

# hablemos de infraestructura



# ¿por qué estoy yo aquí?



[javiercasares.com](http://javiercasares.com)



[ojobuscador.com](http://ojobuscador.com)



[kisslab.com](http://kisslab.com)



[digitalparks.com](http://digitalparks.com)

[javiercasares.com](http://javiercasares.com)

[javier.casares@kisslab.com](mailto:javier.casares@kisslab.com)

2

---

**Si eres tan listo,  
hazlo tú mismo**

# ¿los clientes no tienen ni puta idea de SEO?

18 de diciembre de 2009 13:03 (por La Jefa)

Esta mañana nos hemos encontrado con esta información en Google ¿Nos podríais decir por qué y qué recomendaciones nos dais hacer para que desaparezca?

*NOTA: el aviso es ese de «Este sitio puede provocar daños en su ordenador»*

18 de diciembre de 2009 13:11 (por Yo)

El porqué lo encontraréis en Google Webmaster Tools. Si no recuerdo mal utilizáis un Joomla, Drupal o similar y esos softwares son bastante inseguros... es probable que os hayan hackeado la nueva web...

Para arreglarlo, a parte de corregir el problema que pueda haber (como os digo, lo encontraréis en Webmaster Tools -yo tengo acceso al «com», pero no al «es», por lo que no os puedo mirar que es-).

# ¿los clientes no tienen ni puta idea de SEO?

18 de diciembre de 2009 13:41 (por El Jefe de IT)

Antes de nada comentar que Drupal no es inseguro, muestra de ello es que el site de la Casa Blanca: *www.whitehouse.gov* está hecho con Drupal.

El problema seguramente es que *www.dominio.es* tiene una página con un pequeño script que redirige a nuestra máquina de Drupal. Lo que habría que hacer (y se ha comentado muchas veces) es hacer que *www.dominio.es* apunte directamente a la máquina de Drupal y quitar esa redirección que hay puesta actualmente en .es

18 de diciembre de 2009 13:50 (por Un «Tecniquillo») manda lo siguiente...

# ¿los clientes no tienen ni puta idea de SEO?

Red Eye

"Em dia de vitória, ninguém fica cansado."

Somos: m4U3RiCk - W4N73D - HADES - T4PH0D4 - Sky - Z4i0n

Estamos apenas no começo da evolução se preparem para o que esta por vir!

By Z4i0n - Brasil - 2009

---

# el SEO ha muerto

# el SEO ha muerto

---

**año 1996**

SEO real

- meta-tags

SEO en uso

- ¿qué es SEO?

# el SEO ha muerto

---

## año 1998

### SEO real

- Altavista lee los contenidos
- densidad de palabras

### SEO en uso

- meta-tags

# el SEO ha muerto

---

## **año 1999**

### SEO real

- Google introduce el Pagerank
- enlaces entrantes

### SEO en uso

- densidad de palabras

# el SEO ha muerto

---

## **año 2001**

### SEO real

- buenos contenidos
- buenos enlaces

### SEO en uso

- decenas de veces las keywords
- granjas de enlaces

# el SEO ha muerto

---

## año 2003

### SEO real

- el Pagerank se ha hackeado
- cambio de infraestructura en Google (Florida)

### SEO en uso

- decenas de veces las keywords
- granjas de enlaces

# el SEO ha muerto

---

## año 2006

### SEO real

- Google inventa el «nofollow»
- prima el contenido sobre los enlaces

### SEO en uso

- pensamos en Pagerank y enlaces

# el SEO ha muerto

---

## año 2009

### SEO real

- el «nofollow» era una farsa contra los spammers
- «todo ha de ir muy rápido»

### SEO en uso

- blogs, flash, contenidos, tiempo real

# el SEO ha muerto

---

## **año 2011**

### SEO real

- infraestructura

### SEO en uso

- enlaces
- contenidos
- tiempo real
- «social media»

---

# ¿cómo llegamos ahí?

# quién se ha de encargar de esto



PPC



SEO



Admin.

[javiercasares.com](http://javiercasares.com)

[javier.casares@kisslab.com](mailto:javier.casares@kisslab.com)

17

# pasos para convertir nuestro sitio en el mejor

---

PASO 1: Buena **infraestructura** tecnológica

Bueno para el usuario

Bueno para SEO

Bueno para PPC

PASO 2: Buen **gestor de contenidos** y buenos **contenidos**

Bueno para el usuario

Bueno para SEO

Bueno para PPC

PASO 3: Buen **sistema de seguimiento** del usuario

Bueno para el usuario

Bueno para SEO

Bueno para PPC

---

# ¿comenzamos?



## Paso 1: **Infraestructura**

- Buena conectividad,
- Buena configuración,
- Configuración correcta de IP,
- Configuración correcta de DNS.



## Paso 1: **Infraestructura**

### Modelos:

- Máquinas propias
- Virtualización

### Distribución de Carga

- Todo en Uno
- Varios frontales + SQL



## Paso 1: **Infraestructura**

### Geolocalización:

- IP por país/territorio

### IP Multicasting

- 1 IP en muchos servidores
- No hay distribución DNS



## Paso 2: Sitio Web

- Configurar al 100% el Servidor,
- Cumplir los estándares.
  - robots.txt
  - Sitemaps
  - Redirecciones
  - Códigos Web (200, 30x, 40x, 50x)

## Paso 3: **Contenidos**



- Generar Contenidos,
- Generar Contenidos,
- Generar Contenidos,
- Generar Contenidos.



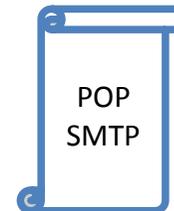
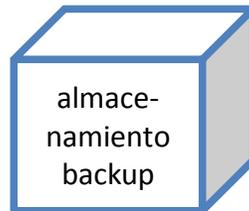
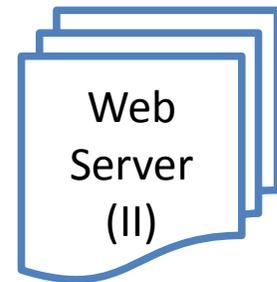
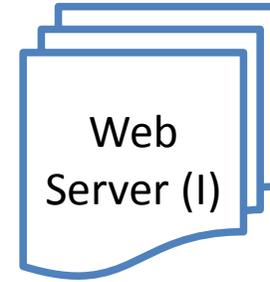
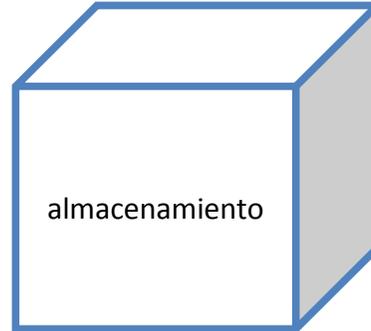
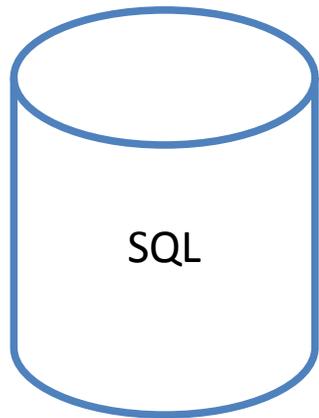
## Paso 4: **Socialización**

- Dejar que la gente te enlace,
- Enlazar a la gente con normalidad,
- Ofrecer sistemas de socialización.

---

# **¿cómo se monta una infraestructura buena, bonita y barata?**

# cómo montar infraestructura



# Cloud Computing y antiSEO

---

- El Cloud Computing permite una distribución de carga en varias máquinas, pero la automatización y desgeolocalización no ayudan al posicionamiento.
- Es mejor utilizar Cloud Privados controlados con nuestra infraestructura. Podemos distribuir carga de máquina según nos interese, incluso distribuyendo en Cloud Público algunos elementos no prioritarios.
- Servicios como Windows Azure, Amazon EC2, Amazon S2 ayudan a la velocidad pero no al SEO.

# Y! Slow y Page Speed

---

- Yahoo! y Google tienen herramientas que analizan el estado de algunos puntos que pueden mejorar el rendimiento de la indexabilidad.
- Yahoo! YSlow  
<http://developer.yahoo.com/yslow/>
- Google Page Speed  
<http://code.google.com/speed/page-speed/>

# Optimización del HTTP

---

## Reducir peticiones HTTP

- Combinar varios CSS y JS en uno
- Combinar imágenes de fondo en los CSS Sprites
- Crear mapas de imágenes en un fichero y gestionarlo por CSS
- Incluir imágenes en el propio HTML con el método «data:»

## Usar sistemas CDN (Content Delivery Network)

- Distribución geográfica de contenidos estáticos

# Optimización del HTTP

## Cabeceras con control de caché y expiración

- Para contenidos estáticos usar un «Never Expires»
- Para contenidos dinámicos usar un «Cache-Control»
- Usar «Cache control: public» para cachear HTTPS
- En Apache usar [mod\\_cache](#), mod\_disk\_cache, mod\_mem\_cache, mod\_file\_cache y htcacheclean

## Comprimir los contenidos posibles

- Usar el protocolo «HTTP/1.1»
- Aceptar codificación «Accept-Encoding: gzip, deflate»
- Aceptar contenido comprimido «Content-Encoding: gzip»

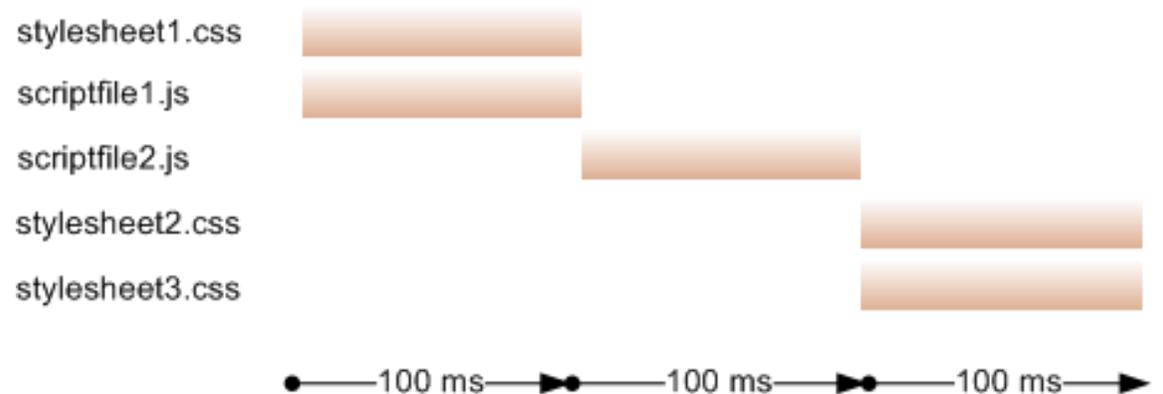
# Optimización del HTTP

## Dominios sin cookies

- Tener dominios sin cookies para estáticos

## Paralelizar peticiones HTTP

- Usar «non-blocking scripts» para poder descargar JS sin que se frene el resto.



# Optimización del HTTP

---

## Externalizar contenidos CSS y JS

- En ficheros .CSS y .JS que son cacheables
- Mejor usar el tag <link> en vez de @import para los CSS

## Incluir los CSS en la parte superior

- Las páginas irán generándose progresivamente en su carga

## Incluir los JS en la parte inferior

- No se recomiendan más de dos peticiones en paralelo
- Intentar usar el atributo «DEFER» para scripts que no tengan «document.write»

# Optimización del HTTP

---

## Reducir las peticiones DNS

- Intentar no hacer llamadas externas a otros dominios
- Cachear scripts externos como Analytics o Adsense

## Evitar redirecciones

- Las redirecciones 301 o 302 no suelen ir acompañadas de «Expires» o «Cache-Control»
- Usar «Alias», «mod\_rewrite» o «DirectorySlash» para automatizar el «/» al final de URL

# Optimización del HTTP

---

## Usar ETags

- Devolver la cabecera «Etag» con un identificador único y cachear contenidos en Apache
- Usar el «Filetimestamp:ChangeNumber» en IIS
- Las peticiones siguientes devuelven un «304 Not Modified»
- Si no se puede usar «Etag», utilizar «Last-Modified»

## Devolver código parcialmente

- Usar la función «flush()» en determinados momentos para una carga de página mayor (por ejemplo, antes del <body>)

# Optimización del HTTP

---

Evitar enlaces a contenidos «404 Not Found»

- Si un JS o CSS llama a una petición «404 Not Found» la carga de otros elementos se ralentizará

Reducir el tamaño de las cookies

- Eliminar cookies que no sean necesarias
- Reducir el tamaño de las cookies al mínimo posible
- Aplicar las cookies al nivel de dominio-subdominio necesario
- Aplicar una fecha de eliminación ni muy lejana ni muy temprana

# Optimización del CSS y JavaScript

## Evitar programación en los CSS

- Internet Explorer permite ciertos JS en los CSS
- Evitar filtros tipo «AlphaImageLoader» en los CSS

## Minimizar CSS y JS

- Reducir el tamaño de CSS y JS con sistemas como «[JSMin](#)» o «[YUI Compressor](#)» (21% - menor riesgo)
- Intentar ofuscar el código (25% - mayor riesgo)
- Unificar todos los JS en un único JS
- Unificar todos los CSS en un único CSS

# Optimización del CSS y JavaScript

---

## Eliminar scripts duplicados

- Usar un sistema programado que sólo mande una vez el script y lo cachee

## Carga de elementos

- Gracias a CSS y Javascript se pueden cargar imágenes u otros textos una vez cargada la página
- Anticipar la carga de elementos antes de siguientes navegaciones (imágenes que se usarán en siguientes páginas...)

# Optimización del CSS y JavaScript

## Optimizar los CSS Sprites

- Es mejor organizar las imágenes en horizontal en vez de vertical
- Combinar mejor los colores para que la imagen quede con GIF 256 o PNG8
- No dejar espacios entre las imágenes y aprovechar al máximo todos los pixels



# Optimización de otros elementos

---

## Reducir el número de elementos DOM

- Un sitio puede cargar por defecto varios miles de elementos DOM que se pueden reducir a menos de 1.000

## Reducir el número de <iframe>

- La página de destino realizará muchas peticiones que ralentizan la carga.

## No escalar las imágenes

- En vez de usar una imagen de 500x500 reducida a 100x100, usar una de 100x100 directamente

# Optimización de otros elementos

## Optimizar las imágenes

- Reducir la cantidad de colores de la paleta de los GIF (por ejemplo, con «[imagemagick](#)»)
- Intentar convertir los GIF en PNG si reducen su tamaño (por ejemplo, con «[imagemagick](#)» un 20%)
- Optimizar los GIF animados (por ejemplo, con «[GIFsicle](#)»)
- Optimizar los PNG (por ejemplo, con «[pngcrush](#)» un 15%)
- Optimizar los JPG (por ejemplo, con «[jpegtran](#)» un 10%)
- Si los JPG pesan menos de 10KB usar «baseline»
- Si los JPG pesan más de 10KB usar «progressive»

# Optimización de otros elementos

---

Usar un «favicon.ico» pequeño y cacheable

- Evitar que el «favicon.ico» de un error «404 Not Found»
- Reducir su tamaño a menos de 1 KB (por ejemplo, con «[Imagemagick](#)»)
- Añadir una cabecera «Expires» para muchos meses

Devolver contenidos de menos de 25KB

- Algunos terminales como iPhone sólo son capaces de cachear 25 KB por elemento
- Cuidar el «gzip» porque tal vez no sea sólo suficiente

---

**no sólo de  
infraestructura  
vive el SEO**

## algunas cuestiones técnicas



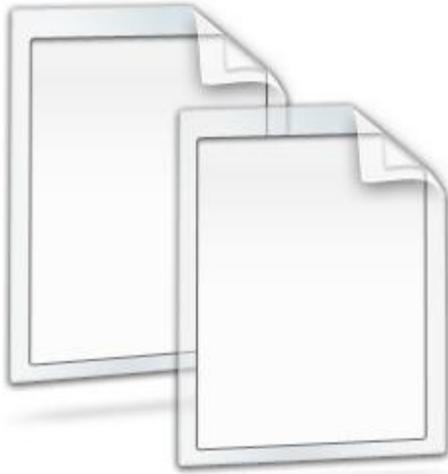
- Intentar tener una IP lo más exclusiva posible, geolocalizada correctamente,
- Configurar las DNS correctas, en el país correcto,
- Usar un servidor web que permita «rewriting» de direcciones URL,
- «Invertir» dinero en una buena infraestructura.

# que las cosas se hagan correctamente



- Utilizar correctamente el robots.txt,
  - no usar Disallow
  - usar Sitemaps
- Comprobar que existe una página de error 404, y si está diseñada, mejor aún,
- Intentar realizar ataques a través de las URL del navegador, añadiendo parámetros o modificando elementos,
- Usar redirecciones de forma controlada, aunque es mejor no tener que usarlas.

# contenidos duplicados



- Hay que evitar tener dos páginas o más iguales,
- Iguales no significa que parte de un contenido esté en otro sino
  - accesos de login en varias URL
  - páginas de resultados, sin resultados
- Títulos, meta-descripciones y URL, únicas por contenido.

```
<meta name="robots" content="noindex" />
```

- Títulos, meta-descripciones y URL, únicas por contenido.

```
<link rel="canonical"  
href="http://example.com/pagina-concreta/" />
```

# los paneles de control de los buscadores



- Google:

<http://www.google.com/webmasters/tools/>

- Yahoo!:

<http://siteexplorer.search.yahoo.com/>

- Bing:

<http://www.bing.com/webmasters/>

- Yandex:

<http://webmaster.yandex.ru/wmconsole/>

# dominios, subdominios o carpetas



- Los dominios geolocalizados (.tld) posicionan mejor en el país en cuestión,
- Si hay que organizar información por idioma, mejor en subdominios, nunca en carpetas,
- Si hay que crear comercio, mejor utilizar los dominios geolocalizados para cada país,
- Cada proyecto es un mundo.

- Si hay que redirigir, que sea con una **Redirección 301**.



ASP:

```
<%  
Response.Status="HTTP/1.1 301 Moved Permanently"  
Response.AddHeader "Location", "http://www.nuevo.ext/"  
%>
```

PHP:

```
<?php  
$url = "Location: http://www.nuevo.ext" . $_SERVER["REQUEST_URI"];  
header($url, true, 301);  
?>
```

Apache:

```
RewriteCond %{THE_REQUEST} ^GET /*/index.(php|html|html) HTTP  
RewriteRule (.*)index.(php|html|html)$ /$1 [R=301,L]
```

- Cuando queremos que no se indexen páginas que no existen en una migración.

ASP:

```
<%  
Response.Status="HTTP/1.1 503 Network Unavaible"  
Response.Status="Retry-After: 3600"  
%>
```

PHP:

```
<?php  
header("HTTP/1.1 503 Service Unavailable");  
header("Retry-After: 3600");  
?>
```



# el Gran Hermano te vigila

- El buscador no vive de Pagerank, sino de otros software espía.



- Barras del navegador

- Sistemas de analítica web

- Usuarios registrados en buscador



- Contenidos duplicados
- Abusar
- Obsesionarse con los enlaces



- Tecnología
- Estándares
- Contenidos



¿Quién hace los buscadores?

Ingenieros

Así que piensa como un ingeniero de Google que está en Googleplex y conseguirás ser un buen SEO.

Recuerda que los robots de búsqueda son máquinas y que siguen lo que se les ha programado.

---

# casos de estudio

## sitio de subastas inversas

---

Necesita tráfico tanto de los que ofrecen la oferta como los que quieren conseguir el trabajo

Lanzado en verano de 2008, llegó a navidad con una caída de tráfico del 80% por “razones desconocidas”.

La revisión del sitio a nivel tecnológico concluyó que su problema principal eran los contenidos duplicados.

La efectividad se vio al cabo de los 10 días de aplicar el informe.

# sitio de subastas inversas

## Motores de búsqueda

14/01/2009 - 10/02/2009



La búsqueda ha enviado 5.164 visitas sin pagar a través de 11 fuentes.

Mostrar: [total](#) | [pagado](#) | [sin pagar](#)

20 de enero

108 visitas/día

9 de febrero

436 visitas/día

incremento x4

javiercasares.com

javier.casares@kisslab.com

57

## sitio de leads de formación

---

Necesita tráfico de gente interesada en formación.

El tráfico se había estancado desde hacía meses. Sus anteriores SEO habían dejado algunos problemas.

La revisión del sitio a nivel tecnológico concluyó que su problema principal eran los contenidos duplicados y mala arquitectura de la información.

La efectividad se vio al cabo de los 30 días de aplicar el informe, que, junto a un análisis de Analítica Web, incrementó ventas.

# sitio de leads de formación

## Motores de búsqueda

01/02/2009 - 30/06/2009



**La búsqueda ha enviado 3.625.605 visitas sin pagar a través de 21 fuentes.**

Mostrar: [total](#) | [pagado](#) | [sin pagar](#)

3 de febrero

25.685 visitas/día

22 de junio

38.274 visitas/día

incremento x1.5

javiercasares.com

javier.casares@kisslab.com

59

# tienda e-commerce venta de videojuegos y gadgets

---

Cuanto más tráfico, más posibilidades de venta

El sitio iba subiendo y bajando de visitas sin sentido aparente

Una revisión hizo concluir que el problema estaba en los enlaces entrantes (linkbuilding) que se había hecho artificial y desde sitios nada apropiados.

La eliminación de esos enlaces (del propio cliente) hizo que el tráfico subiera inmediatamente.

# tienda e-commerce venta de videojuegos y gadgets

## Motores de búsqueda

18/05/2009 - 21/06/2009



La búsqueda ha enviado 10.232 visitas sin pagar a través de 11 fuentes.

Mostrar: [total](#) | [pagado](#) | [sin pagar](#)

26 de mayo

157 visitas/día

9 de junio

610 visitas/día

incremento x4

javiercasares.com

javier.casares@kisslab.com

61

## foro de ayuda sobre tecnología

---

Imagen de marca que lleva tráfico a otras zonas del sitio

Una actualización de un software no actualizado en 5 años hizo que salieran a la luz algunos problemas.

El tráfico cayó en picado debido a problemas de contenidos duplicados y a decenas de miles de páginas de error.

Tras reducir las páginas de error se fue recuperando el tráfico.

# foro de ayuda sobre tecnología

## Motores de búsqueda

15/02/2009 - 15/10/2009



La búsqueda ha enviado 583.994 visitas sin pagar a través de 17 fuentes.

Mostrar: total | pagado | sin pagar

25 de febrero	5.342 visitas /día	
28 de abril	945 visitas /día	decremento x5.5
29 de julio	2.750 visitas /día	incremento x3
6 de octubre	5.441 visitas /día	volvemos, al 100%

javiercasares.com

javier.casares@kisslab.com

63

---

# ¿alguna pregunta?

