
Tool- und Mindset für große Seiten

Julian Cordes

Was ist mit euch?

- Wer betreut große Seiten?
- Was ist groß?
- Wie viele Seiten hat die Domain?
- Wie viele Seiten sind indiziert?
- Wie viele Seiten sollen indiziert sein?

Probleme bei Online-Shops

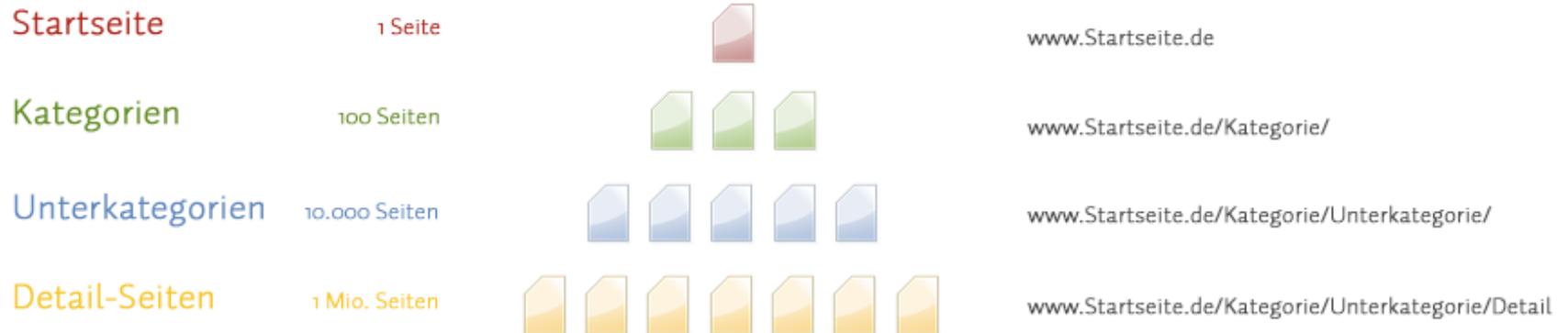
- Pattern für Title und Desc

[Damen günstig kaufen](http://www.roseversand.de) im Online Bike Shop www.roseversand.de
www.roseversand.de/produkte/technik/saettel-und.../saettel/damen/

Damen, die beim Radfahren nicht auf Komfort verzichten wollen, sollten in einen speziellen Damen-Fahrradsattel investieren. Die Ausstattung ist dabei sehr ...

- Herren Sneaker / Damen Sneaker
- <http://www.domain.de/tv/produkt1.html>
- <http://www.domain.de/sale/produkt1.html>

Interne Verlinkung – der größte Hebel



Wo liegt die Power?

- Problem Kanalisierung
 - Lösung: intelligente Verlinkung
- Problem Parameter
 - Lösung: Canonical?
 - und / oder: in WMT Parameter entfernen

Faceted Search – Fluch und Segen

Kat: Fernseher

Typ

- LED
- LCD
- Plasma

Farbe

- Schwarz
- Weiß

Marke

- Philips
- Sony

1 Filter	= 7 (+1) Seiten
2 Filter versch. Typen	= 12 + 4 (+8) = 24 Kombinationen
3 Filter versch. Typen	= 12 (+24) = 36 Kombinationen
2 Filter	= 64 (+8) = 72 Kombinationen
3 Filter	= 729 (+72) = 801 Kombinationen

Totaler Bullshit:

<http://www.domain.de/tv/led/schwarz/sony>

<http://www.domain.de/tv/led/sony/schwarz>

<http://www.domain.de/tv/sony/schwarz/led>

Faceted Search – Fluch und Segen

Produkt-Liste filtern:

Typ:

- LCD-Fernseher (220)
- LED-Fernseher (118)
- Plasma Fernseher (22)
- Portabler Fernseher (9)
- Röhrenfernseher (1)

Bildschirmdiagonale in cm:

- 60 (17)
- 81 (12)
- 56 (6)
- 17.5 (5)
- 61 (5)

▶ [Alle 33 Bildschirmdiagonale in cm im Überblick](#)

Bildschirmdiagonale in Zoll:

- 24 (23)
- 32 (12)
- 22 (10)
- 7 (6)
- 42 (4)

▶ [Alle 27 Bildschirmdiagonale in Zoll im Überblick](#)

Marke:

- LG (113)
- SAMSUNG (67)
- Toshiba (65)
- Philips (44)
- cmx | VCM (29)

▶ [Alle 36 Marken im Überblick](#)

- Nur bewertete Produkte (5)
- Nur reduzierte Produkte (74)
- Nur neueste Produkte (22)

Farben:



Preisspanne:



Faceted Search – Fluch und Segen

Produkt-Liste filtern:

Typ:

- LCD-Fernseher (220)
- LED-Fernseher (118)
- Plasma Fernseher (22)
- Portabler Fernseher (9)
- Röhrenfernseher (1)

Bildschirmdiagonale in cm:

- 60 (17)
- 61 (17)
- 65 (17)
- 17.5 (5)
- 61 (5)

▶ [Alle 33 Bildschirmdiagonale in cm im Überblick](#)

Bildschirmdiagonale in Zoll:

- 24 (23)
- 32 (12)
- 22 (10)
- 7 (6)
- 42 (4)

▶ [Alle 27 Bildschirmdiagonale in Zoll im Überblick](#)

Marke:

- LG (113)
- SAMSUNG (67)
- Toshiba (65)
- Philips (44)
- cmx | VCM (29)

▶ [Alle 36 Marken im Überblick](#)

- Nur bewertete Produkte (5)
- Nur empfohlene Produkte (74)
- Nur neueste Produkte (22)

Farben:



Preisspanne:



Maskieren via

- Javascript
 - onClick
 - nachladen
- Formular

Pagespeed – Augen aufs Detail

- Cookies
- User (CPU, Internetverbindung)
- PHP Version (32bit vs. 64bit)
- Asynchrones JS
- Komplexität HTML, CSS, JS
(Browser-Versionen CSS3, id->class->tag)

Sitemaps – Indizierung messen

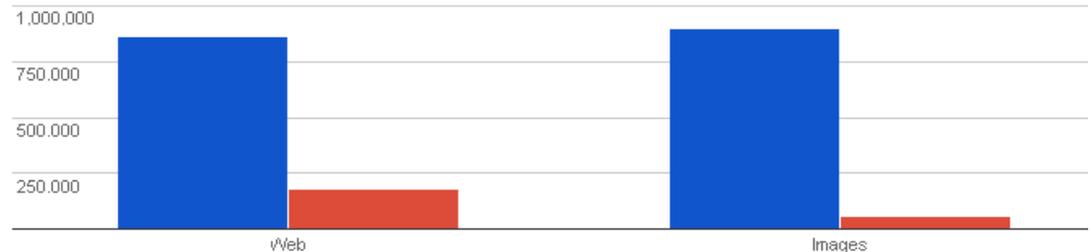
- Standard:
 - Kategorien
 - Produkte
- Besser:
 - Kategorien
 - Produkte
 - Branches
 - sonstiges

Sitemaps

By me (1) All (1)

Sitemaps content

All content types	Web pages	Images
Submitted	865,458 Submitted	897,009 Submitted
Indexed	179,951 Indexed	59,275 Indexed

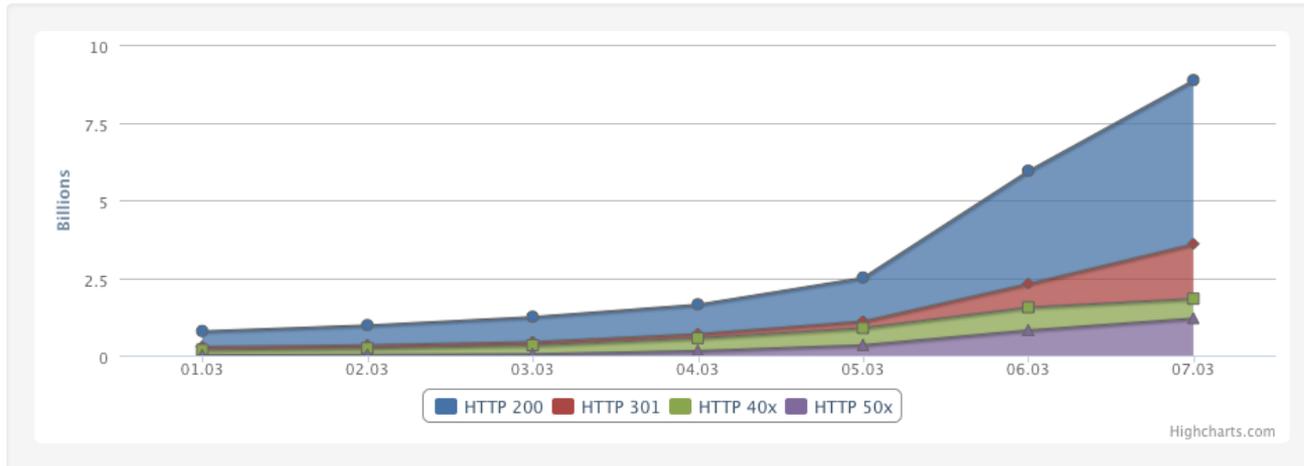


Ideen fürs eigene Tool

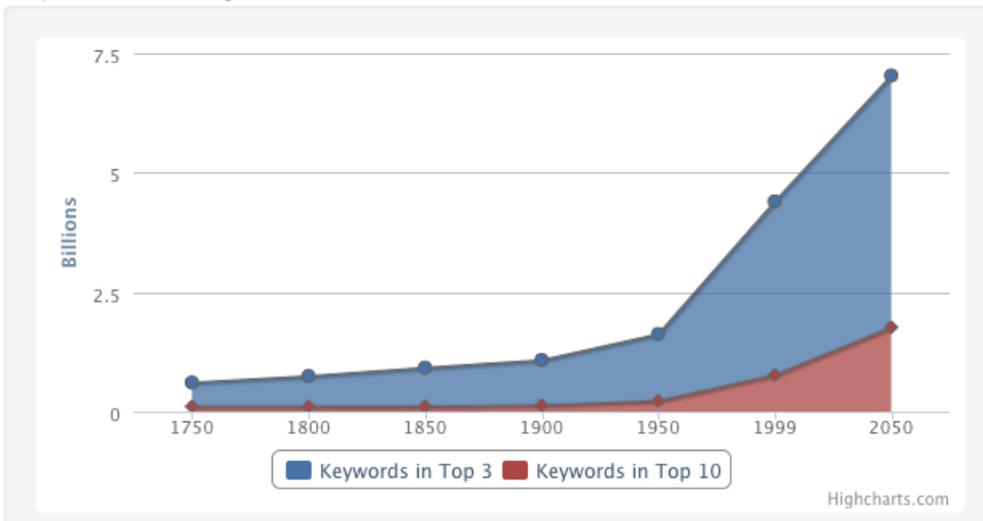
- Pageviews (Besucher / Google)
 - Logs vs. Analytics (Falle: Caching)
 - Statuscodes: 200, 301, 404, 503
- Fehlermeldungen WMT
 - 404
 - robots.txt
 - Sitemap
- Crawling WMT (nicht über API auslesbar)
- Rankings

Ideen fürs eigene Tool

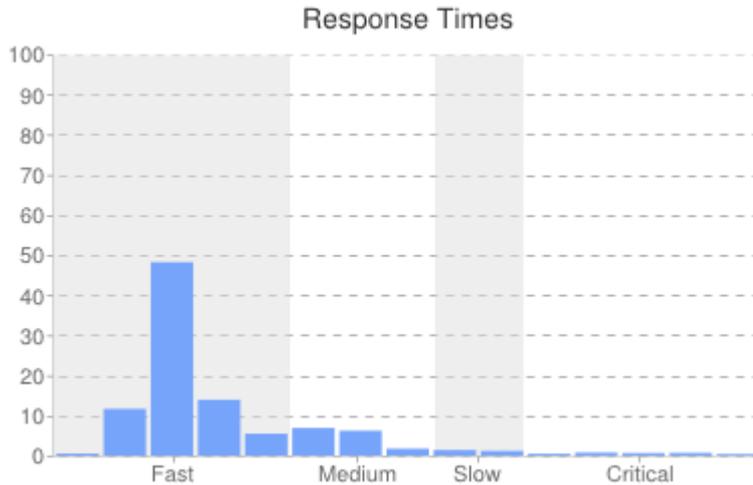
SEO traffic response codes



Keyword rankings



Analyse: Strucr.com



- Verification File notwendig
- Crawling einstellbar
 - Anzahl Pipes
 - wie viele Ebenen
 - User-Agent
 - extra robots.txt
- Kosten (25, 100, 200, 350 Euro)

Hint	Pages
<a> has malformed href ?	54
<a> links relative ?	125,405
<a> more than 100 links ?	515
<a> no internal links ?	5
<h1> and <h2> not found or empty ?	17
<h1> not found or empty ?	25
<html> is more than 100 KB in size ?	49,240
<link rel="canonical"> found ?	116,075

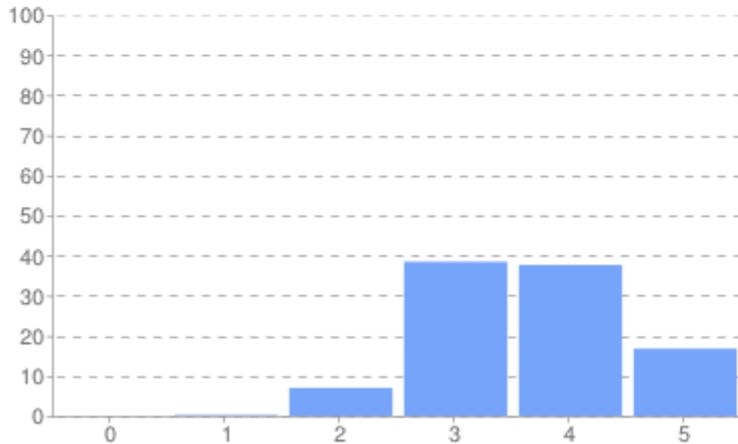
Analyse: Strucr.com

Level	Pages on Level	Multiplier	Percent on Level	Pages Total	Multiplier	Percent Total
0	1	1.0	0.00%	1	1.0	0.00%
1	293	293.0	0.23%	294	294.0	0.23%
2	9,077	31.0	6.97%	9,371	31.9	7.20%
3	50,020	5.5	38.42%	59,391	6.3	45.62%
4	48,987	1.0	37.63%	108,378	1.8	83.25%
5	21,811	0.4	16.75%	130,189	1.2	100.00%

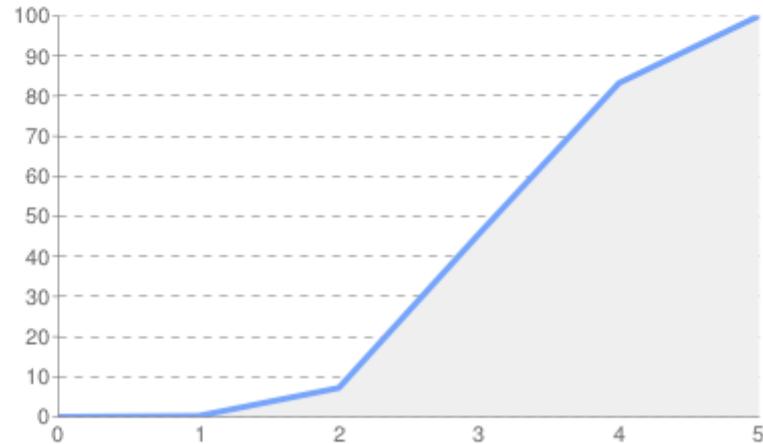
NEW [See all levels](#)

The following diagrams visualize the distribution of pages across levels.

Pages By Level In %



Pages Total By Level In %



Controlling: ServerDensity.com

- Services
 - httpasswd Login möglich
 - Headers setzbar
 - POST sendbar
 - SSL möglich
- Alerts
- Benachrichtigung per
 - eMail
 - iPhone / Android App
 - SMS

Controlling: ServerDensity.com

How often?

Every seconds

Timeout after seconds

Slow threshold milliseconds [?](#)

Check from or round-robin between all

Some other stuff (optional)

Verify SSL? [?](#)

Group

HTTP Auth Username [?](#)

HTTP Auth Password [?](#)

Headers (optional)

Key [?](#)

Value [?](#)

[Add another](#)

POST data (optional)

Key [?](#)

Value [?](#)

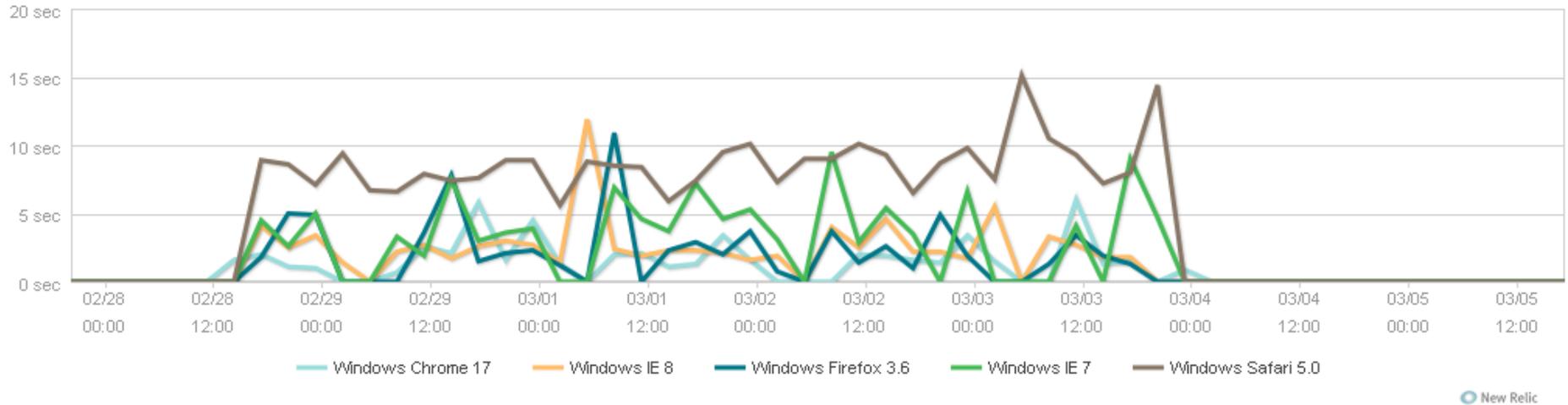
[Add another](#)

Controlling: ServerDensity.com

Check	Trigger threshold	Notification	Add new alert
HTTP - string not found (UA-189[redacted])	-	E-mail	Pause / Edit / Delete / History
HTTP - status code	> 200	E-mail	Pause / Edit / Delete / History
HTTP - string not found (meta name="robots" content="index, follow")	-	E-mail	Pause / Edit / Delete / History
HTTP - string not found (meta name="description" content="Bench Minikl...)	-	E-mail	Pause / Edit / Delete / History
HTTP - string not found (link rel="canonical" href="[redacted]....)	-	E-mail	Pause / Edit / Delete / History

Controlling: New-Relic.com

Top browsers by average front-end time

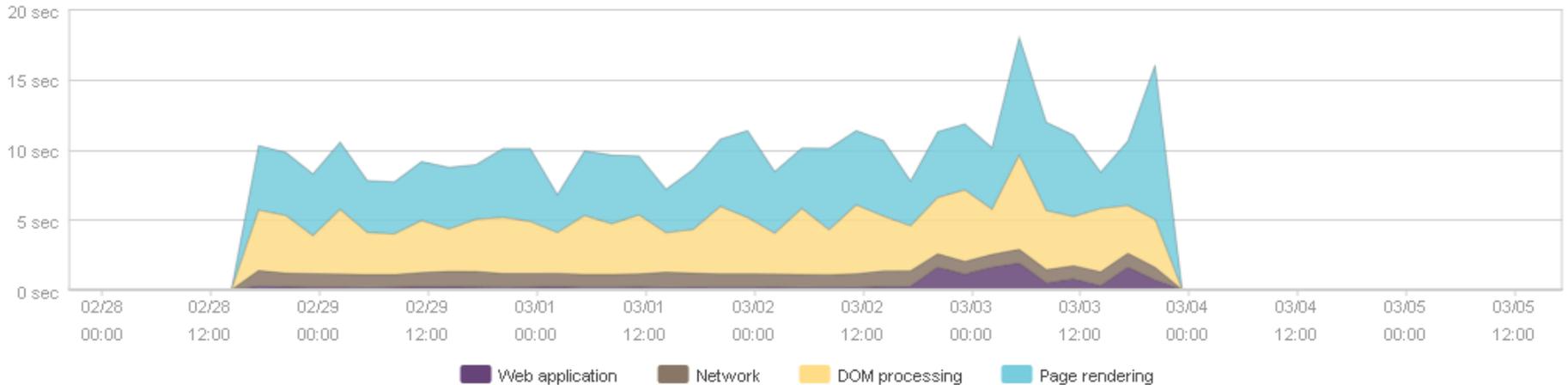


Windows Safari 5.0



Browser average page load time ?

Average: 9.7 sec



Diskussion

- Welche potenziellen Probleme kennt ihr?
- Wie steuert ihr die Indizierung?
- Maskiert ihr Links, wenn ja wie?
- Welche Tools nutzt ihr?