

Fecha: 17/09/2019

Este documento es una guía para saber qué decisiones tomar a la hora de resolver las situaciones más confusas que se encuentren en los errores. No se enlista todos los errores, sino por tipo de error o por situación.

## Errores 404

Qué son? Páginas que en su momento existieron y ya no están live, pero siguen siendo enlazadas desde otra página interna o externa.

Documento: [Páginas con código de estado 4xx](#)

**Primer caso:** comentarios generados en páginas inexistentes

Enlace roto: <https://ed.team/quieres-ganarte-un-ano-gratis-a-edteam>

Página que lo enlaza:

<https://ed.team/publicacion?slug=quieres-ganarte-un-ano-gratis-a-edteam>

El enlace roto está siendo enlazado desde un comentario de su misma página. Pero cada comentario genera una url y sigue estando viva.

Solución inmediata: eliminar la página que enlaza el enlace roto.

Solución definitiva: tecnológicamente diseñar una solución que, cuando se elimine una página, además de eliminarse del sitemap, también se eliminen todos los comentarios generados a partir de dicha página.

**Segundo caso:** enlaces mal enlazados por error humano

Enlace roto: <https://ed.team/github>

Página que lo enlaza: <https://cocha2018.ed.team/>

Solución: simplemente colocar el enlace correcto o eliminarlo, junto con el ícono, de todas las páginas donde se presenta el error.

**Tercer caso:** imágenes inexistentes enlazadas desde el head.

Enlace roto: <https://ed.team/static/images/pwa/edteam-192.png>

Página que lo enlaza: <https://ed.team/>

Solución: reemplazar el enlace de la imagen por una existente. También ya existe una (<https://ed.team/static/images/pwa/edteam-512.png>), así que lo mejor sería evaluar si es necesaria dicha imagen. Si no lo es, eliminarla de toda página donde aparece en el documento aquí citado, si es necesaria dicha imagen, enlazarla correctamente.

**Cuarto caso:** archivos css y js rotos

Enlaces rotos:

1. <https://ux.ed.team/css/prism.css?201803301710>

2. <https://ux.ed.team/js/prism.js?201803301710>

Página que lo enlaza: <https://ux.ed.team/>

Enlazar al archivo correcto o eliminar la solicitud si no es necesario.

**Quinto caso:** imagen inexistente de un perfil de usuario

Enlace roto:

<https://api.ed.team/files/https://avatars1.githubusercontent.com/u/54405720?v=4>

Página que lo enlaza: <https://ed.team/perfil/98620>

Solución: cambiar la imagen por la real, de lo contrario asignar un avatar genérico o random para evitar que existan imágenes rotas o páginas con estado 404, que es lo mismo.

## Páginas con código de estado 5xx

Documento:

<https://docs.google.com/spreadsheets/d/1MMk2QMu9Mo1-WwFfiasnFndP5ChPuSSqRovUD5J93GU/edit#gid=451386669>

Comentario: por algún motivo, que puede ser distinto en cada caso, las urls en el documento citado presentan un error 5xx que es una falla desde el servidor o un error humano y como el recurso no existe en el servidor se reporta como tal.

Solución: cambiar el enlace erróneo por el correcto o eliminar la solicitud del recurso en las páginas donde lo enlazan.

## Contenido duplicado por incompleta migración desde http a https

Documento:

<https://docs.google.com/spreadsheets/d/1ZFqD3q1ENnArbEMs0MYBHJd4W4BH2U47aMnS4wyWFrw/edit>

Comentario: existen pocos recursos con protocolo http que están siendo enlazados desde páginas https.

Solución:

- Migrar dicho recurso con protocolo sin ssl a su versión con ssl.
- Cambiar el enlace del recurso por el nuevo.
- Eliminar el recurso con versión http cuando el proceso esté completo.

## Redirecciones 302

Documento:

[https://docs.google.com/spreadsheets/d/1NzN1H9uBelIBHrguGGVtKE4DgGtL\\_P7uh-eeitAzEGo/edit](https://docs.google.com/spreadsheets/d/1NzN1H9uBelIBHrguGGVtKE4DgGtL_P7uh-eeitAzEGo/edit)

Comentario: no está mal que tengamos redirecciones. Pero en este caso el código es el incorrecto. El código de estado 302 es cuando el recurso está siendo redirigido temporalmente a otra página, porque luego volverá a estar disponible.

Solución: en estos caso la redirección es definitiva, así que lo propio sería hacer la redirección con código de estado 301 y no 302.

## Cadenas de redirecciones

Documento:

[https://docs.google.com/spreadsheets/d/19iTCg64LzBNYWdv84N8GJO4Qa\\_RWlqCDPet6oStDy\\_Y/edit#gid=2047696491](https://docs.google.com/spreadsheets/d/19iTCg64LzBNYWdv84N8GJO4Qa_RWlqCDPet6oStDy_Y/edit#gid=2047696491)

Comentario: esto quiere decir que la URL 1 redirige a la URL 2 y la URL 2 redirige a la URL 3.

Solución: hacer que URL 1, redirija a la URL final. Y la URL 2, también redirija a la URL final.

## No existe el archivo robots.txt

Esto no es un error realmente. Pero es un archivo que no permite la lectura del contenido a los robots a páginas o paths concretos. Hay ciertas páginas donde es probable que nosotros no queramos que acceda a dicho contenido, además de otros métodos.

Si existen páginas que no quieren que los robots (en este caso Google) no acceda al contenido, deberán crearlo. De ser así, me contactan y los ayudo, sin costo extra.

URLs que se deberían bloquear por robots.txt y servidor: documento.

## Etiqueta rel="canonical"

Documento:

<https://docs.google.com/spreadsheets/u/0/d/1RQS3BiUdFktZB32flwebeUkrftYBShvkEeulv0PFVqg/edit?usp=drivesdk>

Comentario: Solamente 50 páginas tienen la etiqueta rel="canonical". Esto sirve para indicar a Google cual es la verdadera (canon = original) página, ya que por diferentes motivos se generan versiones de la misma página.

Ejemplo real de EDteam:

1. <https://ayuda.ed.team/category/6-pagos?sort=popularity>
2. <https://ayuda.ed.team/category/6-pagos/>

La verdadera página es la URL 2, la URL 1 es la misma página, solo que ordena los items en base a su popularidad.

Es una práctica común hacer que todas las páginas que **no** tengan URLs generadas a partir de sí mismas, contengan la etiqueta rel=canonical. Ejemplo, la homepage de EDteam. En el head debería tener <link rel="canonical" href="<https://ed.team/>">. Pero esto no es obligatorio.

Lo que sí es obligatorio que toda página que tenga diferentes versiones (ordenadas por popularidad, alfabéticamente, comentarios, etc), tengan la etiqueta canonical para referenciar cual es la página original y que Google debe indexar.

Y como actualmente solamente hay 50, deberían hallar tecnológicamente una manera de poder implementarlo para evitar este error. Si esto no se hace, están duplicando contenido.

## Errores varios a partir de URLs dinámicas

Documento:

<https://docs.google.com/spreadsheets/d/1Vi6adNEiuQIWkh-NSR0IV-vwxKw89KCUP682-a5PyOE/edit#gid=1485216552>

Comentario: estas URLs se generan a partir de filtros, comentarios, paginaciones, etc. Son variaciones de las páginas canon u originales. En el documento citado se encuentran varias URLs dinámicas con diferentes errores y sus soluciones también allí indicadas.

## URLs demasiado largas

Documento:

[https://docs.google.com/spreadsheets/d/1kFno4kbH-1n0S9SIImJI8N4\\_LiIDVKb80npP4OixYak/edit?usp=drive\\_open&oid=108905468882640061457](https://docs.google.com/spreadsheets/d/1kFno4kbH-1n0S9SIImJI8N4_LiIDVKb80npP4OixYak/edit?usp=drive_open&oid=108905468882640061457)

Comentario: la gran mayoría de estas páginas son post generados por los usuarios que es difícil de controlar. Una manera de hacerlo es asignar un patrón de URL. Ejemplo: root/perfil/fecha/hash.

Esto, además de solucionar con el problema de URLs largas, también hará que si un miembro de la comunidad postea un artículo que puede competir con una página interna

que generen beneficios, como la url no está SEO optimizada, será más difícil que supere en los rankings a la página nos interesa posicionar.