

NetMonitor

**Výzkum sociodemografie návštěvníků internetu
Výzkum návštěvnosti internetu**

Základní metodika
verze s platností od 06/2010

**Vypracovala MEDIARESEARCH, a. s.
ve spolupráci s GEMIUS, S. A.**



OBSAH

1	Úvod	4
2	Projekt ve stručných heslech	5
3	Zařazování médií do výzkumu	7
3.1	Pravidla pro zařazování médií ve výzkumu NetMonitor	7
3.2	Pravidla pro umístování měřících skriptů	8
4	Základní popis technologie měření na straně webové stránky (site-centric)	12
4.1	Technologie gemiusTraffic	12
4.2	Hit-kolektor zajišťující sběr dat	13
5	Základní popis technologie měření na straně internetového prohlížeče (user-centric)	14
5.1	Aplikace netSoftware	14
5.1.1	Sběr dat (logy BHO)	15
5.1.2	Modifikace HTTP hlavičky User-Agent	15
6	Výzkum návštěvnosti	16
6.1	Algoritmus pro reportování reálného uživatele (site-centric) ...	16
7	Výzkum sociodemografie	19
7.1	SEKV	19
7.2	Pop-up Panel	20
7.3	SW Panel	21
7.4	Rektutace respondentů	22
7.5	Komunikace s panelem a odměny	22
7.6	Joint Panel	23
7.6.1	Princip stanovení reportujícího panelu	24
7.6.2	Vlastnosti reportujícího panelu	25
7.7	Normy a reprezentativita	26
7.8	Vážení	27
8	Doménový filtr, definice sekcí, kategorizace	29
8.1	Doménový filtr	29
8.2	Definice sekcí	29
8.2.1	Webové rozhraní pro definici sekcí médií na základě URL masek	30
8.3	Kategorizace médií	31
9	Modulární rozšíření projektu	33
9.1	Měření streamovaného videa	33
9.2	Měření používání komunikátorů	33
9.3	Měření přístupů z mobilů	33

9.4	Měření rychlosti připojení	33
10	Analytické nástroje	34
10.1	On-line rozhraní gemiusTraffic	34
10.2	On-line aplikace NetMonitor	37
10.3	Lokální aplikace gemiusExplorer	39
11	Data a reporty	42
11.1	Výsledky a data z měření návštěvnosti	42
	11.1.1 On-line reporty	42
	11.1.2 Denní reporty	43
11.2	Výsledky a data z výzkumu sociodemografického profilu návštěvníků spojená s daty o návštěvnosti	44
	11.2.1 Měsíční reporty o návštěvnosti	44
	11.2.2 Měsíční report o sociodemografii	44
	11.2.3 Měsíční data pro gemiusExplorer	45
12	ukazatele a sociodemografické proměnné	46
12.1	Definice ukazatelů	46
12.2	Sociodemografické proměnné v datech	48

1 ÚVOD

Tento dokument je základní verzí metodiky projektu NetMonitor – nejrozsáhlejšího výzkumu internetu v České republice.

Materiál popisuje základní principy a parametry projektu, které vycházejí z dohod mezi realizátorem projektu a zadavatelem projektu.

2 PROJEKT VE STRUČNÝCH HESLECH

Realizátor projektu: Realizátorem je společnost MEDIARESEARCH, a. s. (www.mediaresearch.cz) ve spolupráci s polskou společností Gemius S.A. (a její dceřinou společností Gemius s.r.o.) www.gemius.com

Zadavatel a vlastník dat: SPIR, z.s.p.o. www.spir.cz

Období výzkumu: 1. 2. 2008 – 31. 5. 2011

Výzkumné techniky: kombinovaná metodika zahrnující:

- měření návštěvnosti na straně měřeného serveru (site-centric)
- panelové šetření realizované kombinací metod:
 1. Pop-up panel – měření na základě pop-up dotazníků na straně měřeného serveru (site-centric)
 2. SW panel - měření aktivity uživatelů na straně internetového prohlížeče (user-centric)

Použitá technická platforma:

- site-centric - systém gemiusTraffic
- user-centric - lokální aplikace netSoftware

Technologie použitá na straně měřeného serveru (site-centric): javascriptové kódy gemiusTraffic vkládané do stránek.

Technologie použitá na straně internetového prohlížeče (user-centric): aplikace netSoftware s plug-inem do internetových prohlížečů *Microsoft Internet Explorer*, *Mozilla Firefox* a *Opera* zaznamenávající aktivitu uživatele na internetu.

Měřená média: Měřeným médiem může být libovolný webový server, který má zájem zúčastnit se výzkumu NetMonitor a jehož zařazení není v rozporu s Etickým kodexem SPIR. Primárně je projekt NetMonitor určen pro výzkum Médii, které jako celek, případně alespoň jejich část splňují definici Reklamního média. Tedy na svých stránkách nabízejí placené umístění reklamního obsahu (mají definován ceník umístění reklamy). Měřené médium je uvedeno zadavatelem na seznamu reportovaných médií. Zadavatel má právo odmítnout zařazení webového serveru do projektu.

Média (servery, resp. internetové stránky) jsou do projektu aktivně zapojeny, tedy mezi nimi a zadavatelem existuje smluvní vztah a servery jsou vybaveny měřicími skripty systému gemiusTraffic. Aktivně zapojené servery jsou měřeny metodou site-centric i user-centric. Podrobná definice viz odstavec 3.1.

Zdroj dat pro strukturální vážení: Kontinuální výzkum TVM (ATO – Mediaresearch).

Dodávka dat:

- data o návštěvnosti - on-line přes rozhraní gemiusTraffic či následující den po skončení měřeného období (den, týden, měsíc) v „on-line rozhraní OLA, případně 1. pracovní den po 24. dni po ukončení měřeného měsíce jako měsíční off-line report s reportovanými statistikami na bázi ukazatele RU – Reálný Uživatel,
- data s přepočtem na reálné uživatele a sociodemografická data – 1. pracovní den po 24. dni po ukončení měřeného měsíce ve formátech
 - pro software gemiusExplorer a v ASCII pro ostatní softwary (k dispozici oprávněným uživatelům na URL předané po produkci), viz kap. 11.2.3 a
 - jako informace o přepočtené návštěvnosti v měsíční zprávě.

Doména s výsledky z výzkumu: netmonitor.cz
(reporty s výsledky, on-line aplikace OLA).

Dostupnost výsledků: Základní výsledky návštěvnosti bez propojení na sociodemografické výsledky jsou dostupné veřejně na www.netmonitor.cz. (viz kap. 11.2.1 a 11.2.2.). Další podrobnější údaje může poskytnout výhradně vlastník dat.

3 ZAŘAZOVÁNÍ MÉDIÍ DO VÝZKUMU

3.1 Pravidla pro zařazování médií ve výzkumu NetMonitor

Provozovatel

Společnost, která zapojí internetové médium na doméně druhého řádu do měření NetMonitor. Mění-li médium Provozovatele a Zadavatel nemá proti této změně námitky, je nutno tuto změnu oznámit Realizátorovi nejpozději do 20. dne měsíce, v němž k této změně došlo. V případě, že se tak stane později, není možné tuto změnu aplikovat na strukturu dat již za tento měsíc.

Médium

Souhrn www stránek, které společně tvoří obsahový celek, a všechny běží na jedné stejné doméně druhého řádu (např. Médium *server.cz* obsahuje *cokoliv.server.cz* i *server.cz/cokoliv*). Do výzkumu je možné na žádost Provozovatele zařadit libovolný, jím provozovaný, webový server, pokud jeho obsah není v rozporu s Etickým kodexem SPIR. Zadavatel si vyhrazuje právo zamítnout zařazení serveru do projektu NetMonitor. Do projektu jsou primárně zapojovány servery, které jako celek odpovídají, případně alespoň jejich část odpovídá definici Reklamního média. 2 různá Média jsou disjunktní. Každé médium může patřit jen jednomu Provozovateli.

Reklamní médium

Médium, na kterém je míra zobrazování reklamního obsahu na jednotlivých stránkách větší nebo rovna 90 %. Prostor pro zobrazování reklamního obsahu je zároveň komerčně nabízen k reklamnímu využití a má stanoven veřejný ceník internetové reklamy.

Sekce

Libovolná část média. Každá sekce smí být součástí jen jedné nadřazené sekce, resp. jen jednoho média. Libovolné 2 různé sekce na stejné úrovni jsou disjunktní.

Pack (nebo balíček)

Libovolná kombinace médií a sekcí. V případě, že jednotlivé části Packu patří více Provozovatelům, je pro zřízení Packu třeba souhlas všech dotčených Provozovatelů.

Definice složení Packů a požadavky na změny složení již existujících Packů je nutno Realizátorovi zaslat do 20. dne měřeného měsíce (pokud má být součástí packu sekce média, musí tato sekce již

existovat v OLA), jehož se změny ve struktuře Packu mají týkat. V případě, že se tak stane později, nemusí být možné tuto změnu aplikovat na strukturu dat již za tento měsíc.

Hlavním (primárně zveřejňovaným) žebříčkem je žebříček médií. V případě, že médium ve výzkumu definuje své veřejné sekce, jsou tyto spolu s příslušnými statistikami návštěvnosti součástí tohoto žebříčku. Vedle tohoto žebříčku jsou zveřejňovány rovněž reporty za Provozovatele a Packy.

Médium zařazené do výzkumu je vždy měřeno kompletní (všechny stránky daného média nesou měřící skript). Provozovatel sám určí jaká část měřeného média je měřena standardním způsobem, tj. kompletně ve shodě s touto metodikou, zejména pak v otázkách týkajících se emise dotazníků a zobrazování výsledků ve veřejných datech.

Zbývající část, tedy část nereklamní, je měřena nestandardním způsobem a je možné domluvit s realizátorem specifické podmínky měření. Například nezobrazování dotazníků, nezobrazování informací o této části serveru v rámci veřejných dat, případně poskytnutí dalších služeb. Sjednání podmínek zapojení nestandardně měřené části média je plně v kompetenci realizátora.

V případech, kdy část serveru, která je měřena standardním způsobem nedosahuje 90% míry zobrazování reklamního obsahu v rámci všech svých měřených stránek, si realizátor vyhrazuje právo zapojit tento server do projektu až po sjednání specifických podmínek.

Realizátor a zadavatel si vyhrazují z metodických důvodů právo korigovat rozhodnutí provozovatele o rozdělení média na standardně a nestandardně měřenou část.

Aby bylo možno pro médium (případně jeho sekce) za příslušný měřený měsíc v oficiálních měsíčních výstupech (viz kapitola 11) reportovat statistiky návštěvnosti a sociodemografický profil jeho uživatelů, musí toto médium být měřeno (mít nasazený měřící skripty na všech stránkách) po období celého měsíce. V opačném případě nebude v těchto výstupech médium reportováno a výsledky za toto médium neobdrží ani jeho Provozovatel.

3.2 Pravidla pro umístování měřících skriptů

Měřící skript je tvořen 2 částmi:

- hlavním skriptem (ve formě .js), který bude po uložení na server dosažitelný pro vlastní měřící kódy.

- měřicími kódy (sčítacími skripty), které je třeba vložit do html kódu všech stránek, které mají být měřeny.

Prvním krokem k umístění skriptů do stránek je přihlášení do účtu systému gemiusTraffic, ve kterém lze stáhnout hlavní skript a vygenerovat sčítací skripty. Sčítací skript má podobu:

```
<!-- (c) 2000-2007 Gemius SA ver pro 11.0 strona glowna -->
<script language="javascript1.2" type="text/javascript">
<!-- var gemius_identifier = new String('IDENTIFIER');
//-->
</script>
<script language="javascript1.2" src="/js/gemius.js"></script>
```

Provozovatel Média je umístuje na měřená Média Kód ve spolupráci s Realizátorem nejpozději 5 dní před začátkem výzkumu. Při umísťování měřících skriptů je třeba dodržovat následující pravidla:

1. Měřicí kód nesmí být umístěn ve stránkách, které se obnovují (refreshují) automaticky, tzn. bez akce uživatele. Rovněž nesmí být měřicí kód umístěn do stránek, které obsahují automatické přesměrování na jiné stránky. Provozovatel média se zavazuje informovat realizátora o existenci takovýchto stránek. (URL takovýchto stránek by neměla vstupovat do váženího procesu při použití SW panelu.)
2. Měřicí kód nesmí být umístěn ve stránkách (nových oknech), které se otevírají automaticky bez explicitního povolení (resp. nastavení uživatele). Provozovatel média se zavazuje informovat realizátora o existenci takovýchto stránek. (URL takovýchto stránek by neměla vstupovat do váženího procesu při použití SW panelu.)
3. Je-li stránka tvořena více framy (rámy) za použití např. elementu *iframe* nebo *frame*, může být měřicí kód umístěn pouze do jednoho z nich. Měřicí kód musí patřit médiu, které má doménu druhého řádu shodnou s doménou druhého řádu v adresním řádku prohlížeče.
4. Do html kódu jedné měřené stránky je povoleno umístit pouze jeden měřicí kód. Pokud realizátor neurčí jinak.
5. Není povoleno provádět jakékoli modifikace měřícího kódu.
6. Měřicí kódy není dovoleno vkládat dovnitř kódů sloužících pro zobrazování reklamního obsahu.

7. Provozovatel média se zavazuje informovat realizátora o stránkách obsahující java applety, flashové hry, chaty apod. V těchto případech je nutné využít speciálních metod umístování měřících skriptů do stránek založených na podobných technologiích.
8. V případě využití technologie AJAX pro zobrazování obsahu v rámci měřených stránek, kdy je obsah načítán asynchronně a není možné stránky vybavit měřícími skripty standardním způsobem, je provozovatel povinen informovat realizátora a využít speciálních metod pro umístění měřících skriptů. V takovémto případě je možné započítat zobrazení PV (PageView) pouze v takovém případě, kdy po akci uživatele, typicky po kliknutí myší, je načten a zobrazen velikostně významný obsah.
9. Všechny stránky příslušného média musí mít vložen měřící kód, včetně předdefinovaných chybových stránek. Pokud provozovatel média nechce umístit měřící kód na všechny běžně přístupné stránky a jedná se o větší rozsah stránek, nebudou pro dané médium jako celek ani pro jeho případné sekce stanoveny sociodemografické statistiky. Provozovatel média se zavazuje informovat realizátora o stránkách, které nejsou vybaveny měřícím kódem. (URL takovýchto stránek by neměla vstupovat do vážícího procesu při použití SW panelu.)
10. V případě, že stránka s měřícím kódem není stránkou plnohodnotnou, je její provozovatel povinen vyvinout maximální úsilí, aby byl možný přístup na tuto stránku výhradně a bezprostředně přes médium, kterému je návštěvnost této stránky započítávána. A to zejména v případech jde-li o spolupráci jednotlivých provozovatelů. Plnohodnotnou stránkou se rozumí taková stránka, jejíž okno obsahuje adresní řádek (URL bar) s kompletním URL dané stránky. Pouze v případě, že nově otevřené okno má vždy stejnou doménu druhého řádu jako okno, ze kterého bylo otevřeno, nemusí toto nově otevřené okno obsahovat adresní řádek. Pokud provozovatel chce mít okno stránky se skrytým adresním řádkem, a toto okno je přístupné z jiné domény druhého řádu než je doména tohoto nově otevíraného okna, nemůže být ve zdrojovém kódu nově otevíraného okna umístěn měřící kód. Zároveň se provozovatel zavazuje informovat realizátora o takových stránkách.

Pro každé Médium je vytvořen zvláštní účet v gemiusTraffic pro účely sběru dat o návštěvnosti tohoto Média. Je zakázáno umístit měřící

skripty z tohoto účtu do stránek, jejichž doména druhého řádu se liší od domény druhého řádu Média.

Realizátor je povinen zajistit, že kódy umístěné do Médii Provozovatele Média v souladu popsaným v žádném případě neovlivní dostupnost a funkčnost těchto Médii.

Média porušující uvedená pravidla mohou být z Výsledků okamžitě vyloučena, podmínky pro opětovné zařazení stanoví Zadavatel po projednání s Realizátorem.

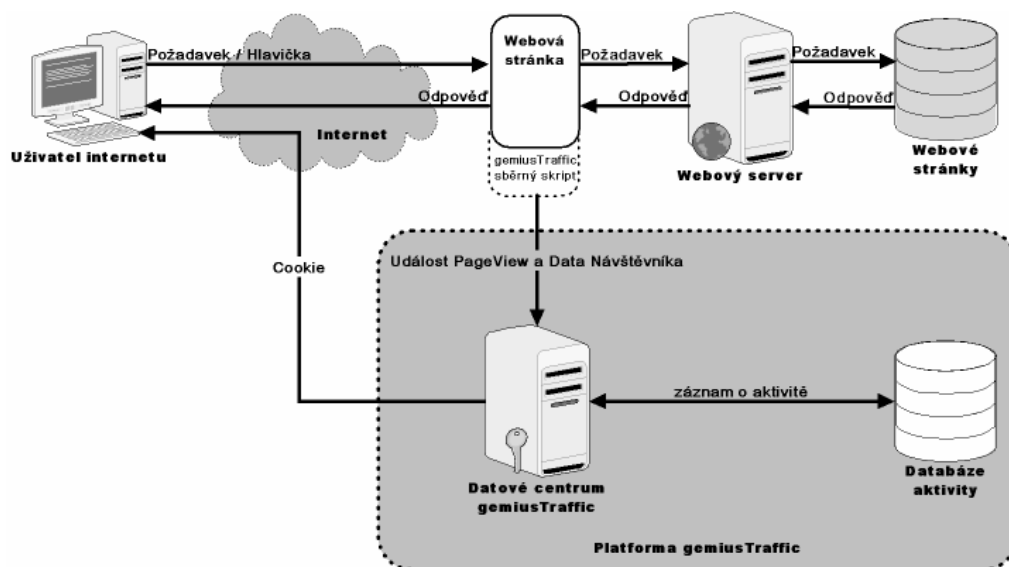
4 ZÁKLADNÍ POPIS TECHNOLOGIE MĚŘENÍ NA STRANĚ WEBOVÉ STRÁNKY (SITE-CENTRIC)

4.1 Technologie gemiusTraffic

Prvním výzkumným nástrojem používaným v realizované metodice pro sběr dat je proprietární technologie gemiusTraffic založená na měření přístupů k webovým stránkám.

Systém gemiusTraffic je nainstalován na straně serveru. Sbírá kompletní data o užití webových stránek (aktivně zapojených do výzkumu). Eviduje všechna zobrazení všech návštěvníků dané stránky. Nejedná se tedy o žádný vzorek internetových uživatelů, ale o všechny uživatele, a proto zde nedochází k žádné statistické chybě.

Systém umožňuje měřit webové stránky nezávisle na použité technologii (statické HTML, dynamicky generované HTML (PHP, ASP, Java, Perl a další), včetně Adobe Flash).



obrázek 1 Architektura gemiusTraffic

V rozhraní gemiusTraffic (viz kapitola 7.1) mají vlastníci serverů možnost definovat logickou stromovou strukturu měření svých stránek, a následně sledovat v reálném čase statistiky pro jednotlivé nadefinované uzly (stránky či sekce vytvořené struktury).

Pro jednotlivé měřené stránky je vygenerován javascriptový kód - **sběrný skript**. Ten je vložen do HTML kódu příslušné stránky, kterou si vlastník serveru přeje monitorovat.

System gemiusTraffic užívá javascript, protože tím automaticky na rozdíl od technologie „blank-pixelů“:

- Minimalizuje započítávání provozu způsobeného indexovacími roboty a dalšími automatickými skripty do reportovaných statistik.
- Eliminuje vliv cachování na započítávání provozu tak, že dynamicky modifikuje URL použité pro kontakt se sběrným serverem.
- Umožňuje sledovat technické parametry počítačů uživatelů včetně použitého operačního systému, webového prohlížeče, jazykové verze, rozlišení obrazovky, hloubky barev i verzí plug-inů instalovaných do internetového prohlížeče jako např. Flash plug-in.
- Integruje se s dalšími složkami metodiky pro měření návštěvnosti a sociodemografie (viz dále).

4.2 Hit-kolektor zajišťující sběr dat

Hit-kolektor je jednoduché zařízení (hardware+software), které zajišťuje sběr dat ve výzkumu. Hit-kolektor registruje „hity/zásahy“ (počet vykonání kódu sběrného skriptu). Dále podle získaných dat rozhoduje, zda příslušnému návštěvníku internetové stránky zobrazí rekruční dotazník (viz dále metodika Pop-up Panelu).

Pro účely projektu jsou primárně použity dva hit-kolektory (hlavní a záložní), umístěné na páteřní síti s dostatečnou propustností. Oba hit-kolektory jsou ve vlastnictví realizátora projektu, který je zodpovědný zejména za jejich technickou funkčnost, bezpečnost a údržbu.

5 ZÁKLADNÍ POPIS TECHNOLOGIE MĚŘENÍ NA STRANĚ INTERNETOVÉHO PROHLÍZEČE (USER-CENTRIC)

5.1 Aplikace netSoftware

Druhým výzkumným nástrojem používaným v realizované metodice pro sběr dat je měřicí software netSoftware – lokální aplikace registrující adresy prohlížených stránek a čas strávený jejich prohlížením. Technologie je použita pro realizaci panelového výzkumu SW panel.

Aplikace netSoftware je nainstalována na počítači internetového uživatele doma nebo na pracovišti. Program netSoftware sbírá kompletní data o prohlížení všech (aktivně i pasivně zapojených) webových stránek na vzorku uživatelů internetu (tj. mezi členy panelu).

Vzhledem ke skutečnosti, že mnohé počítače mohou být používány větším počtem jednotlivců, monitorovací software identifikuje jednotlivce, kteří v daném okamžiku počítač používají, a rozlišuje data generovaná těmito jednotlivými uživateli.

Monitorovací software sbírá data o návštěvách webových stránek členy panelu a využívá k tomu události BHO v Internet Exploreru, resp. ekvivalent těchto událostí v prohlížeči Mozilla Firefox. Konkrétně, tyto události umožňují definovat začátek stahování webové stránky (i když se sestává z několika komponent HTML) a konec tohoto stahování. Prostřednictvím těchto událostí je možné identifikovat zobrazení stránky (pageview), čímž se rozumí zobrazení celého HTML dokumentu, který je prohlížen v okně internetového prohlížeče. Zobrazení stránky se nezapočítává, pokud je stažena jen část HTML dokumentu (např. jeden rámec z celého dokumentu). Monitorovací software registruje veškeré události typu BHO, které jsou pak interpretovány a zpracovávány na zobrazení stránek generované danými členy panelu na webových stránkách, které tento člen panelu navštíví. Registrovaná data jsou uložena na lokální diskovou jednotku člena panelu (jsou komprimována a následně zašifrována) a poté jsou přenesena na servery realizátora. Tento proces byl navržen tak, aby využíval HTTP protokol (metoda GET). Tímto je zabezpečena dostupnost spojení s datovým centrem (pokud uživateli funguje spojení s webovými servery poskytovatelů obsahu, zároveň funguje i spojení se servery realizátora).

5.1.1 Sběr dat (logy BHO)

Data naměřená aplikací netSoftware jsou ukládána ve formátu logů BHO. Základní logy BHO, kde je zaznamenávána aktivita člena panelu, reprezentují sérii zobrazení (page view) a pro každé z nich je dáno následující:

1. čas zobrazení stránky podle GMT (časový údaj)
2. číslo člena panelu
3. URL (v textovém formátu)
4. další parametry (události BHO)

5.1.2 Modifikace HTTP hlavičky User-Agent

Kromě výše uvedeného způsobu sběru dat aplikací netSoftware lze na SW panelu využít i alternativní postup při získávání dat o navštěvovaných stránkách. netSoftware modifikuje HTTP hlavičku User-Agent všech požadavků vykonaných prohlížečem Mozilla Firefox, Opera a Internet Explorer vč. aplikací využívajících jádro IE bez uživatelské modifikace hlavičky User-Agent. Hlavička se modifikuje přidáním jednoznačné identifikace respondenta SW panelu.

Takto modifikovaná hlavička se dostane na hit-kolektor. Díky ní mohou být při zpracování efektivně odlišeny všechna zobrazení vykonaná uvnitř uvedených prohlížečů a aplikací. Toho se využije pro získání informace o navštívených URL a následnou produkci Výsledků v případě prohlížeče Opera plošně (pro všechna zobrazení vykonaná respondenty SW panelu), a volitelně na základě předešlého požadavku Provozovatele, který je schválen Zadavatelem a Realizátorem, v případě PV specifických Médii či jejich Sekcí, na nichž není měření prostřednictvím BHO zpráv (resp. jejich ekvivalentu v Mozilla Firefox) dostatečné. *Příkladem takové sekce může být Welcome screen zobrazovaná v klientu ICQ.*

6 VÝZKUM NÁVŠTĚVNOSTI

Měření návštěvnosti metodou **site-centric** se provádí pouze pro ty servery, které se aktivně účastní výzkumu, a tedy jsou vybaveny měřicími skripty systému gemiusTraffic. Současně je prováděn sběr dat o návštěvnosti metodou **user-centric** (pro všechny měřené servery - aktivně i pasivně zapojené).

Údaje o návštěvnosti získané z výzkumu se vyhodnocují ve dvou základních režimech:

- Návštěvnost založená na přímém použití **všech cookies** (měřená site-centric technologií)
 - Ukazatele jsou reportovány v reálném čase
 - Není možné konstruovat sociodemografický profil návštěvníků
- Návštěvnost **reálných uživatelů** (měřená site-centric i user-centric technologií)
 - Finální počet reálných uživatelů je získán systémem zpracování. Algoritmy zpracování eliminují nepřesnosti, které vznikají mazáním cookies, používáním jednoho počítače více uživateli resp. používáním více počítačů jedním uživatelem.
 - Z povahy aplikovaného systému zpracování plyne, že tato data jsou k dispozici se zpožděním.
 - Při výpočtu reálných uživatelů je možné konstruovat i jejich sociodemografický profil.

6.1 Algoritmus pro reportování reálného uživatele (site-centric)

Ve výzkumu návštěvnosti všech návštěvníků internetu se často používá identifikace návštěvníka pomocí jednoznačného identifikátoru „cookie“. Bohužel často nastávají situace, kdy je identifikace pomocí cookies nemožná (uživatelé počítačů nepovolí ukládání cookies na svůj počítač, cookies na svém počítači pravidelně či nepravidelně mažou, reinstalují systém a tím cookie odstraní či např. správce lokální sítě nastaví politiku mazání cookies pro všechny uživatele nebo je možné cookies modifikovat na firewallu a tím zabránit jejich ukládání). Občas je identifikace pomocí „cookies“ velmi nedokonalá (jeden počítač používá více uživatelů či jeden uživatel používá více počítačů nebo více prohlížečů na jednom počítači).

Z toho vyplývá, že počet cookies zaznamenaných na měřeném serveru při měření návštěvnosti není roven počtu skutečných uživatelů internetu. Není tedy možné věrohodným způsobem stanovit reach měřených serverů pouze na základě počtu cookies. Proto je v rámci výzkumu NetMonitor implementována metodika, která na základě informací o chování cookies stanovuje počet Reálných Uživatelů (RU), tedy odhad počtu skutečných lidí. Princip metodiky vychází z požadavku, že nechceme informaci o návštěvnosti „cookies“, ale o návštěvnosti „reálných osob“.

Algoritmus využívá informace o chování cookies, které jejich „vlastníci“ (uživatel počítače) nemažou, o chování cookies, které jsou mazány, a dále informace o celé internetové populaci. Využití těchto údajů umožňuje poměrně přesně určit počty reálných osob – návštěvníků jednotlivých serverů.

Algoritmus pro reportování reálného uživatele pracuje v několika krocích:

- 1) Výpočet počtu cookies (tzv. estimated cookies), které by systém naměřil, kdyby neexistoval problém mazání cookies.

Systém analyzuje vztah počtu všech cookies a počtu všech zobrazení vůči počtu tzv. validních cookies (označované rovněž jako good cookies) a počtu jimi generovaných zobrazení (tzv. validních zobrazení). Za validní cookies jsou považovány takové cookies, u nichž systém zaznamenal aktivitu před měřeným měsícem, v něm a rovněž do 15 dnů po jeho skončení. Z výše uvedených charakteristik lze vypočítat počet cookies, které by systém naměřil, kdyby bylo možné pomocí cookies dlouhodobě a jednoznačně označit všechny počítače připojící se k internetu.

$EC = \min [aC; gC * (aPV / gPV)]$, kde:

EC – Estimated Cookies

aC – všechny naměřené cookies

gC – počet nsmazaných (validních) cookies

gPV – zobrazení generovaná nesmazanými (validními) cookies

aPV – všechna zobrazení generovaná všemi cookies

Funkce minima se používá pro zamezení situace, kdy by počet EC překročil všechny naměřené cookies.

Počet EC lze stanovit jak pro celý internet (všechny aktivně měřené servery dohromady), tak i pro každý server či jeho sekci jednotlivě.

2) Stanovení ukazatele Reach pro jednotlivé servery a sekce:

$$R(s) = EC(s) / EC(i), \text{ kde}$$

R(s) – Reach serveru (sekce)

EC(s) – počet Estimated Cookies serveru (sekce)

EC(i) – počet Estimated Cookies internetu

3) Aplikace velikosti internetové populace (Univerzum) na spočtený Reach serveru (sekce) podle bodu 2). Zde je nutno zmínit fakt, že ve výzkumu nejsou zahrnuty všechny internetové servery, avšak pouze ty aktivně zapojené, které projevily o účast ve výzkumu zájem. Tato skutečnost má vliv na to, že část české internetové populace nenavštíví ani jednou v daném měsíci jakýkoli z aktivně měřených serverů (navštíví pouze servery aktivně neměřené). Velikost universa o tuto skupinu lidí tedy ponížíme – aktuálně o 2,3% - říkáme, že Reach NetMonitoru je 97,7%. Tento jev je průběžně monitorován, dojde-li k jeho hmatatelnému vychýlení je tento parametr modifikován.

$$RU(s) = R(s) * U * \text{Reach(NM)}, \text{ kde}$$

RU(s) – počet Reálných uživatelů pro daný server (sekci)

R(s) – Reach serveru (sekce)

U – velikost internetové populace z externího výzkumu

Reach(NM) - Reach NetMonitoru

Výpočet denních a týdenních RU probíhá dle obdobného mechanismu jako výpočet měsíčních hodnot. Denní a týdenní EC pro jednotlivé servery se počítají podle výše uvedeného vzorce, Reach pro jednotlivé servery či sekce je stanovován z denních a týdenních EC a je vztažen k měsíční internetové populaci.

7 VÝZKUM SOCIODEMOGRAFIE

Výzkum sociodemografického profilu návštěvníků aktivně měřených serverů a jejich sekcí je v NetMonitoru realizován na panelu respondentů. Do srpna 2009 byla tato část výzkumu realizována metodikou Pop-up Panel. Od září 2009 využívá sociodemografický výzkum model, ve kterém je využívána kombinace SW a Pop-up panelu. Součástí nového panelu je velká část SW panelistů, která je doplněna o vybrané Pop-up panelisty s cílem maximalizovat velikost nového panelu a zaručit reprezentativitu na internetovou populaci ČR. Takový kombinovaný panel je nazýván Joint Panel (JP).

Pro konstrukci panelu je využívána znalost struktury internetové populace z KV ATO (Kontinuální výzkum realizovaný pro účely Výzkumu sledovanosti televize pro ATO) a informace o chování internetové populace a o znacích toto chování ovlivňujících z výzkumu SEKV (viz následující kapitola).

7.1 SEKV

Strukturální Exkluzivní Kontinuální Výzkum (SEKV) je reprezentativní šetření realizované pro SPIR společností MEDIARESEARCH, ve kterém jsou dotazovány internetové domácnosti (tj. domácnosti, kde má alespoň jeden z členů starších 10 let přístup na internet alespoň jednou měsíčně), jejichž struktura je reprezentativní vzhledem k oficiálním údajům o internetových domácnostech získaných z Kontinuálního výzkumu TV-metrového výzkumu ATO (KV ATO)

V prvním roce výzkumu je dotazováno ve dvou vlnách 5050 domácností, v dalších letech pak 2550 domácností ve třech vlnách.

Cílem SEKV je popsat strukturu internetových domácností, jejich vybavenost a internetové chování jejich členů spolu s hlavními faktory ovlivňujícími toto internetové chování. SEKV je zároveň základním podkladem pro stanovení ukazatelů, které musí panel splňovat k tomu, aby byl reprezentativní. Pro stanovení těchto ukazatelů jsou použity specializované statistické postupy, mimo jiné CH.A.I.D. analýza, případně jiná obdobná statistická metoda. Výstupem je stanovení kontrolních znaků (základních charakteristik panelu, které nejvíce ovlivňují internetové chování populace) internetového panelu a jejich norem. Kontrolní znaky se mohou změnit maximálně jednou ročně, na základě CH.A.I.D. analýzy prováděné ročně na datech SEKV.

V případě zařazení nové proměnné mezi kontrolní znaky je téměř jisté, že její hodnoty nebudou v ten moment pro jednotlivé panelisty známy. Tato nová proměnná bude zařazena do rekrutačního dotazníku a u již existujících panelistů dojde ke zpětnému dotazování na tuto proměnnou. Vzhledem k problematice malé návratnosti při zpětném dotazování lze seznam kontrolních znaků panelu dočasně přehodnotit. Pokud nebude návratnost dodatečného dotazování dostatečná, bude nutné provést revizi dodatečných otázek a tím i množiny kontrolních znaků. Po získání dostatečně velké databáze panelistů, kteří vyplní dodatečnou otázku, bude možné zařadit danou charakteristiku mezi kontrolní znaky.

7.2 Pop-up Panel

Realizovaný systém používá běžnou metodu získávání informací o sociodemografii návštěvníků webové stránky založenou na zobrazování pop-up dotazníků náhodnému vzorku návštěvníků webové stránky. Nevýhodou Pop-up panelu je jeho výrazně nižší stabilita (větší obměna mezi měsíci, zejména proto, že členství v panelu je podmíněno platností původní cookie z doby registrace) a možnost použití jen pro účely výzkumu aktivně zapojených serverů. Výhodou je pak jeho velikost, která může velikost SW panelu několikanásobně přesáhnout.

Systém používá jedinečné cookies pro každého uživatele internetu, který navštíví webový server monitorovaný prostřednictvím platformy gemiusTraffic. Jestliže uživatel vyplní pop-up dotazník zjišťující jeho demografické charakteristiky na jednom libovolném serveru, pak je jeho profil spárován s obdrženou cookie, a tento profil je použit pro všechny návštěvy všech serverů aktivně se účastnících projektu.

Demografické profily jsou tak sdíleny celým systémem a uživatelé nejsou zbytečně obtěžováni totožnými pop-up dotazníky na různých serverech. To zároveň umožňuje výpočty překryvů mezi jednotlivými servery.

Měřený server musí mít na všech běžně přístupných stránkách umístěn měřící skript platformy gemiusTraffic, který zároveň slouží pro emitaci dotazníků. Pokud toto pravidlo není dodrženo, není možné generovat sociodemografický profil daného serveru jako celku, respektive jeho sekcí. Pokud by server mohl selektivně na některých svých částech dotazníky neemitovat, mohl by tak ovlivňovat svůj výsledný sociodemografický profil.

V Pop-up panelu se používají následující pravidla pro validaci:

- Panelová validace – respondent musí být členem panelu po celé měřené období (měsíc).
- Demografická validace dat pro detekci “chybných” odpovědí z pop-up dotazníků, které by měly být vyřazeny z výsledných datových souborů (*např. uživatelé deklarují užívání internetu v práci, ale jsou mladší 15 let*).
- Behaviorální validace pro detekci netypického chování, které lze považovat za činnost robotů a nikoli za skutečnou aktivitu uživatelů internetu.

Vzhledem k omezením v datech ze strukturálního výzkumu nezahrnuje Pop-up panel osoby mladší 10 let.

7.3 SW Panel

SW panel je výzkumem založeným na měření internetové aktivity na straně uživatele (resp. jeho prohlížeče). Výhodou SW panelu je jeho vyšší stabilita umožňující lepší a náročnější analýzy na výsledných datech a rovněž jeho schopnost poskytnout data i o serverech, jež nejsou v NetMonitoru aktivně měřeny pomocí skriptů gemiusTraffic.

Členové SW panelu jsou rekrutováni obdobným způsobem jako členové Pop-up Panelu. Rekrutační dotazníky jsou zobrazovány náhodně vybraným uživatelům (cookie), kteří navštíví rekrutační místa (stránky, které se účastní měření site-centric a jsou vybaveny měřicími skripty systému gemiusTraffic). Po vyplnění rekrutačního dotazníku si respondenti volitelně nainstalují na své počítače monitorovací software (aplikaci netSoftware), který monitoruje adresy (podrobněji viz kapitola 5) jimi navštívených webových stránek. Pokud si respondent měřicí aplikaci po vyplnění dotazníku nenainstaluje, stává se pouze členem Pop-up Panelu. Tj. rekrutace pro Pop-up Panel probíhá pouze v rámci rekrutace do SW panelu.

Metoda pro sběr dat od členů panelu prostřednictvím monitorovacího softwaru umožňuje provádět měření pouze, pokud člen panelu:

- ✓ používá internet prostřednictvím počítače (a ne prostřednictvím TV set-top boxu, PDA nebo mobilního telefonu) a
- ✓ používá počítač s operačním systémem rodiny Microsoft Windows (Windows 95 / 98 / Me/ NT / 2000 / XP / Vista).

Počet internetových uživatelů, kteří nesplňují tato kritéria, je nízký, tj. toto omezení není podstatné při odhadování výsledků vzhledem k celé internetové populaci, a je průběžně monitorován společně

s technologickými a metodickými možnostmi implementace měření. Podrobné údaje na toto téma lze na vyžádání poskytnout.

Vzhledem k právním omezením, omezením při rekrutaci a způsobu využívání internetu (např. velký počet osob s různou demografickou charakteristikou používá stejný počítač - např. v internetové kavárně) výzkum nezahrnuje následující:

- ✓ osoby připojující se méně často než jednou za 30 dní,
- ✓ osoby připojující se z jiných míst než domov nebo pracoviště (monitorovací software není instalován např. ve školách, univerzitních laboratořích, internetových kavárnách, veřejných terminálech apod.). Potenciální člen panelu zde není rekrutován, protože politika mazání cookies na takovýchto místech prakticky vylučuje, aby internetový uživatel obdržel rekrutační dotazník.

Vzhledem k omezením v datech ze strukturálního výzkumu nezahrnuje SW panel osoby mladší 10 let.

7.4 Rekrutace respondentů

Rekrutace respondentů probíhá pomocí dotazníku, který je emitován náhodně vybraným návštěvníkům měřených serverů pomocí měřících skriptů. Vyplněním rekrutačního dotazníku se respondent stává členem tzv Pop-up Panelu a je mu zobrazena výzva k instalaci měřící aplikace netSoftware. Přistoupí-li respondent na tuto výzvu a aplikaci si nainstaluje, stává se členem SW Panelu.

Realizátor dostává na pravidelné bázi od subdodavatele, který realizuje online rekrutace, zprávy o aktuálním stavu těchto online rekrutací a aktuálním stavu stávajících panelistů (včetně charakteristik nových a odcházejících panelistů), a to jak pro Pop-up tak pro SW Panel. Na základě těchto zpráv je možné operativně rozhodnout o případném okamžitém zásahu do panelu či rekrutací.

7.5 Komunikace s panelem a odměny

Se členy SW Panelu je komunikováno tak, aby byla obměna mezi členy panelu byla rovnoměrná, aby byli aktivováni neaktivní členové panelu a aby byla členům panelu poskytnuta potřebná technická podpora. Komunikace probíhá prostřednictvím e-mailu či telefonicky.

E-maily jsou zasílány těm členům panelu, kteří zadali svoji e-mailovou adresu a souhlasili (zvolili možnost v dotazníku) s příjmem e-mailové korespondence. Členové panelu jsou rozděleni do různých

skupin dle stupně jejich angažovanosti na výzkumu a podle potřebné technické podpory. Komunikace s každou skupinou je přizpůsobena konkrétním potřebám a charakteristice skupiny.

Pro členy panelu je navíc organizována soutěž, v níž mají možnost vyhrát hmotné ceny. Hlavním cílem soutěže je podpořit zájem o spolupráci v SW Panelu a motivovat členy panelu k dlouhodobé spolupráci.

Za rok proběhne 5 kol soutěže, v každém z nich:

- Je vylosováno 5-7 soutěžících,, a ti
- získají ceny v souhrnné hodnotě 20-30 000 Kč.

Soutěžícím se může stát každý člen SW Panelu, který spolupracuje na výzkumu (má nainstalovanou a aktivní aplikaci netSoftware) déle než dva měsíce pro dvouměsíční část a déle než jeden rok pro roční část soutěže. Účastník soutěže musí vyplnit a odeslat přihlašovací formulář, který je mu zobrazen prostřednictvím měřící aplikace.

Následně je z množiny soutěžících náhodným procesem vybráno n respondentů (n odpovídá počtu rozdělovaných cen), kteří jsou telefonicky (telefonický kontakt respondent uvedl v přihlašovacím formuláři) kontaktováni a je jim položena soutěžní otázka. V případě správné odpovědi se kontaktovaná osoba stává výhercem jedné z možných cen. Pokud po oslovení všech n respondentů nejsou rozděleny všechny ceny, vybírá se zbývajících m respondentů a proces se opakuje až do rozdělení všech n cen.

7.6 Joint Panel

Vlastní SW panel může dle zkušeností Realizátora za současných podmínek rekrutace nabýt hrubé velikosti (neohlížíme se na splněnost norem kontrolních znaků a následné reprezentativity) cca 11 – 13 000 respondentů. Při této hrubé velikosti může být problematické dosažení naplnění všech nebo alespoň většiny potřebných norem pro zajištění reprezentativity panelu (viz kapitola 7.7). Podmnožina SW panelu, která těmto normám vyhovuje, podle striktních pravidel může být výrazně menší než cílová velikost panelu (10 000 respondentů). S klesající velikostí panelu a s klesající velikostí (z hlediska návštěvnosti) serveru/sekce zároveň roste pravděpodobnost, že nebude možné tento server/sekci navázat na návštěvnost.

Z těchto důvodů se přistupuje k výběru podmnožiny SW panelistů, která je následně doplněna o vybrané členy Pop-up panelu. Při výběru je kladen důraz na maximalizaci velikosti subpanelu jak u SW panelu tak u

Pop-up panelové části. Tento vzniklý reportující panel (Joint Panel) poté splňuje požadavky na naplnění norem i velikost. Ta se aktuálně u výsledného panelu pohybuje na úrovni 30 000 - 33 000 respondentů.

Minimální počet JP panelistů, pro které je možno konstruovat sociodemografický profil návštěvníků příslušného Média či sekce je 45.

7.6.1 Princip stanovení reportujícího panelu

Pro reportující panel je určen seznam panelistů, kteří mohou být zahrnuti do dat za reportovaný měsíc a to jak z pop-up panelu, tak ze softwarového panelu. Nejprve je potřeba vyřešit problematiku zdvojení informace pro členy panelu, kteří jsou zapojeni jak v pop-up panelu, tak i v softwarovém panelu. Vzhledem k tomu, že informace ze softwarového panelu je v tomto případě podrobnější (je popsána veškerá aktivita respondenta a nikoliv pouze aktivita na měřených serverech), je možné informaci z pop-up panelu zcela zanedbat bez ztráty jakékoliv informace. Respondenti vyskytující se v obou seznamech jsou proto z pop-up panelu vyloučeni.

Do reportujícího JP (máme na mysli panel, jež je použit pro konstrukci sociodemografického profilu návštěvníků Média) mohou být zařazeni SW panelisté, jejichž měřící aplikace byla nainstalována v měsíci předchozím a zároveň byl aplikací zaslán alespoň jeden datový log v měsíci produkčním (aktuálním). Následně musí platit, že měřící aplikace nebyla v průběhu celého měřeného měsíce odinstalována. Panelista pochopitelně musí projít validačními, sociodemografickými testy.

V dalším kroku je z redukovaného souboru možných panelistů proveden výběr panelistů do reportujícího panelu, který splňuje následující požadavky:

- vybraný reportující panel dosahuje maximální možné reprezentativity vzhledem k požadovaným kontrolním znakům (minimalizovány odchylky od norem)
- reportující panel obsahuje alespoň 10 000 softwarových panelistů
- zahrnutí softwaroví panelisté jsou doplněni pop-up panelisty, kteří splňují přísnější kritéria zaručující vyšší stabilitu celkově vzniklého reportujícího panelu. Konkrétně u pop-up panelistů požadujeme:

- panelista se připojil do pop-up panelu s cookie, která vznikla alespoň 3 měsíce před 15. dnem¹ měsíce následujícího po reportovaném měsíci (zvyšuje se pravděpodobnost, že bude členem panelu i další měsíc) – např. pro výběr panelu v měsíci prosinci uvažujeme pouze pop-up panelisty, jejichž cookie vznikla před 15. říjnem.
- v případě dostatečně velké základny pop-up panelistů lze přidat podmínku na existující emailový kontakt (umožní případnou komunikaci s panelistou).
- z obou panelů je vybrán maximální možný počet panelistů, který umožní splnění všech kontrolních znaků na reportujícím panelu v χ^2 (chí kvadrát) normě nebo jí podobné s dostatečnou významností. Obvykle je požadováno splnění norem na hladině významnosti 0.01.

K výběru je použit software, který využívá k dosažení maximálně možné reprezentativity iterační algoritmus. Software vybírá z dané množiny panelistů co nejreprezentativnější vzorek, který má minimální odchylku od požadovaných relativních četností.

7.6.2 Vlastnosti reportujícího panelu

Hlavní předností reportujícího panelu, který je získán výše popsáním postupem, je jeho vysoká reprezentativita v požadovaných kontrolních znacích a maximální velikost. Výsledkem je možnost kvalitnějšího navázení dat s lepšími vlastnostmi, s minimálním poklesem počtu uzlů (serverů a sekcí) navážených na návštěvnost. Tyto dva aspekty hrají klíčovou roli ve zkvalitnění odhadů sociodemografie na internetu.

Další nespornou výhodou je vyšší stabilita reportujícího panelu. Využitím softwarového panelu v kombinaci s „kvalitními“ pop-up panelisty získáváme panel, který stabilitou překonává v minulosti používaný pop-up panel.

Doplněním o pop-up panelisty zvyšujeme pravděpodobnost, že respondent, který vyplní v dotazníku využívání internetu v práci/ve škole

¹ Do tohoto dne jsou sbírána data panelistů k vyhodnocení návštěvnosti v reportovaném měsíci.

(kontrolní znak), v daném okamžiku reálně využívá pracovní/školní počítač. Je totiž pravděpodobnější, že v práci / ve škole respondent spolupracuje spíše formou pop-up panelu než instalací software. Máme tak větší možnost v návštěvnosti internetu zahrnout rovněž profil uživatelů v práci nebo ve škole. Rozdílnost internetového chování této skupiny panelistů od chování softwarových panelistů může být tímto faktem postihnuta.

7.7 Normy a reprezentativita

Kontrolní znaky panelu stanovené analýzou SEKV definují požadavky na vlastnosti cílového panelu. Jejich plnění je kontrolováno pomocí stanovených norem a umožňuje kontrolovanou reprezentativitu cílového panelu.

Norma panelu (tj. požadovaný počet jednotlivců v panelu) v každé kategorii jednotlivých kontrolních znaků se určí jako relativní četnost příslušné kategorie mezi internetovými uživateli x velikost panelu.

Odchyly panelu od norem jsou posuzovány pomocí χ^2 (chí kvadrát) testu dobré shody. Panel je konstruován tak, aby odchyly panelu od norem byly minimalizovány (tj. aby bylo dosaženo co nejlepšího výsledku vzhledem k možnostem ochoty v populaci). Odchyly panelu se určuje jednou měsíčně na základě aktuálního stavu panelu k datu prováděné kontroly. Normy panelu budou v případě potřeby aktualizovány na základě dat z Kontinuálního výzkumu ATO a ze SEKV. Report o stavu panelu je měsíčně předkládán po distribuci dat Realizátorem a obsahuje informaci o stavu obou panelů (SW a Pop-up) před výběrem a o vybraném panelu.

Plnění norem je kontrolováno jednou měsíčně, vždy k patnáctému dni daného měsíce. V rámci kontroly jsou stanoveny aktuální odchyly od norem, které jsou měsíčně reportovány. Cílem kontroly je rovněž vytvoření seznamu panelistů, kteří jsou pro aktuální měsíc zařazeni do reportujícího panelu. Tento seznam je vytvářen tak, aby vzniklý subpanel splňoval podmínku minimální velikosti reportujícího panelu a zároveň byly minimalizovány odchyly panelu od norem.

Následující tabulka uvádí souhrn kontrolních znaků vybraných v 1. fázi výzkumu SEKV (a potvrzený výsledky 2. fáze). Kromě samotných norem je také uveden zdroj, který bude sloužit k aktualizaci norem.

Zdroj pro normy	kontrolní znak
Kontinuální výzkum SEKV	Starám se o bezpečnost PC
	Instant messaging (ICQ, AOL,...)
	Přístup v práci / ve škole
	Čas věnovaný zahraničním serverům
Kontinuální výzkum TV metrového projektu	Pohlaví respondenta
	Věk respondenta
	Nejvyšší ukončené vzdělání
	Velikost sídla
	Kraj
	Rodinný stav
	Počet členů domácnosti
Ekonomicky aktivní	

V průběhu roku 2009 byly přidány 3 nové kontrolní znaky ze SEKV. U těchto znaků došlo ke zpětnému dotazování (viz kap. 7.1) a byly zařazeny mezi hlídané v momentě, kdy byl k dispozici dostatečný počet panelistů s těmito vyplněnými znaky.

7.8 Vážení

Vážící znaky panelu jsou znaky, pro které chceme, aby jejich odchylky od skutečnosti byly co možná nejmenší. Toho nelze v praxi dosáhnout jinak než metodou dovažování získaných výsledků. Vážení zabezpečuje nevychýlenost výsledného odhadu návštěvnosti za předpokladu, že odhady v jednotlivých vážících třídách jsou již nevychýlené.

Z hlediska výzkumu se zajímáme o návštěvnost jednotlivců, tedy vážící znaky bereme vzhledem k jednotlivcům. Konkrétní seznam vážících proměnných byl stanoven po zpracování zaváděcího výzkumu a je aktualizován na základě výsledků kontinuálního výzkumu (viz kap. 7.1 a 7.7). V rámci Joint panelu do vážení vstupují kromě sociodemografických proměnných (typu pohlaví, věk, vzdělání,...) také behaviorální proměnné (navazování na celkové naměřené objemy návštěvnosti měřených stránek tak, jak ji zachytil systém měření návštěvnosti gemiusTraffic).

Vzhledem ke komplikovanějšímu strukturálně-behaviorálnímu vážení je požadováno, aby vlastní váhy přiřazené jednotlivcům v panelu splňovaly následující vlastnosti:

- tvar rozdělení vah by měl být soustředěný kolem hodnoty 1 (vzhledem k behaviorálnímu vážení však nelze zajistit doporučené rozdělení ve tvaru zvonu)
- součet všech vah je větší nebo roven počtu respondentů reportujícího panelu a
- váhy nesmí překročit hodnotu 10, případně nesmí u vybrané podmnožiny demografie překročit hodnotu 5.

V případě nemožnosti splnit tato kritéria je možné přistoupit k dalším krokům vedoucím ke zlepšení kvality vah:

- ze zpracování mohou být vyřazeny ty servery, které způsobují vysoké váhy
- může být proveden další výběr subpanelu tak, aby takto vzniklý subpanel splňoval podmínku minimální velikosti reportujícího panelu a zároveň byl řešen problém vysokých vah.

Vzhledem k velkému množství vážících proměnných (hlavně behaviorální) může nastat situace, kdy není možné určité servery dovážet s danou přesností (obvykle měřeno χ^2 statistikou na hladině 0.05 nebo 0.01). Servery, které není možné dovážet, jsou z vážení vyloučeny.

Tím, že je celý panel konstruován z hlediska kontrolních znaků, které zajišťují reprezentativitu, jsou zabezpečeny dobré vlastnosti výsledných vah a jejich snadnější numerická konstruovatelnost. Není statisticky nezbytné konstruovat normy panelu příliš těsně (plně postačující jsou odchylky v normě χ^2 statistiky), protože korektnost z hlediska jejich nevychýlení je zajištěna dovážením.

Vážení se provádí vzhledem k teoretickým relativním četnostem internetové populace, které jsou získávány z Kontinuálního výzkumu ATO. Dále do vážení vstupují charakteristiky, které byly reálně naměřeny na internetových serverech (behaviorální charakteristiky).

8 DOMÉNOVÝ FILTR, DEFINICE SEKCI, KATEGORIZACE

8.1 Doménový filtr

V souladu s odstavcem 3.2 tohoto dokumentu, je na sesbírání data o návštěvnosti jednotlivých Médii aplikován tzv. „**Doménový filtr**“ Jedná se o opatření, které z procesu zpracování veřejně přístupných a oficiálních dat vylučuje veškerý traffic, který pochází z jiných domén druhého řádu, než z domény „definované“. Definovanou doménou rozumíme takovou doménu druhého řádu, jejíž název se shoduje s doménou, na které obsah Média skutečně běží a pod kterou je Médium prezentováno ve výstupech.

Očištění doménovým filtrem tak podléhají data v online aplikaci na <http://online.netmonitor.cz>, reporty na www.netmonitor.cz a data do analytické aplikace gemiusExplorer (rovněž ASCII data pro užití v dalších SW). Doménový filtr nemá vliv na data v aplikaci gemiusTraffic.

Provozovatelům je Realizátorem na měsíční bázi zasílán report pro jejich jednotlivá média s informací o procentu trafficu pocházejícím z nejsilnějších jiných domén (těch, které odchytil filtr).

8.2 Definice sekci

Pro účely měření návštěvnosti je v NetMonitoru možno médium dělit na menší části (sekce). Sekce jsou vzhledem k použití aplikace netSoftware (jež sbírá informace o URL stránek navštívených panelisty) pro účely konstrukce sociodemografického profilu uživatelů serverů a sekci definovány pomocí tzv. URL masek. Takto definované sekce jsou pak součástí veřejných a oficiálních výstupů NetMonitoru.

URL maska je předpisem vymezujícím sekci jako skupinu stránek, jejichž URL obsahuje řetězec (řetězce) v masce definovaný. Sekci „sport“ na Médium server.cz lze tedy např. definovat následujícím způsobem:

Médium: server.cz; Sekce: sport; URL maska: server.cz/sport, sport.server.cz

Dle této masky bude sekce „sport“ tvořena všemi stránkami, jejichž URL obsahuje příslušný řetězec, tedy *např. i* cokoli.server.cz/sport/cokoli či sport.server.cz/cokoli.

Stromy sekcí definovaných v gemiusTraffic a sekcí definovaných URL maskami jsou naprosto nezávislé. Stromová struktura skriptů a sekcí v gemiusTraffic tak slouží výhradně pro vnitřní potřeby média a nemá žádný vliv na strukturu a hodnoty ukazatelů návštěvnosti sekcí ve veřejných a oficiálních výstupech.

Jednotlivé definované sekce je možno označit jako veřejné nebo neveřejné. Veřejné sekce a jejich data o návštěvnosti jsou ve výstupech přístupné všem uživatelům, odběratelům kompletních dat je přístupná i informace o sociodemografickém profilu návštěvníků příslušné sekce. Sekce neveřejné, resp. jejich statistiky návštěvnosti, jsou přístupné výhradně Provozovateli, a to prostřednictvím online aplikace na <http://online.netmonitor.cz> po přihlášení. Sociodemografická informace o neveřejné sekci není přístupná ani Provozovateli.

8.2.1 Webové rozhraní pro definici sekcí médií na základě URL masek

Webové rozhraní slouží k vytváření a úpravám definic sekcí médií na základě URL masek. Hlavní účel je poskytnout řešení, které bude možné použít pro aktualizaci OLA i produkci měsíčních dat NetMonitoru. Rozhraní běží na url <http://treecz.gemius.com>.

Identifikace klienta bude realizována systémem Login/Heslo využívajícím http protokol. Login bude odvozen od názvu měřeného média v souladu se zobrazením v OLA. Ve druhé fázi dojde k implementaci „loginu Provozovatele“, který umožní ovládání více účtů.

Pokud se provozovatel rozhodne, že si přeje médium rozdělit na podsekce, zadá tento požadavek společnosti Mediaresearch, která poté, co médium potvrdí nasazení měřících skriptů, zašle médiu přístupové údaje do webového rozhraní. K nově vytvořenému účtu bude automaticky přiřazena povolená doména a název účtu.

Uživatelé budou mít možnost definovat následující informace o každém uzlu stromu:

- Název uzlu

- URL definice pro každý uzel
- Status uzlu (privátní/veřejný)
- Typ uzlu "Ostatní" bude generován automaticky pro každé médium a bude obsahovat traffic, který nebyl přiřazen žádné sekci na základě definice. Jméno tohoto uzlu může být uživatelem změněno.
- Informace o propojení mezi historickým stromem (na základě gT skriptů) a nově vytvořeným/upraveným uzlem.

Doporučená omezení týkající se vytváření uzlů a jejich definic:

- Médium si může nadefinovat maximálně 200 uzlů.
- Uzel smí obsahovat maximálně 1000 definic.

Rozhraní poskytne mechanismy pro kontrolu a opravy definic stromu médií. Existují tři stupně kontroly a oprav:

- Ověření, zda nově vytvořená definice nekoliduje se zbytkem definic. V případě špatné definice, bude tento fakt uživateli oznámen s návrhem opravy.
- Reportovacím mechanismem bude ověřena správnost celého stromu, s informacemi o tom, co by mělo být ve stromu opraveno.
- Týdenní kontroly na straně realizátora. Chyby budou opraveny realizátorem.

Úprava definic je ve webovém rozhraní možná kdykoli s výjimkou času potřebného pro pravidelnou údržbu. Pro aktualizace budou použity zkontrolované a opravené definice, které byly uloženy do databáze do nedělní půlnoci. Proces aktualizace probíhá každé pondělí o půlnoci. Pokud pondělí nebude pracovním dnem, proběhne aktualizace o půlnoci nejbližšího následujícího pracovního dne.

Pro účely produkce dat budou použity zkontrolované a opravené definice, které byly uloženy do databáze do půlnoci posledního dne v měsíci, pro který mají být produkována data.

8.3 Kategorizace médií

V rámci výstupů NetMonitoru jsou jednotlivá média a některé jejich veřejné sekce řazeny do některé z kategorií dle svého obsahu a zaměření. Aktuální seznam kategorií spolu s pravidly pro zařazování Médíí a sekcí do těchto kategorií je přístupný na www.spir.cz

Provozovatel Realizátorovi zašle návrh ke kategorizaci média či sekce. Tento návrh by měl zaslat nejpozději 5. den měsíce, ve kterém si přeje

mít médium či sekci zakategorizovanou. Jedná-li se o sekci, sekce by v té době již měla existovat v OLA. Kategorizaci poté odsouhlasí kategorizační komise (KK). Nedojde-li mezi KK a Provozovatelem k dohodě, je za rozhodující považováno stanovisko KK SPIR.

9 MODULÁRNÍ ROZŠÍŘENÍ PROJEKTU

9.1 Měření streamovaného videa

Metodika měření streamovaného videa bude doplněna po schválení metodiky ze strany orgánů SPIR (Streamovací pracovní skupina, MK SPIR, a Výkonná rada).

9.2 Měření používání komunikátorů

Měření používání komunikátorů bylo rozhodnutím MK SPIR odročeno.

9.3 Měření přístupů z mobilů

Metodika měření přístupů z mobilů bude doplněna po schválení metodiky ze strany orgánů SPIR (Mobilní pracovní skupina, MK SPIR, a Výkonná rada).

9.4 Měření rychlosti připojení

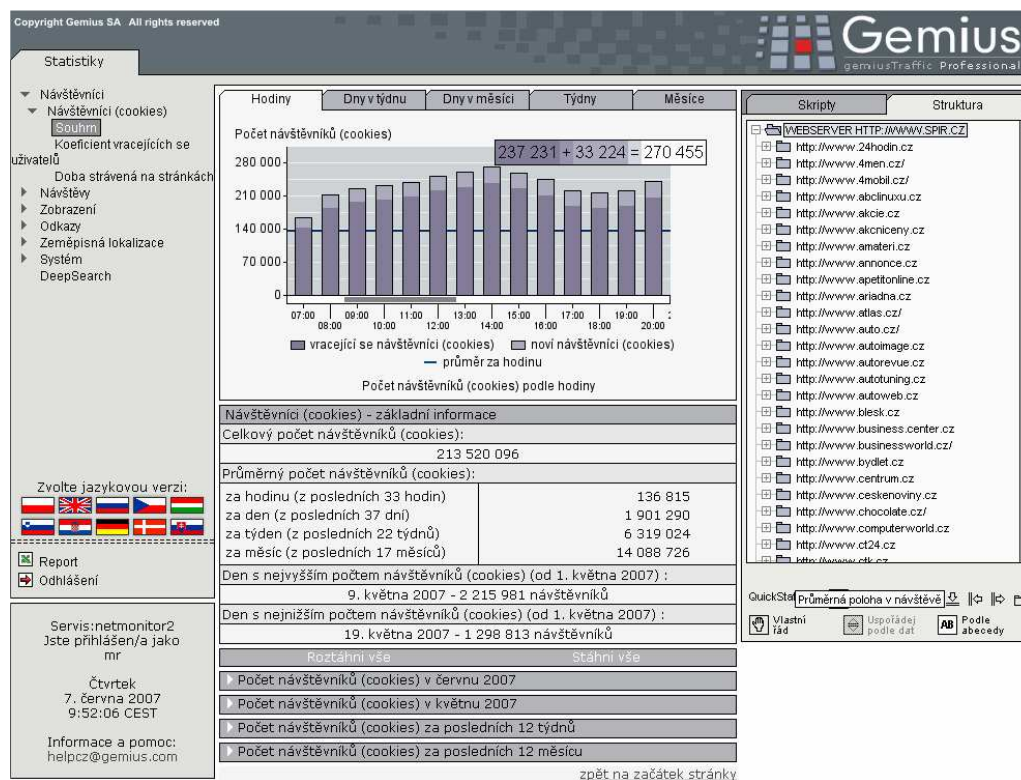
Rekrutační dotazník obsahuje možnost pro respondenta změřit rychlost svého připojení k internetu. Pokud ji respondent využije, je mu tato rychlost připojení změřena po zodpovězení všech otázek rekrutačního dotazníku a tento údaj je následně uložen do databáze spolu s dalšími informacemi o respondentu. Vlastní měření probíhá zasláním dat o velikosti 1MB do respondentova prohlížeče a současným měřením času které toto zaslání dat potrvá. Pokud dojde k přenosu celého objemu do 2,5 sekundy, dojde k následnému stažení dvou souborů o velikosti 3MB. Pokud ke stažení prvního 1MB souboru dojde v delším čase, je velikost dvou následujících stahovaných souborů rovna opět 1MB. Výsledné časy stahování všech tří souborů jsou do databáze zapsány zvlášť, uživateli je pak zobrazena jako výsledek průměrná hodnota. Respondenti, kteří se stali členy panelu v době, kdy toto měření rychlosti ještě nebylo součástí rekrutačního dotazníku, je systémem zobrazeno zvláštní pop-up okno s nabídkou změření rychlosti připojení, aby byl tento údaj zpětně doplněn i u těchto respondentů. Jakmile bude množství doplněných měření rychlosti na panelu dostatečné, bude tato proměnná zahrnuta mezi reportované charakteristiky.

10 ANALYTICKÉ NÁSTROJE

10.1 On-line rozhraní gemiusTraffic

Rozhraní systému gemiusTraffic umožňuje detailní analýzu chování návštěvníků internetových stránek, dává k dispozici nástroje na správu přístupu k datům a dále umožňuje logické členění stránek pro účely měření.

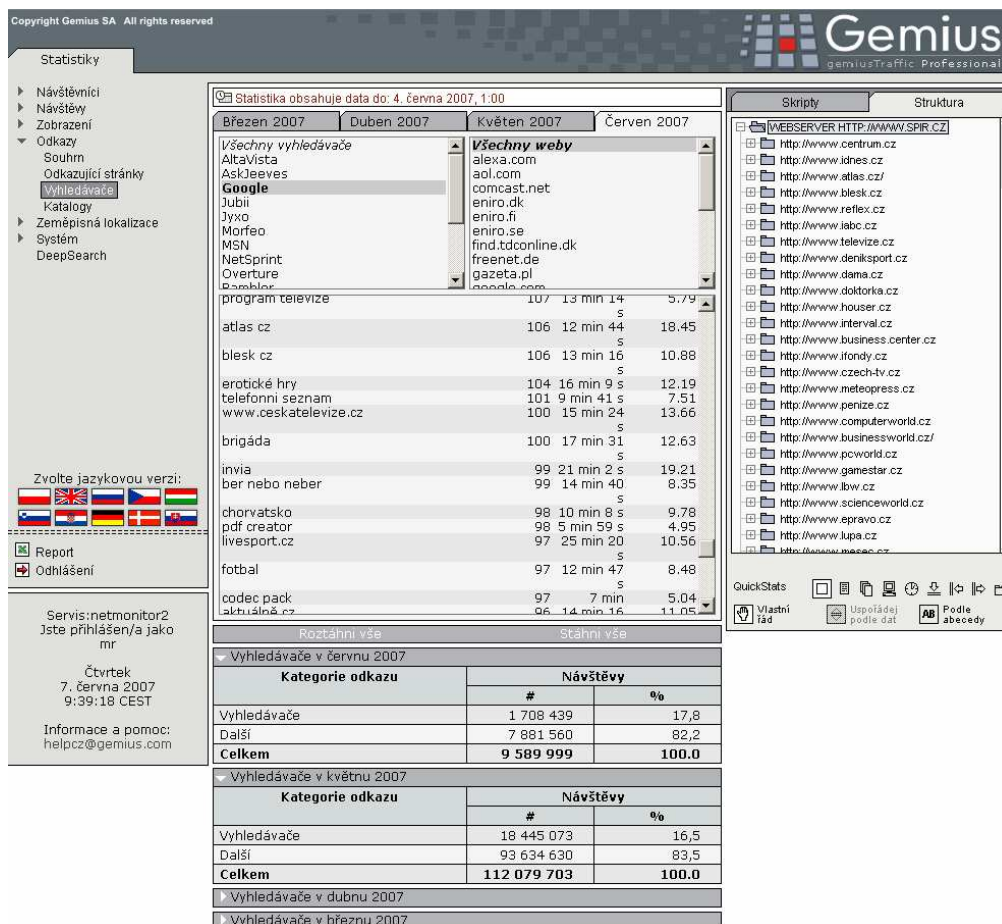
Rozhraní gemiusTraffic nabízí analýzu dat na bázi cookies v rozsahu statistik uvedených v kapitole 12.



obrázek 2 Ukázka rozhraní gemiusTraffic

Pomocí rozhraní gemiusTraffic mohou být statistiky návštěvnosti analyzovány až do úrovně hodin.

Rozhraní gemiusTraffic umožňuje analyzovat například i kompletní informaci o odkazujících stránkách. Identifikuje katalogy, vyhledávací stroje a servery a použitá klíčová slova.



Copyright Gemius SA All rights reserved

Statistika obsahuje data do: 4. června 2007, 1:00

Březen 2007 Duben 2007 Květen 2007 Červen 2007

Vyhledávač	Program televize	čas	průměr
Všechny vyhledávače		107	13 min 14 s
AltaVista			
AskJeeves			
Google		106	12 min 44 s
Jubii			
Jyx			
Morfeo			
MSN			
NetSprint			
Overture			
Search.nl			
program televize		107	13 min 14 s
atlas.cz		106	12 min 44 s
blesk.cz		106	13 min 16 s
erotické hry		104	16 min 9 s
telefonni seznam		101	9 min 41 s
www.ceskatelevize.cz		100	15 min 24 s
brigada		100	17 min 31 s
invia		99	21 min 2 s
ber nebo neber		99	14 min 40 s
chorvatsko		98	10 min 8 s
pdf creator		98	5 min 59 s
livesport.cz		97	25 min 20 s
fotbal		97	12 min 47 s
codec pack		97	7 min
aktualna.cz		96	14 min 16 s

Všechny weby

webové stránky	čas	průměr
alexa.com		
aol.com		
comcast.net		
eniro.dk		
eniro.fi		
eniro.se		
find.tdconline.dk		
freenet.de		
gazeta.pl		
google.com		

Vyberte jazykovou verzi:

Report

Odhlášení

Servis: netmonitor2
Jste přihlášen/a jako mr

Čtvrtek
7. června 2007
9:39:18 CEST

Informace a pomoc:
helpcz@gemius.com

QuickStats

Vlastní řád

Uspořádej podle dat

Podle abecedy

Kategorie odkazu	Návštěvy	
	#	%
Vyhledávače	1 708 439	17,8
Další	7 881 560	82,2
Celkem	9 589 999	100,0

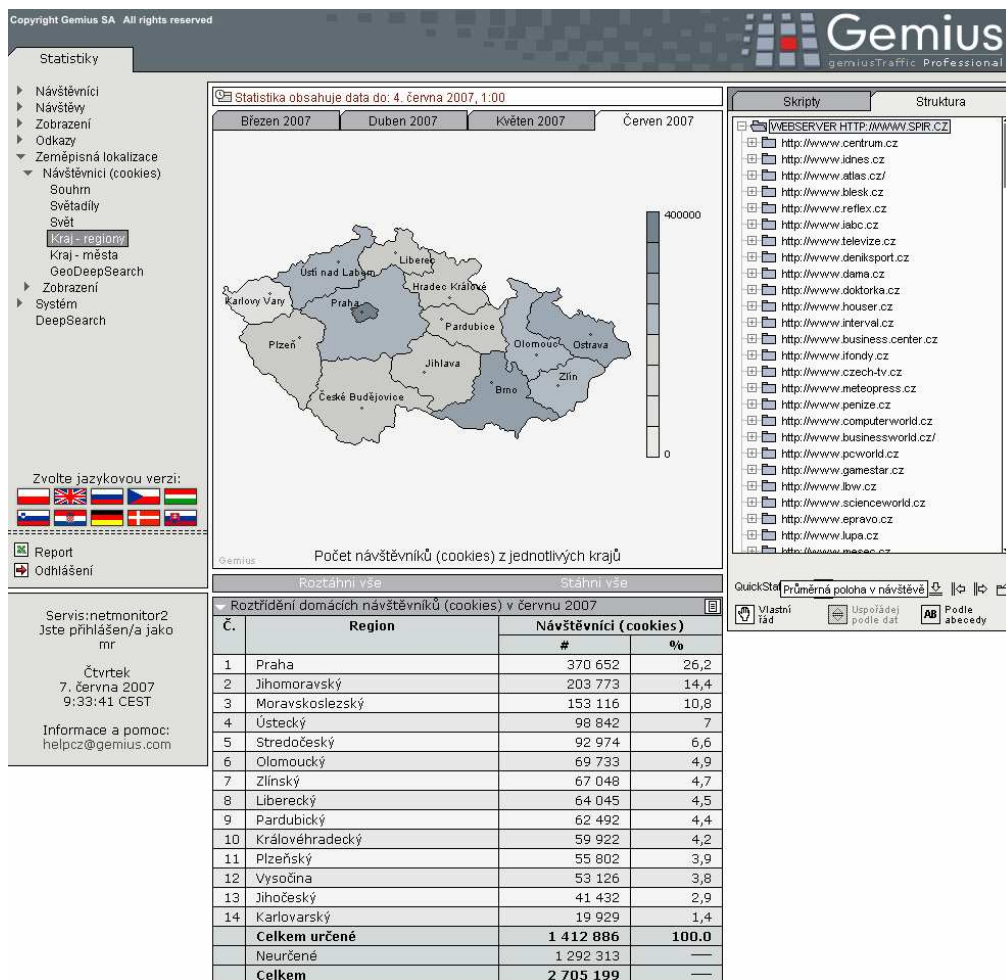
Kategorie odkazu	Návštěvy	
	#	%
Vyhledávače	18 445 073	16,5
Další	93 634 630	83,5
Celkem	112 079 703	100,0

Vyhledávače v dubnu 2007

Vyhledávače v březnu 2007

obrázek 3 Rozhraní gemiusTraffic - statistika vyhledávacích nástrojů

Analýza dat pomocí rozhraní přináší velmi přesné (až do úrovně měst) geografické rozložení návštěvníků stránek.



obrázek 4 Rozhraní gemiusTraffic – geografická lokalizace

10.2 On-line aplikace NetMonitor

V této aplikaci, která je veřejně přístupná prostřednictvím adresy <http://online.netmonitor.cz/>, jsou k dispozici vybrané on-line statistiky (počet návštěvníků (RUest, RU, příp. cookies), počet zobrazení (pageviews), počet návštěv a informace o stráveném čase a délce návštěvy). Na rozdíl od aplikace gemiusTraffic jsou v této aplikaci současně zobrazeny informace o všech měřených serverech (aktivně se účastnících výzkumu). Je tedy možné provádět základní porovnávání serverů mezi sebou.

Vzhledem k tomu, že v rámci sociodemografických dat je reportován ukazatel založený na počtech "Reálných uživatelů" **RU** je důležité i v rámci online dat reportovat ukazatel na stejné bázi. Vzhledem k metodice stanovení počtu **RU** ale není možné tento ukazatel reportovat bezprostředně po skončení daného měřeného období (dne, týdne, měsíce), tedy na on-line bázi. Z tohoto důvodu se v aplikaci přistupuje k výpočtu a zobrazování "Odhadů počtu reálných uživatelů" **RUest**, který je k dispozici ihned druhý den po ukončení daného období. Ukazatel **RUest** je počítán pro jednotlivé servery a sekce z naměřených cookies na těchto serverech či sekcích za použití koeficientu mezi naměřenými cookies a spočtenými RU v minulých obdobích. Měsíční RUest se tedy počítají jako:

$$RUest = Cm(n) / Qm(n-1), \text{ kde}$$

$Cm(n)$ – naměřené měsíční cookies na serveru (sekci) v měsíci n

$Qm(n)$ – koeficient (podíl) naměřených měsíčních cookies v měsíci $n-1$ a RU za měsíc $n-1$ z těchto cookies spočtených

Týdenní a denní RUest jsou počítána opět z naměřených cookies v těchto obdobích. Koeficient pro jejich výpočet je získáván jako průměrná hodnota poměrů týdenních (denních) spočtených RU a naměřených cookies na příslušném serveru (sekci) v měsíci $n-2$ (použije se vážený průměr vůči počtu cookies v jednotlivých obdobích). Měsíc $n-2$ je na rozdíl od výpočtu měsíčních RUest brán v úvahu proto, že se jedná o nejaktuálnější dostupnou informaci v momentě uveřejnění RUest.

Příklad: RUest za březen jsou dostupná již 1. dubna. V tento den jsou k dispozici již údaje o RU za únor (údaj uveřejněný 25. března) a tedy můžeme pro určení březnových RUest použít koeficient stanovený na základě únorových dat (měsíc $n-1$).

Pro stanovení RU_{est} za den např. 15. března (uveřejňuje se následující den, tedy 16. března) však únorová data ještě k dispozici nejsou a nejaktuálnějšími daty pro stanovení denního koeficientu jsou tak data lednová (měsíc n-2).

Vlastní ukazatel RU je pro jednotlivá Média a jejich sekce doplněn do rozhraní po distribuci měsíčních dat NetMonitoru

Ukazatel počtu cookies, který je znám z rozhraní gemiusTraffic je v aplikaci rovněž zobrazován, pouze však přihlášeným uživatelům, typicky provozovatelům serverů účastnících se měření. Přihlášenému uživateli je rovněž k dispozici informace o návštěvnosti neveřejné části struktury (sekcí) některých serverů (typicky Provozovatel má přístup ke statistikám neveřejných sekcí svých serverů - pokud jsou tyto definovány). Přístupové údaje pro přihlášení do aplikace obdrží příslušný subjekt po dohodě s Realizátorem.

Aplikace umožňuje zobrazení následujících ukazatelů vždy pro zvolené období (den, týden, měsíc) pro všechny měřené servery a jejich veřejné sekce:

- **RU_{est}** odhad počtu reálných uživatelů
- **RU** počet reálných uživatelů
- **U** počet uživatelů (cookies) (pouze po zalogování)
- **PV** počet zobrazení (pageviews) vygenerovaných uživateli
- **V** počet návštěv vykonaných uživateli
- **TTS** celkový strávený čas všemi uživateli
- **ATS_{est}** odhad průměrného času stráveného jedním uživatelem
- **ATS_{RU}** průměrný čas strávený jedním reálným uživatelem
- **AVT** průměrný čas trvání jedné návštěvy

Aplikace pracuje ve dvou režimech:

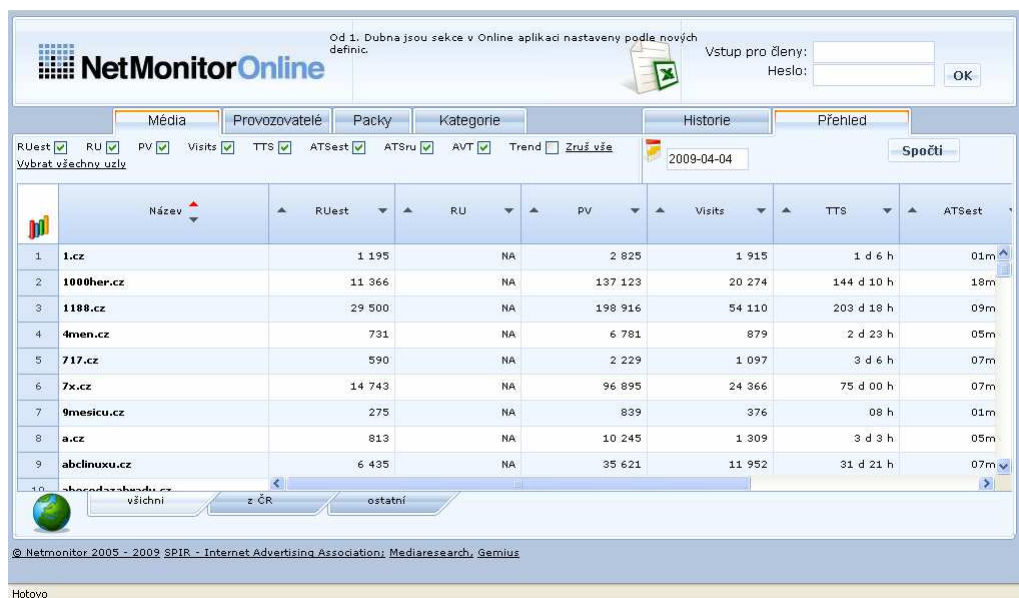
- **Režim Přehled**

V tomto režimu je možné vybrat libovolnou kombinaci výše popsaných ukazatelů a tyto ukazatele jsou zobrazeny pro příslušné časové období. Časové období může být jeden den, týden, nebo měsíc. Data lze zobrazit pro návštěvníky z ČR, pro všechny návštěvníky anebo pro ostatní návštěvníky, tedy ty návštěvníky, kteří nejsou z ČR. Rovněž je možné zobrazit ke každému ukazateli jeho trend, což je procentuální změna oproti předchozímu stejnému období. V případě trendu pro jeden den je trend vztažen k předchozímu stejnému dni v minulém týdnu, tedy např. středa vůči minulé středě. Definici pohledu potvrdíme tlačítkem "Spočti" a následně se nám zobrazí požadovaná tabulka. Klikáním na ikony v záhlaví tabulky lze měnit řazení

údajů v tabulce. Kliknutím na jednotlivé řádky tabulky provádíme jejich výběr a následným kliknutím na ikonku grafu vybrané řádky (resp. jejich hodnoty můžeme zobrazit v grafu ve zvláštním okně). K dispozici je rovněž možnost zobrazit statistiky návštěvnosti pro skupiny Provozovatelů, Packů a Kategorii.

- **Režim Historie**

V tomto režimu je možné vybrat si jeden z možných ukazatelů a ten zobrazit pro definovaný interval dnů, týdnů, respektive měsíců. Tedy je možné například získat počty reálných uživatelů RU na serverech po dnech v období mezi 10.8. a 20.8. Ostatní volby jsou obdobou voleb z režimu "Přehled".



Od 1. Dubna jsou sekce v Online aplikaci nastaveny podle nových definic.

Vstup pro členy:
Heslo:

Média Provozovatelé Packy Kategorie Historie Přehled

RUest RU PV Visits TTS ATSeSt ATSRu AVT Trend Zruš vše

Vybrat všechny uzly

2009-04-04

	Název	RUest	RU	PV	Visits	TTS	ATSeSt
1	1.cz	1 195	NA	2 825	1 915	1 d 6 h	01m
2	1000her.cz	11 366	NA	137 123	20 274	144 d 10 h	18m
3	1188.cz	29 500	NA	198 916	54 110	203 d 18 h	09m
4	4men.cz	731	NA	6 781	879	2 d 23 h	05m
5	717.cz	590	NA	2 229	1 097	3 d 6 h	07m
6	7x.cz	14 743	NA	96 895	24 366	75 d 00 h	07m
7	9mesicu.cz	275	NA	839	376	08 h	01m
8	a.cz	813	NA	10 245	1 309	3 d 3 h	05m
9	abclinuxu.cz	6 435	NA	35 621	11 952	31 d 21 h	07m
10	skopceskouda.cz						

z ČR ostatní

© Netmonitor 2005 - 2009 SPIR - Internet Advertising Association; Mediaresearch, Gemius

Hokovo

obrázek 5 Rozhraní On-line aplikace NetMonitor

10.3 Lokální aplikace gemiusExplorer

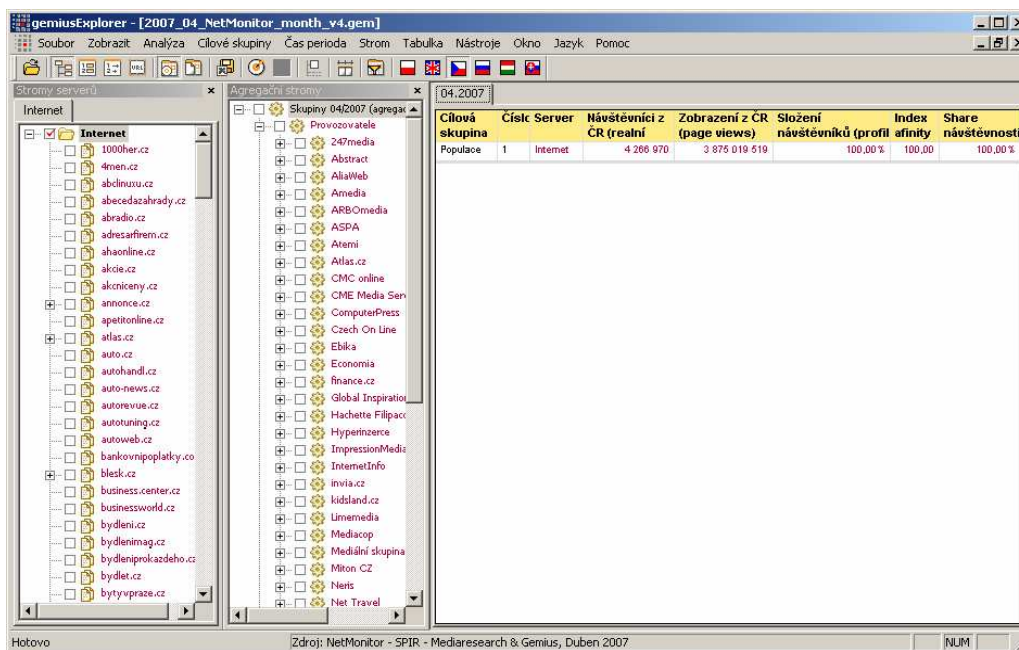
GemiusExplorer je speciální analytický software, který umožňuje analýzy návštěvnosti webových stránek, profilu jejich uživatelů a jejich online aktivity.

Aplikace gemiusExplorer je distribuována s výsledky výzkumu o návštěvnosti a sociodemografickém profilu reálných uživatelů.

Práce s aplikací je uživatelsky jednoduchá, ovládání je obdobné standardním aplikacím. GemiusExplorer pracuje na počítačích s operačním systémem MS Windows

Uživatelské rozhraní je velmi srozumitelné a logické. Data jsou zobrazována ve dvou ovládacích panelech: Na levé straně je stromová struktura médií, resp. stromová struktura tzv. agregačních stromů (například struktura provozovatelů a jednotlivých serverů pod ně spadajících). V této stromové struktuře si uživatel může vybrat, které agregační stromy nebo webové servery chce analyzovat. Zobrazovací panel na pravé straně je připraven ve formě tabulky s vybranými statistikami pro analyzované servery.

Statistiky dostupné v gemiusExplorer dávají uživateli kompletní obraz o analyzované návštěvnosti webových serverů. Statistiky zahrnují mezi mnoha jinými: počet návštěvníků (reálných uživatelů), počet zobrazení, průměrný čas návštěvy, dosažený reach nebo afinitu.



obrázek 6 Uživatelské rozhraní gemiusExplorer

Významnou výhodou programu gemiusExplorer je jeho optimalizace na data z online měření – z čehož vyplývá velmi vysoká interaktivita a rychlost získání požadovaného výsledku.

Podrobné informace o software jsou obsaženy v uživatelské příručce a oprávnění uživatelé jsou proškoleni Realizátorem výzkumu. Podrobný

popis software je nad rámec této základní metodiky, informace je možné získat u Realizátora.

Kromě dat *.gem pro užití v software gemiusExplorer jsou dodávána též ASCII data pro konverzi do jiných softwarů, například do Data Analyzeru společnosti Median.

Data do gemiusExploreru obsahují kompletní informaci o návštěvnosti jednotlivých Médii, jejich sekcí či vyšších skupin (např. souhrn za všechny servery Provozovatelů) na bázi RU a rovněž informaci o sociodemografickém profilu návštěvníků všech těchto uzlů. Přístup k těmto datům je pak řízen licencí (uživatelským jménem a heslem), která opravňuje uživatele ke studiu statistik návštěvnosti všech uzlů, avšak přístup k sociodemografii umožňuje jen u uzlů, k nimž má uživatel přístup (typicky Provozovatel má možnost zkoumat návštěvnost cílových skupin jen na svých serverech). Toto omezení se nevztahuje na licence subjektů, jež jsou odběrateli kompletních dat NetMonitoru. Jejich licence je opravňuje ke zkoumání celého datového souboru (všech uzlů, jejich návštěvnosti i sociodemografie) bez omezení.

Každý Provozovatel má nárok na 3 licence ke gemiusExploreru zdarma, odběratelé kompletních dat pak na 10 licencí. Každá licence nad tento limit je Zadavatelem zpoplatňována.

11 DATA A REPORTY

11.1 Výsledky a data z měření návštěvnosti

Data o návštěvnosti jsou reportována oprávněným uživatelům v reálném čase prostřednictvím on-line rozhraní **gemiusTraffic**. Dále jsou reportována data o hodinové aktivitě v **denních reportech**. Navíc jsou základní ukazatelé o všech (aktivních) měřených serverech k dispozici v „**On-line aplikaci NetMonitoru**“ Data s přepočtem na reálné uživatele jsou k dispozici ve formátu pro **software gemiusExplorer** a dále jsou součástí **měsíčních (off-line) reportů**. Základní data o návštěvnosti jsou rovněž k dispozici v **ASCII** formátu pro ostatní softwary.

11.1.1 On-line reporty

Podrobná on-line data přístupná přes rozhraní gemiusTraffic jsou **privátní** (tj. data o jednom médiu či skupině médií jsou přístupná pouze vlastníku těchto médií).

Základní on-line data přístupná v rozhraní on-line aplikace NetMonitoru jsou **veřejná**. Přístup k těmto datům je umožněn každému zájemci, navíc vlastníci jednotlivých médií mají po přihlášení možnost analyzovat vlastněné médium v rámci celé hloubky definovaného stromu.

Výpočetní ukazatele poskytované on-line v rozhraní gemiusTraffic:

Data jsou zpracovávána okamžitě a průměrný čas mezi koncem aktivity internetového uživatele a jejím zobrazením v rozhraní je cca 15 minut. S výjimkou technické přestávky každou středu od 8 do 10 hodin, či mimořádně 45 minut mimo stanovený čas technické přestávky. Některé statistiky (např. „Odkazující stránky“) si žádají delší (hromadné) zpracování a jejich výsledky jsou v rozhraní k dispozici následující den.

Většina statistik je k dispozici v následujících časových rozlišeních:

- rozlišení hodin (pro předchozí a aktuální den)
- rozlišení dnů (pro předchozí a aktuální měsíc)
- rozlišení týdnů (pro předchozí a aktuální kvartál)
- rozlišení měsíců (pro předchozí a aktuální rok)

K dispozici jsou následující statistiky:

<ul style="list-style-type: none"> • Návštěvníci (cookies) → souhrn • Návštěvy → souhrn • Návštěvy → délka • Návštěvy → rozsah 	<ul style="list-style-type: none"> • Zeměpisná lokalizace → souhrn • Zeměpisná lokalizace → světadíly • Zeměpisná lokalizace → státy • Zeměpisná lokalizace → ČR kraje
--	--

<ul style="list-style-type: none"> • Návštěvy → doba • Návštěvy → podrobně • Zobrazení → souhrn • Zobrazení → doba • Strávený čas → souhrn • Odkazy → souhrn • Odkazy → odkazující stránky • Odkazy → vyhledávače • Odkazy → Katalogy 	<ul style="list-style-type: none"> • Zeměpisná lokalizace → ČR města • Systém → operační systémy • Systém → prohlížeče • Systém → jazykové verze • Systém → rozlišení • Systém → hloubka barev
--	--

Výpočetní ukazatele poskytované v on-line aplikaci NetMonitoru:

Data o předchozím dnu jsou zpracovávána následující den ráno. Stejným způsobem se kompletují data po skončení delšího období, např. data za týden jsou zpracovávána během následujícího pracovního dne (ráno). Aplikace umožňuje analyzovat zvolené ukazatele za dané časové období a nebo zobrazit vývoj jednoho zvoleného ukazatele v delším časovém úseku a to včetně trendů. Aplikace pracuje s daty na denní, týdenní a měsíční bázi.

K dispozici jsou následující ukazatelé:

- Návštěvníci (cookies)
- Zobrazení (pageviews)
- Návštěvy
- Délka všech návštěv
- Délka návštěv uživatele (ATS)
- Průměrná délka návštěvy

11.1.2 Denní reporty

Denní reporty jsou **privátní** (tj. data o jednom médiu či skupině médií jsou přístupná pouze vlastníku těchto médií). Denní report obsahuje informace o hodinové aktivitě návštěvníků (cookies) na příslušném serveru. Tento report je generován ráno následující den o dni předchozím a je k dispozici ke stažení prostřednictvím webového rozhraní.

K dispozici jsou následující ukazatelé:

- Návštěvníci (cookies)
- Zobrazení (pageviews)
- Návštěvy

11.2 Výsledky a data z výzkumu sociodemografického profilu návštěvníků spojená s daty o návštěvnosti

11.2.1 Měsíční reporty o návštěvnosti

Měsíční reporty s informací o přepočtené návštěvnosti reálných uživatelů (off-line reporty) jsou **veřejné**, volně k dispozici na stránkách netmonitor.cz každého 25-tého dne následujícího měsíce (či následující pracovní den) o předchozím měsíci (např. 25. dubna o březnu). K dispozici jsou informace o všech reportovaných serverech, případně jejich veřejných sekcích, dále pak o skupinách serverů jednotlivých Provozovatelů, o volně definovaných skupinách serverů (Packy) a o Kategoriích včetně pořadí uzlů podle RU v těchto kategoriích. Informace jsou rozděleny na část zahrnující uživatele přistupující z České republiky (českých IP adres) a na všechny uživatele (uživatelé přistupující z libovolných IP adres).

K dispozici jsou následující ukazatelé:

- Počet RU (reálných uživatelů) za měsíc
- Průměrný počet RU (reálných uživatelů) za týden
- Průměrný počet RU (reálných uživatelů) za den
- Průměrný počet RU (reálných uživatelů) za pracovní den
- Průměrný počet RU (reálných uživatelů) za den o víkendu
- Počet zobrazení (pageviews)
- Průměrný strávený čas za měsíc
- Průměrný strávený čas za den
- Počet návštěv
- Průměrná délka návštěvy

11.2.2 Měsíční report o sociodemografii

Měsíční reporty s informací o sociodemografickém profilu jsou **privátní** a jsou k dispozici pouze Provozovatelům příslušného serveru. Tyto reporty obsahují základní informace o serveru rozčleněné podle jednotlivých socio-demografických charakteristik (např. pohlaví, vzdělání apod.) Je možné porovnání s rozložením na všech měřených serverech vzatých jako jeden celek. Odběratelům je k dispozici každého 25-tého dne následujícího měsíce (nebo následující pracovní den) o předchozím měsíci. Souhrnný report o sociodemografickém profilu všech (aktivních) měřených médií celkem je **veřejný** a je k dispozici na netmonitor.cz. Z důvodu chybějící informace o sociodemografické struktuře zahraničních uživatelů tyto reporty obsahují informaci pouze za RU z ČR.

K dispozici jsou následující ukazatelé:

- Počet RU (reálných uživatelů) (absolutně i v procentech)
- Počet zobrazení (pageviews) (absolutně i v procentech)
- Reach
- TRP
- Afinita

11.2.3 Měsíční data pro gemiusExplorer

Data do gemiusExploreru obsahují kompletní informaci o návštěvnosti jednotlivých Médii, jejich veřejných sekcí či vyšších skupin (např. souhrn za všechny servery Provozovatelů) na bázi RU a rovněž informaci o sociodemografickém profilu návštěvníků všech těchto uzlů (kompletní data). Přístup k těmto datům je pak řízen licencí (uživatelským jménem a heslem), která opravňuje uživatele ke studiu statistik návštěvnosti všech uzlů, avšak přístup k sociodemografii umožňuje jen u uzlů, k nimž má uživatel přístup (typicky Provozovatel má možnost zkoumat návštěvnost cílových skupin jen na svých serverech). Toto omezení se nevztahuje na licence subjektů, jež jsou odběrateli kompletních dat NetMonitoru. Jejich licence je opravňuje ke zkoumání celého datového souboru (všech uzlů, jejich návštěvnosti i sociodemografie) bez omezení.

Data umožňují definici vlastních cílových skupin (překřížením základních skupin) a rovněž v rámci daného měsíce i definici vlastního časového období, a to i nespojitého (např. jen všechny středy či víkendy). Nejnižší báží, na které lze analýzy provádět je báze jednotlivých dnů, nejvyšší pak měsíční,

Data jsou odběratelům dodávána každého 25-tého dne následujícího měsíce (či nejbližší následující pracovní den) o předchozím měsíci. Data jsou připravována ve formátu .gem pro zpracování v analytické aplikaci gemiusExplorer a ve formátu ASCII vhodném pro načtení do ostatních analytických nástrojů dostupných na trhu.

Z důvodu chybějící informace o sociodemografické struktuře zahraničních uživatelů tato data obsahují informaci pouze za RU z ČR.

12 UKAZATELE A SOCIODEMOGRAFICKÉ PROMĚNNÉ

12.1 Definice ukazatelů

Celkový strávený čas – součet časů strávených návštěvníky (reálnými uživateli) z dané cílové skupiny na příslušném webovém serveru nebo jeho části. Ukazatel je zobrazen v hodinách.

Návštěva – sled zobrazení na daném webovém serveru, nepřerušovaný časovým úsekem delším než 30 minut.

Návštěvníci (reální uživatelé) – počet uživatelů internetu v příslušné cílové skupině, kteří vygenerovali alespoň jedno zobrazení na vybraném webovém serveru během daného měsíce (týdne, dne). Tento ukazatel odpovídá skutečnému počtu lidí (ne počítačů, cookies nebo IP adres), kteří navštívili webový server v daném měsíci (týdnu, dnu).

Zobrazení (page views) – počet zobrazení vygenerovaných příslušnou cílovou skupinou na vybraném webovém serveru během daného měsíce (týdne, dne).

Průměrný strávený čas na návštěvníka – celkový strávený čas průměrným návštěvníkem (reálným uživatelem) z příslušné cílové skupiny na vybraném webovém serveru během daného měsíce.

Průměrný čas trvání zobrazení – průměrný čas mezi dvěma po sobě jdoucími zobrazeními jedné návštěvy. Tato zobrazení jsou generována příslušnou cílovou skupinou na vybraném webovém serveru během daného měsíce (týdne, dne).

Průměrný počet zobrazení na návštěvníka – počet zobrazení generovaných průměrným návštěvníkem z příslušné cílové skupiny na vybraném webovém serveru během daného měsíce (týdne, dne).

Průměrný počet návštěv na návštěvníka – počet návštěv generovaných průměrným návštěvníkem z příslušné cílové skupiny na vybraném webovém serveru během daného měsíce (týdne, dne).

Reach – poměr počtu návštěvníků (reálných uživatelů) z příslušné cílové skupiny, kteří vygenerovali alespoň jedno zobrazení na vybraném

webovém serveru během daného měsíce (týdne, dne) k celkovému počtu uživatelů internetu za daný měsíc. Je zobrazen jako procento.

Složení návštěvníků (profil reache) – poměr počtu návštěvníků (reálných uživatelů) z příslušné cílové skupiny, kteří navštívili vybraný webový server k celkovému počtu návštěvníků (reálných uživatelů), kteří navštívili vybraný webový server během daného měsíce (týdne, dne). Je zobrazen jako procento.

Složení zobrazení (profil zobrazení) – poměr počtu zobrazení vygenerovaných příslušnou cílovou skupinou na vybraném webovém serveru k počtu zobrazení vygenerovaných všemi návštěvníky (reálnými uživateli) na vybraném webovém serveru během daného měsíce (týdne, dne). Je zobrazen jako procento.

Afinita reache – poměr složení návštěvníků příslušné cílové skupiny pro vybraný webový server ke složení návštěvníků příslušné cílové skupiny pro skupinu všech měřených serverů.

Share návštěvnosti – poměr počtu návštěvníků (reálných uživatelů) z příslušné cílové skupiny, kteří navštívili vybraný webový server k počtu návštěvníků (reálných uživatelů) ze stejné cílové skupiny, kteří navštívili skupinu všech měřených serverů. Je zobrazen jako procento.

Share zobrazení – poměr počtu zobrazení vygenerovaných příslušnou cílovou skupinou na vybraném webovém serveru k počtu zobrazení vygenerovaných stejnou cílovou skupinou na skupině všech měřených serverů. Je zobrazen jako procento.

Share stráveného času – poměr celkového stráveného času příslušnou cílovou skupinou na vybraném webovém serveru k celkově strávenému času stejnou cílovou skupinou na všech měřených serverech. Je zobrazen jako procento.

Duplikace návštěvnosti (%) – procento z návštěvníků (reálných uživatelů) příslušného webového serveru, kteří také navštívili alespoň jeden z ostatních vybraných webových serverů.

Populace – celkový počet návštěvníků (reálných uživatelů) za daný měsíc.

12.2 Sociodemografické proměnné v datech

- Všechna místa používání internetu
- Způsob připojení doma
- Měřená četnost užívání internetu
- IT znalosti
- Pohlaví
- Věk
- Nejvyšší dokončené vzdělání
- Rodinný stav
- Počet osob v domácnosti
- Hospodyně
- Hlava domácnosti
- Ekonomická aktivita
- ABCDE národní socioekonomická klasifikace
- Čistý příjem domácnosti
- Velikost místa bydliště
- Kraj
- Vybavenost
- Ekonomicky aktivní jako
- Charakteristika zaměstnání
- Podnikatelská činnost
- Počet podřízených
- Ekonomicky neaktivní jako
- Činnost respondenta na počítači
- Pravidelné aktivity na internetu
- Čas trávený na zahraničních serverech v cizích jazycích vyjma slovenštiny