



Erkenntnntnisse aus dem Suchverhalten im Web "Muster in Suchanfragen"

Nadine Schmidt-Mänz



Gliederung

1. Einleitung und Motivation
2. Kernergebnisse der Umfrage:
"Suchmaschinen und Suchverhalten im Internet"
3. Analyse von Suchanfragen:
Fireball, Lycos, MetaGer und Metaspinner
4. Dauerbrenner und Eintagsfliegen
5. Impulse und Events
6. Handlungsempfehlungen und weitere Forschung



1. Einleitung und Motivation

- Suchmaschinen sind bedeutende Anlaufstellen.
- Suchmaschinennutzer müssen in den Fokus gerückt werden.
- Tiefgreifende Analysen des Suchverhaltens unumgänglich.



Figure 1: "First they do an on-line search."

Mit der Kenntnis dieses Verhaltens

- können Störfaktoren der Internetnutzer bei der Suche im Internet eliminiert werden.
- können Web-Angebote direkt an dem Verhalten der Suchenden ausgerichtet werden.
- werden Verbesserungspotentiale für Suchmaschinen aufgedeckt.
- werden Möglichkeiten geboten, den Suchenden bei der Recherche zu unterstützen.
- können Marketingaktionen zielgerichteter durchgeführt werden.



2. Kernergebnisse der Umfrage "Suchmaschinen und Suchverhalten im Internet"

Deskriptiva:

- Selbstselektiver Charakter
- Altersunterschied niedriger, mehr männliche Teilnehmer
- Wenige Suchbegriffe und Ergebnisseiten
- Recherche nach Personen und Produkten
- Sehr oft komplexe Suche oder Eingabe von Phrasen
- Keine Personalisierung der Suchmaschinen
- Ablehnung von Werbung

Untersuchung verschiedener Gruppen:

- Bei Gruppierung nach Wissen unterschied sich der Aufbau der Suchanfragen überdurchschnittlich.
- Die Angestellten zeigten gegenüber den Studenten eine zeitoptimiertere Gestaltung bei der Recherche auf.
- Frauen gegenüber Männern stöbern mehr in News.
- Nutzer die mehr über Suchmaschinen wissen, lehnen Werbung stärker ab und gehen zielgerichteter vor.



3. Analyse von Suchanfragen

- Suchanfrage, Suchterme und Operatoren
- Observationsperiode $[1; R]$ bzw. $[0; T]$
- Zeitpunkt $r \in [1; R]$, Zeitintervall $t \in [0; T]$
- Suchanfrage SQ_{ZSt_r} mit Zeitstempel ZSt_r
- t repräsentiert eine Stunde, vier Stunden oder auch einen Tag (Taktung)
- $\vec{S}_{ZSt_r} = (s_1, \dots, s_m, \dots, s_{M_{ZSt_r}})$
- $\vec{V}_{ZSt_r} = (v_1, \dots, v_n, \dots, v_{N_{ZSt_r}})$
- s_m, v_n Häufigkeit des Erscheinens
- $\|\vec{S}_{ZSt_r}\|_1 = \sum_{m=1}^{M_{ZSt_r}} s_m$ Bruttoanzahl Anfragen
- $\|\vec{V}_{ZSt_r}\|_1 = \sum_{n=1}^{N_{ZSt_r}} v_n$ Bruttoanzahl Suchterme
- M, N Nettoanzahl Suchanfragen/-terme
- $\vec{S}^{(t)}, \vec{V}^{(t)}$ am Ende des Intervalls t

t	r	SQ_{ZSt_r}	\vec{S}_{ZSt_r}	\vec{V}_{ZSt_r}	$\vec{S}^{(t)}$	$\vec{V}^{(t)}$
1	1	(term ₁ , term ₂ , term ₃)	(1)	(1, 1, 1)	–	–
	2	(term ₂ , term ₁ , term ₃)	(1, 1)	(2, 2, 2)	–	–
	3	(term ₁ , term ₃)	(1, 1, 1)	(3, 2, 3)	–	–
	4	(term ₄ , term ₅)	(1, 1, 1, 1)	(3, 2, 3, 1, 1)	–	–
	5	(term ₁ , term ₃)	(1, 1, 2, 1)	(4, 2, 4, 1, 1)	$\vec{S}^{(1)}$	$\vec{V}^{(1)}$
2	6	(term ₆)	(1, 1, 2, 1, 1)	(4, 2, 4, 1, 1, 1)	–	–
	7	(term ₁ , term ₂ , term ₃)	(2, 1, 2, 1, 1)	(5, 3, 5, 1, 1, 1)	$\vec{S}^{(2)}$	$\vec{V}^{(2)}$
3	8	(term ₇)	(2, 1, 2, 1, 1, 1)	(4, 2, 4, 1, 1, 1, 1)	–	–
	9	(term ₄ , term ₅)	(2, 1, 2, 2, 1, 1)	(4, 2, 4, 2, 2, 1, 1)	–	–
	10	(term ₆)	(2, 1, 2, 2, 2, 1)	(4, 2, 4, 2, 2, 2, 1)	–	–
	11	(term ₁)	(2, 1, 2, 2, 2, 1, 1)	(5, 2, 4, 2, 2, 2, 1)	–	–
	12	(term ₁ , term ₃)	(2, 1, 3, 2, 2, 1, 1)	(6, 2, 5, 2, 2, 2, 1)	–	–
	13	(term ₂ , term ₁ , term ₃)	(2, 2, 3, 2, 2, 1, 1)	(7, 3, 6, 2, 2, 2, 1)	$\vec{S}^{(3)}$	$\vec{V}^{(3)}$
4	14	(term ₆)	(2, 2, 3, 2, 3, 1, 1)	(7, 3, 6, 2, 2, 3, 1)	$\vec{S}^{(4)}$	$\vec{V}^{(4)}$
⋮	⋮	⋮	⋮	⋮	⋮	⋮

Table 1: Beispiel für eingegangene Suchanfragen



3. Analyse von Suchanfragen

In Tabelle 2 sind die genauen Beobachtungszeiträume dargestellt.

Suchmaschine	Erster Tag	Letzter Tag	Zeitraum	Anzahl Tage
Fireball	18.08.2004	20.09.2005	13 Monate	399
Lycos	14.08.2004	20.09.2005	13 Monate	403
Metager <small>(Top-4000)</small>	11.11.2004	20.09.2005	10 Monate	314
Metaspinner	28.09.2004	20.09.2005	12 Monate	358

Table 2: Observationsperioden der einzelnen Suchmaschinen

In Tabelle 3 sind die Brutto- und Nettoanzahlen der erhobenen Suchterme, bzw. Suchanfragen aufgezeigt.

Search Engine	$\ \vec{S}\ _1$	M	\bar{s}	$\ \vec{V}\ _1$	N	\bar{v}
Fireball	132.833.007	17.992.069	7,4	241.833.877	6.296.833	38,4
Lycos	189.930.859	29.322.366	6,5	344.242.099	11.232.710	30,6
Metager <small>(Top-4000)</small>	4.407.566	678.655	6,5	7.333.343	430.338	17,0
Metaspinner	4.089.731	1.287.417	3,2	7.853.501	627.507	12,52

Table 3: Netto- und Bruttozahlen



3. Analyse von Suchanfragen

Die **durchschnittliche Länge** von Suchanfragen
ähnlich lang:

1,6 (Metager), 1,7 (Lycos) und 1,8 (Fireball und
Metaspinner) Terme pro Suchanfrage
(Bruttomenge der Suchanfragen)

1,4 (Metager), 2,0 (Metaspinner), 2,5 (Lycos) und 2,7
(Fireball) Terme pro Suchanfrage
(Nettomenge der Suchanfragen).

Prozentualer Anteil von allen Suchanfragen mit
Phrasensuche:

Fireball 2,1%, bei Lycos 1,7% und bei Metaspinner
2,1%

Nutzung **spezieller Funktionen** bzw. Features:

Suchbereich	Fireball	Metaspinner
Bilder	24,5	1,6
Deutsche Suche	66,4	88,1
Internationale/Weltweite Suche	9,1	4,2
ohne Suchbereich	0,0	1,9
(Online) Auktionen	–	3,7
sonstige Suchbereiche	–	0,4

Table 4: Suchbereiche (in Prozent)



3. Analyse von Suchanfragen

Wie man sieht liegt das Tageshoch zwischen 14.00 und 16.00 Uhr. Der absolute Tiefpunkt ist wird in den frühen Morgenstunden um 05.00 Uhr erreicht.

An den Montagen werden Informationen des Wochenendes verarbeitet. Der **Montagseffekt** ist deutlich sichtbar.

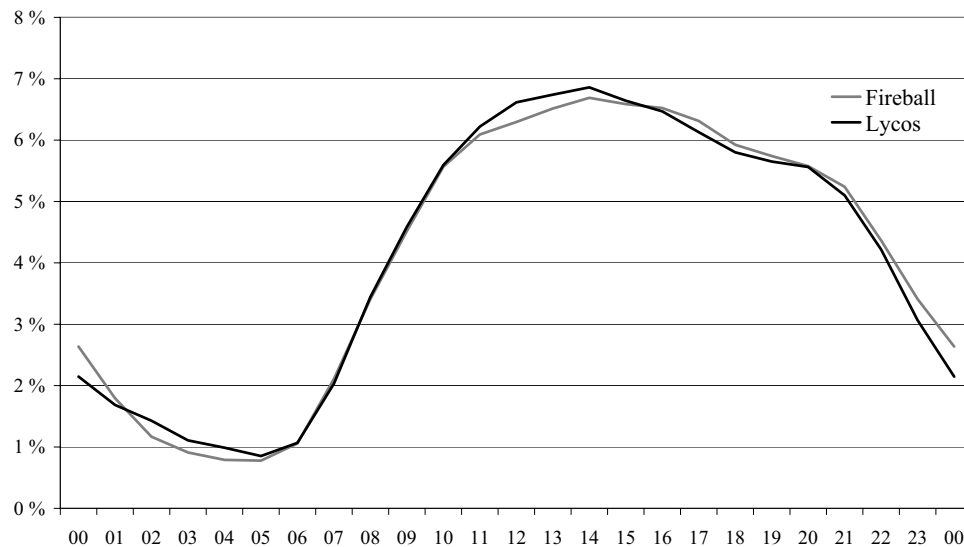


Figure 2: Mittlere prozentuale Anzahl pro Stunde

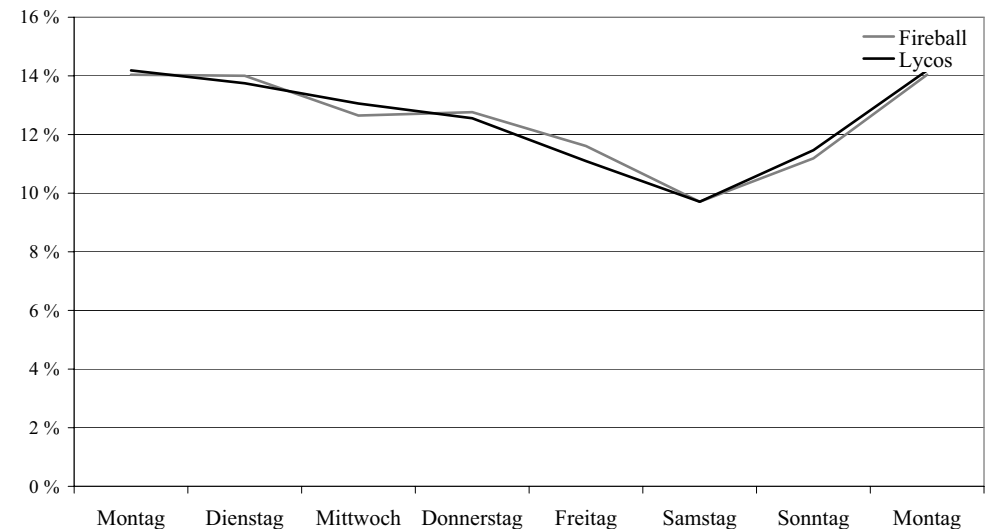


Figure 3: Mittlere prozentuale Anzahl pro Woche



3. Analyse von Suchanfragen

Betrachtung der Top-10 der Suchanfragen nach unterschiedlich langen Zeitperioden:

Top-10 der Suchterme nach einer langen Periode und an verschiedenen Tagen:

Rang	Fireball	Lycos	Metager (Top-4000)	Metaspinner
1.	sex	lycos	ebay	link
2.	porno	link:http://www.	google	weiterentwicklung
3.	hentai	sex	routenplaner	routenplaner
4.	lycos	hentai	telefonbuch	sex
5.	lack	porno	gmx	gebrauchtwagen
6.	fkf	ebay	web.de	jahreswagen
7.	erotik	google	sex	london
8.	anal	erotik	freenet	hotel
9.	bondage	plexiglasgehäuse	www.ebay.de	versicherung
10.	muschi	christina aguilera	www.google.de	autobewertung

Table 5: Top-10 der Suchanfragen Dezember 2005

Rang	Fireball	Lycos	Metager (Top-4000)	Metaspinner
1.	sex	lycos	ebay	in
2.	download	sex	google	download
3.	in	in	der	link
4.	free	der	routenplaner	der
5.	porno	und	download	berlin
6.	dvd	hotel	und	hotel
7.	hotel	von	in	und
8.	kostenlos	für	für	für
9.	der	porno	berlin	dvd
10.	bilder	free	sex	von

Table 7: Top-10 der Suchterme Mai 2005

Rang	Fireball	Lycos	Metager (Top-4000)	Metaspinner
1.	sex	lycos	ebay	link
2.	porno	sex	google	sex
3.	hentai	link:http://www.	routenplaner	routenplaner
4.	lycos	hentai	gmx	gebrauchtwagen
5.	fkf	porno	telefonbuch	hotel
6.	erotik	ebay	www.ebay.de	london
7.	lack	erotik	web.de	jahreswagen
8.	anal	google	sex	autobewertung
9.	titten	fkf	www.google.de	versicherung
10.	bondage	christina aguilera	wetter	werkswagen

Table 6: Top-10 der Suchanfragen (Mai 2005)

Rang	05.09.2004	10.12.2004	23.12.2004	17.09.2005
1.	+	+	+	damion
2.	lycos	lycos	lycos	mail
3.	in	sex	sex	+
4.	sex	in	im	link
5.	jackie	download	bluetooth-software	lycos
6.	der	paris	hotel	neil
7.	hentai	der	der	sex
8.	mailto	hilton	lastminute	in
9.	the	bilder	porno	add
10.	free	porno	online	hentai

Table 8: Top-10 der Suchterme für verschiedene Tage



4. Dauerbrenner und Eintagsfliegen

- Eintagsfliege (Mayfly): Term, der in nur sehr wenigen Zeitintervallen t auftritt.
- Dauerbrenner (Evergreen): Term, der in vielen Intervallen t vorkommt.
- Weiterhin wird (wahlweise) ein Schwellenwert f definiert.

Term	c_i	Term	c_i	Term	c_i	Term	c_i
(o) +	243	(e) erotik	242	(e) fkk	239	(o) or	233
(o) -	243	(e) girls	242	(f) mit	239	(e) titten	233
(r) berlin	243	(e) nackt	242	(sm) ebay	238	(m) video	233
(m) bilder	243	(e) sexy	242	(e) gay	238	(m) software	232
(f) der	243	(f) von	242	(f) im	238	(e,r) hilton	231
(f) die	243	(a) de	241	(sm) google	237	(e) hardcore	228
(m) download	243	(e) frauen	241	(sh) kostenlos	237	(o) &	227
(m) free	243	(r) hotel	241	(e) nude	237	(e) manga	227
(e) hentai	243	(f) of	241	(e) porn	237	(m) chat	226
(f) in	243	(m) online	241	(sh) kostenlose	236	(sh) test	226
(sm) lycos	243	(e,r) paris	241	(m) mp3	236	(sh) gratis	224
(e) porno	243	(o) and	240	(r) hamburg	235	(m) musik	222
(e) sex	243	(f) für	240	(e) pics	235	(r) münchen	222
(f) the	243	(f) 2	239	(e) girl	234	(a) auto	221
(o) und	243	(m) dvd	239	(sh) shop	234	(e) gina	220
						(f) das	219

Table 9: Dauerbrenner der Terme bei Lycos, für $f = 400$. Diese Terme lagen an 90% der erhobenen Tage t über dem Schwellenwert.



4. Dauerbrenner und Eintagsfliegen

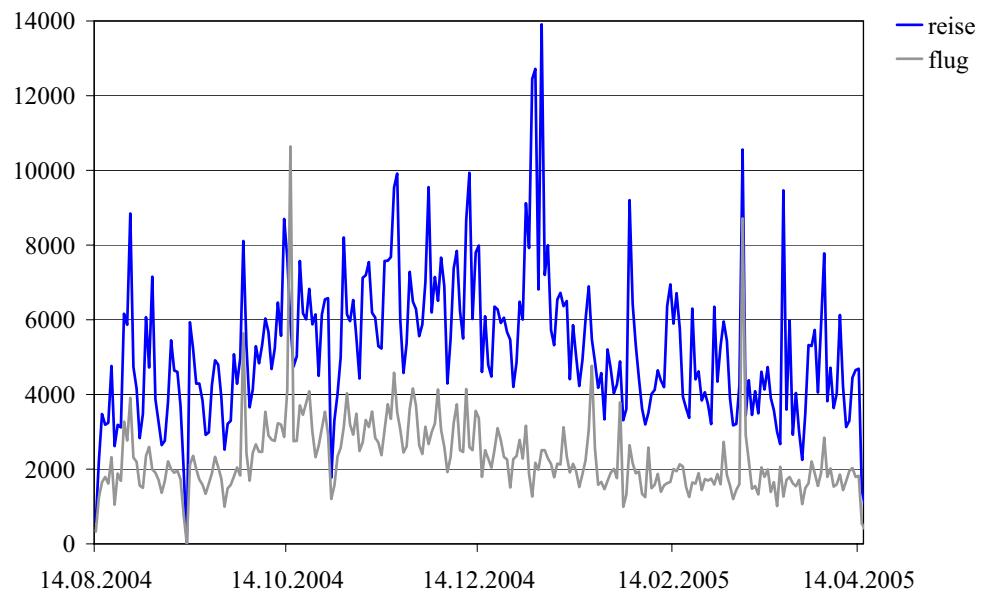


Figure 4: Beispiel für Dauerbrenner: 'sex' und 'berlin'

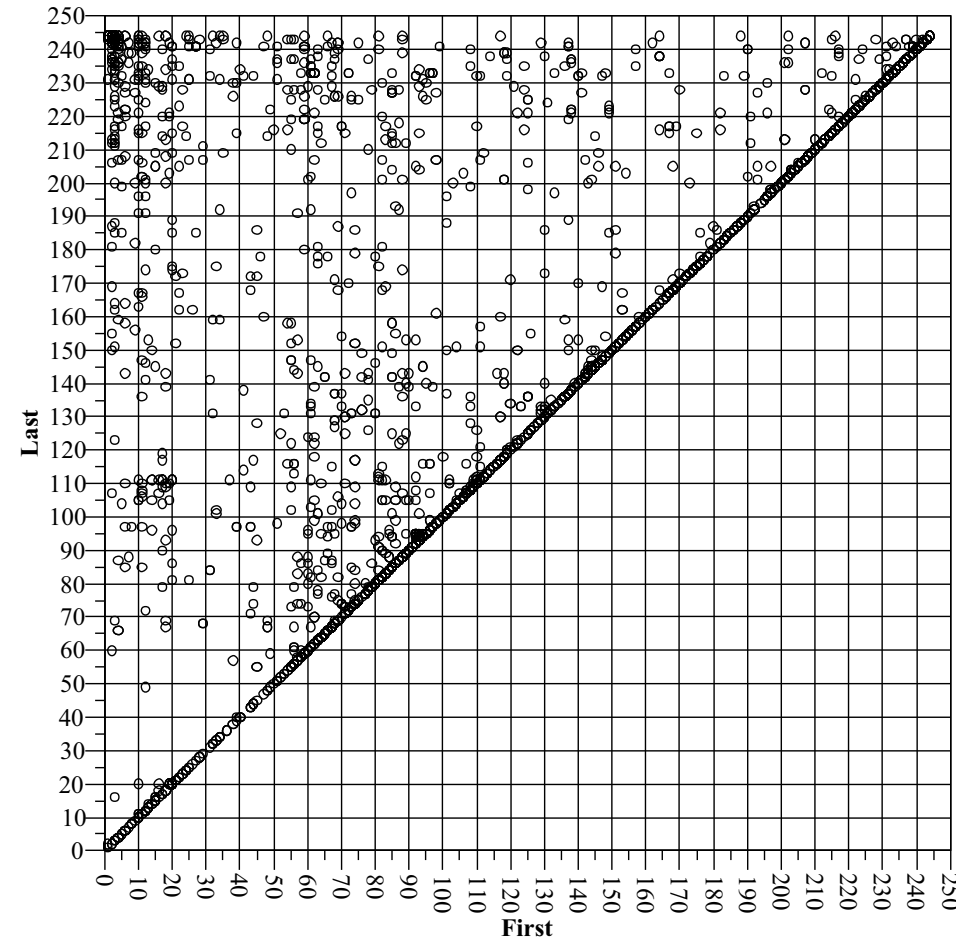


Figure 5: Erstes vs. letztes Auftreten eines Termes



5. Impulse und Events

- Event:
Vorhersehbare oder angekündigte Ereignisse

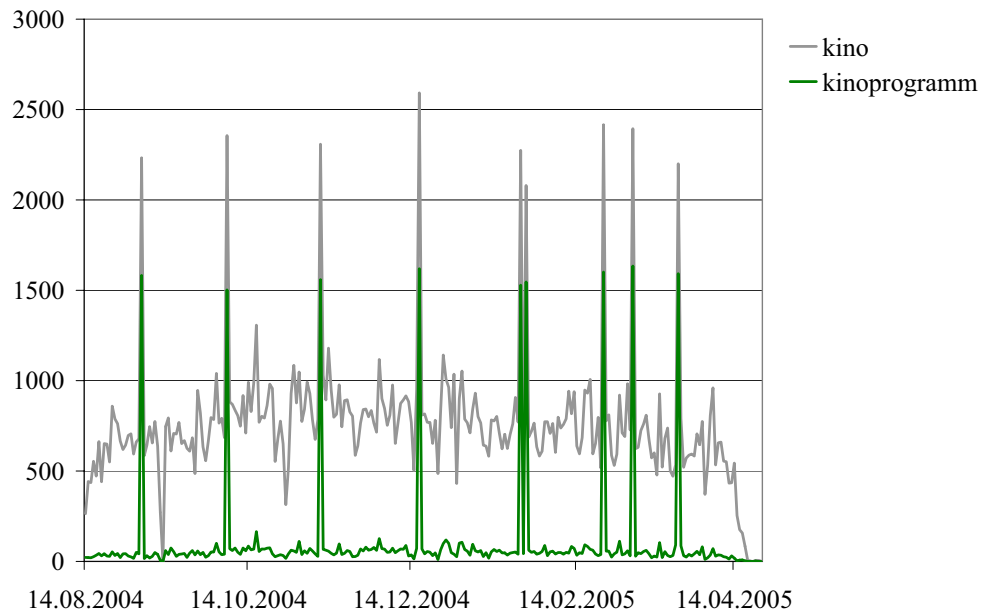


Figure 6: Beispiel für Event: Kino

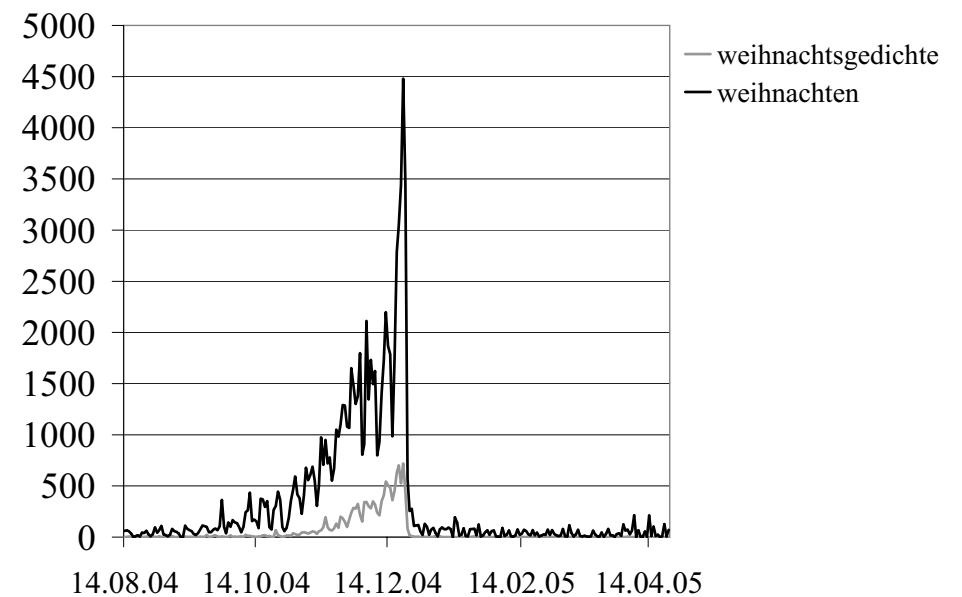


Figure 7: Beispiel für Event: Weihnachten



5. Impulse und Events

- Impuls:
Unvorhersehbare oder unangekündigte Ereignisse

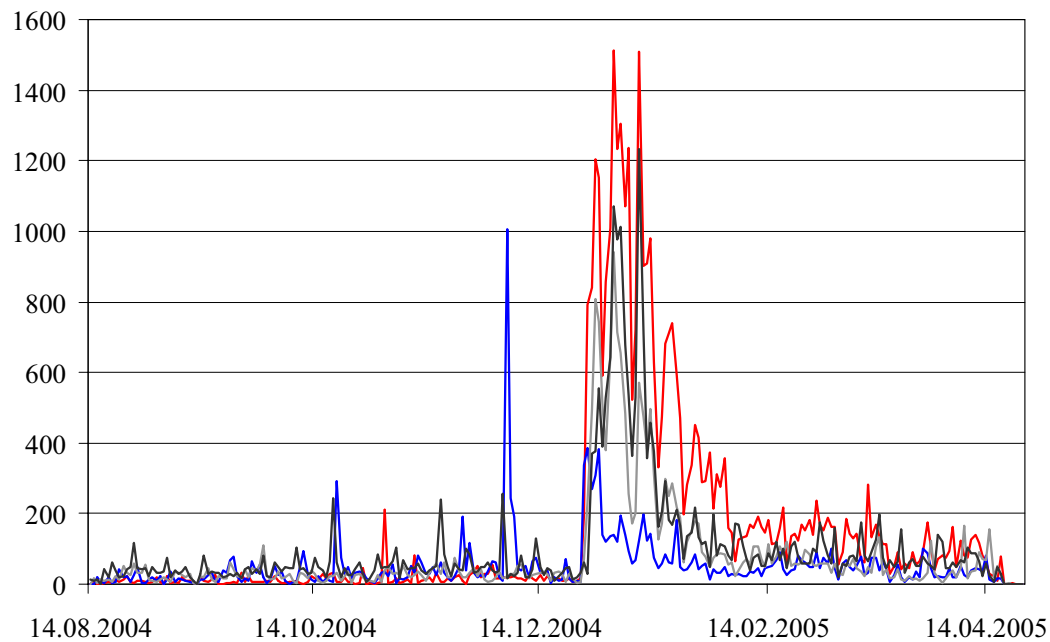


Figure 8: Beispiel für Impuls: Tsunami

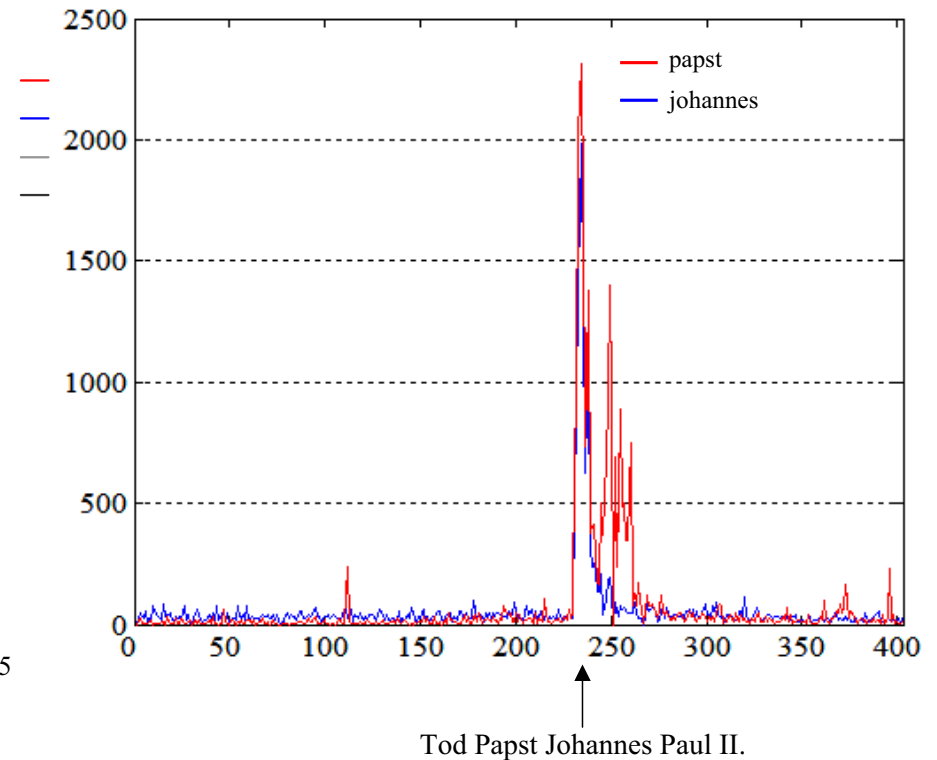


Figure 9: Beispiel für Impuls: Papst (Tod und Wahl)



6. Handlungsempfehlungen und weitere Forschung

- Einführung von Tabu-Suche (Ebay, Porno)
- Vorsicht bei Werbung (Pop-Ups)
- Suchmaschinen vereinfachen → Empfehlungsmaschinen

- Generelle charakteristische Klassen für Terme
- Clusteralgorithmus zu Identifizierung zeitlich ähnlicher Terme (implementiert)
- Caching Strategien für Dauerbrenner
- Impuls adWORDtising, aber auch Dauerbrenner und Events
- Optimierte inhaltliche Ausrichtung von Webseiten (Online Marketing)

- Weitere Erforschung der online Informationsverarbeitung
- Mikrotrends und Periodizitäten (Fourier)
- Weitere mathematische Formulierung der unterschiedlichen Termklassen



Kontakt:

Nadine Schmidt-Mänz

Institut für
Entscheidungstheorie und Unternehmensforschung
Universität Karlsruhe (TH)
home marketing.wiwi.uni-karlsruhe.de

mail nadine.maenz@wiwi.uni-karlsruhe.de

cell +49 176 600 17 0 15

phone +49 721 608 4770

skype n.maenz