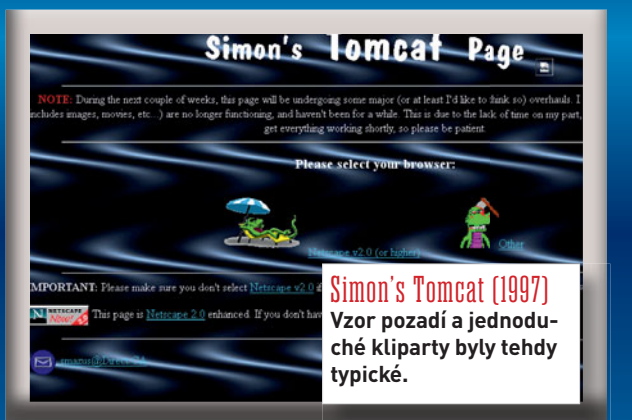
Archive team (archiveteam.org)

Tento anonymní archivní projekt funguje zejména díky týmové práci. Pomocí nahrání zprávy se na archivační práci může podílet každý uživatel. Styl tohoto projektu je tedy podobný například Wikipedii. K dispozici ke stažení je sto gigabajtů materiálu, a to včetně relativně nového archivu z mikroblogovací služby Tumblr.



OoCities (oocities.org)

Webhostingový projekt Yahoo s názvem GeoCities asi není nutné většině uživatelů představovat. Jeho poslední dny se podařilo zachytit archivnímu projektu OoCities, který uložil přibližně dva miliony jeho stránek. Data jsou uspořádána podle GeoCities „obcí“ do 29 témat.





fkoč, místa setkání absolventů, stránky politických aktivistů a klubů, stránky s legračními obrázky i soukromé domovské stránky s osobními daty a vzorovými pracemi. Často vám na první pohled budou tyto stránky připadat zvláštní a primitivní. Prvotní verze stránek (např. eBay nebo Amazon) z konce devadesátých let byly totiž vytvořeny tak jednoduše, že by je v současnosti dokázal mnohem lépe naprogramovat i žák základní školy. Přesto je k nim nutné vzhlížet s respektem, protože jsou součástí naší historie.

Záchrana pro bity a bajty (2200)

Do budoucna ale nepřežijí jen pozůstatky po GeoCities. V databázích služby Internet Archives: Wayback machine najdete přes 240 miliard stránek, archivovaných už od roku 1996. Podobných služeb existuje na internetu více, žádná z nich však nenabídne archivaci dat v takovém rozsahu. A jak je na tom domácí internet?

V roce 2000 byl zahájen projekt internetového archivu pro Českou republiku, který nyní najdete na internetové adrese www.webarchiv.cz. O shromažďování dat se stará Národní knihovna ČR ve spolupráci s Moravskou zemskou knihovnou a Ústavem výpočetní techniky Masarykovy univerzity. Vzhledem k tomu, že nedostatečná legislativa nutí knihovnu uzavírat smlouvy o zveřejnění díla se všemi, kdo na internetu publikují, může veřejnost využívat jen část těchto archivů. Proto při pokusech o vyhledání historie většiny webů získáte jen lakonickou zprávu „Přístup zamítnut. Z důvodu omezení Zákonem č. 121/2000 Sb. (autorský zákon) může být tento dokument zobrazen pouze v prostorách knihovny.“ a paradoxně i odkaz na mezinárodní Internet Archive, který si se zákony příliš hlavu neláme a používá odlišný přístup – ze svého archivu vyřadí ty, kterým se tento postup nelíbí.

Celý tento právní problém je poněkud rozporuplný, uvědomíme-li si, že podobnou práci jako archiváři dělají každý den na webu vyhledávače. Například Google má na svých serverech po omezenou dobu kopie miliard stránek, které si lze zpětně prohlédnout v Google Cache. Ve Wayback machine se stránka objeví až přibližně po devíti měsících.

Experti se ale domnívají, že budoucí generace by měla mít možnost disponovat nějakým materiálem, po jehož přečtení by se dozvěděla, co lidé na počátku internetové éry považovali za zajímavé, jak byl takový internet v plenkách organizován a jak byla nová technologie zvaná internet vůbec využívána.

PETR.KRATOCHVIL@CHIP.CZ

DESET PRAVIDEL STRÁNEK Z MINULÉHO TISÍCLETÍ:

V počátcích internetu vládla mezi tvůrci webů zcela jiná pravidla, a pokud jste chtěli být in, bylo nutné se jich držet.

1) Zapomeňte na soukromí

Čím více jste toho o sobě prozradili, tím zajímavější byly vaše stránky. Nejdůkladnější uživatelé nejen vytvářeli seznam svých oblíbených jídel, ale kromě svého telefonu často přidávali i datum narození a adresu.

2) Úvodní stránka

Jen naprostí amatéři neměli dokonalou úvodní stránku s obrovským obrázkem na pozadí a tlačítkem VSTUPE, na které bylo nutné kliknout, aby se nahrála samotná osobní stránka.

3) Blikající animované ikony

Pro epileptiky nikdy nebylo surfování těžší – téměř na žádném webu nechyběly blikající a animované ikony. I ti největší konzervativci měli na odkazu k e-mailu alespoň otáčející se zavináč.

4) Stránky ve výstavbě

Důkazem dynamiky a pokrokovosti webu byl nápis „Stránka ve výstavbě“, obvykle (v souladu s předchozím pravidlem) doplněný animovaným dělníkem s krumpáčem či sbíječkou.

5) Třetí rozměr

3D bylo hitem už v minulém století. Vzhledem k tomu, že představy Jamese Camerona byly ještě v nedohlednu, stal se nástrojem vizionářů Word a jeho WordArt.

6) Počítadlo

Mírou úspěšnosti každého webmastera bylo počítadlo, které ukazovalo, kolik surfařů si prohlédlo jeho úžasně stránky. Pokud na něm nebylo alespoň trojmístné číslo, bylo nutné to po nocích klikáním napravit.

7) Optimalizace

Zatímco dnes se snaží vývojáři optimalizovat web pro široké spektrum návštěvníků, v minulosti se museli surfaři přizpůsobovat tvůrcům. Oblíbená byla hláška „Tento web je optimalizován pro Internet Explorer a rozlišení 800 x 600“.

8) Odkazy

Správný majitel webové stránky měl spousty přátel a nebál se to dát vědět. Obvyklou formou byly textové odkazy na stránky přátel, ti nejpokrokovější webmasteri pak pro ně vytvářeli kreativní ikonky.

9) Kouzla s textem

Na burze nebo ve zpravodajství u důležitých událostí běží v dolní části obrazovky text. Tuto techniku využívali už dřívější majitelé webů a důležitost zprávy typu „Jste na úžasném Jirkově webu“ ještě zdůrazňovali blikáním. Navíc, pokud si surfař chtěl zprávu přečíst, strávil na webu mnohem více času.

10) Multimédia

V minulém tisícletí prožívala multimédia svůj rozmach a jen outsider pohrdl možností zpestřit návštěvníkovi prohlížení stránky přehráváním audio(midi)souboru, ideálně ve smyčce, aby si surfař hudební zážitek užil co nejvíce.

