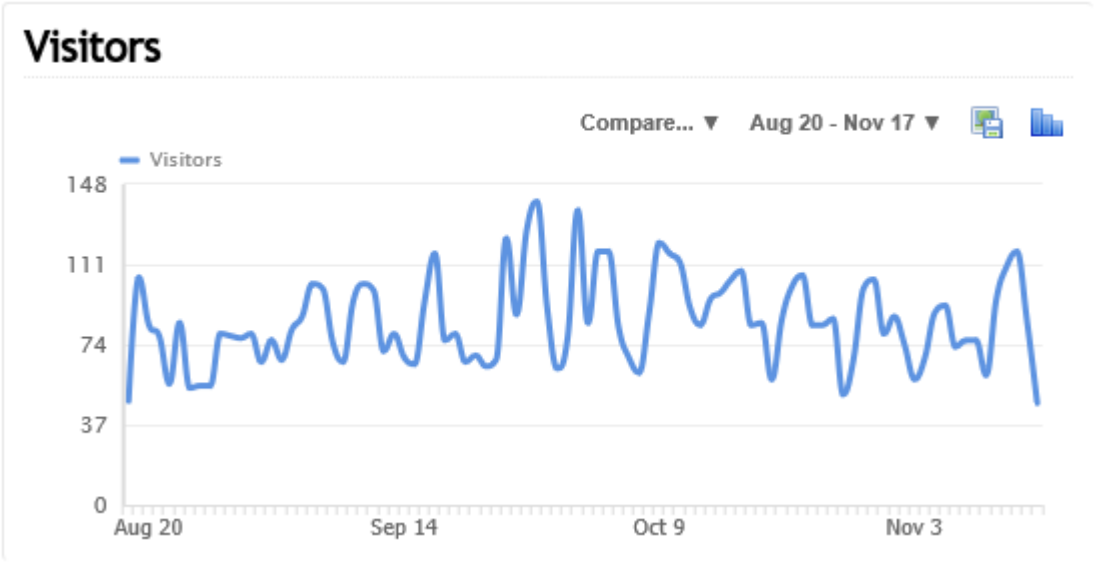
How to use Web Text Analyzer to look at the text of a home page

This tutorial will look at analyzing the text on a page. We will use a real website - http://sleep-deprivation.com/ . This site receives around 100 unique visits a day and is well established. There have been no changes to the site for at least 2 years.



The top keywords bringing traffic to the site over the last 90 days are –

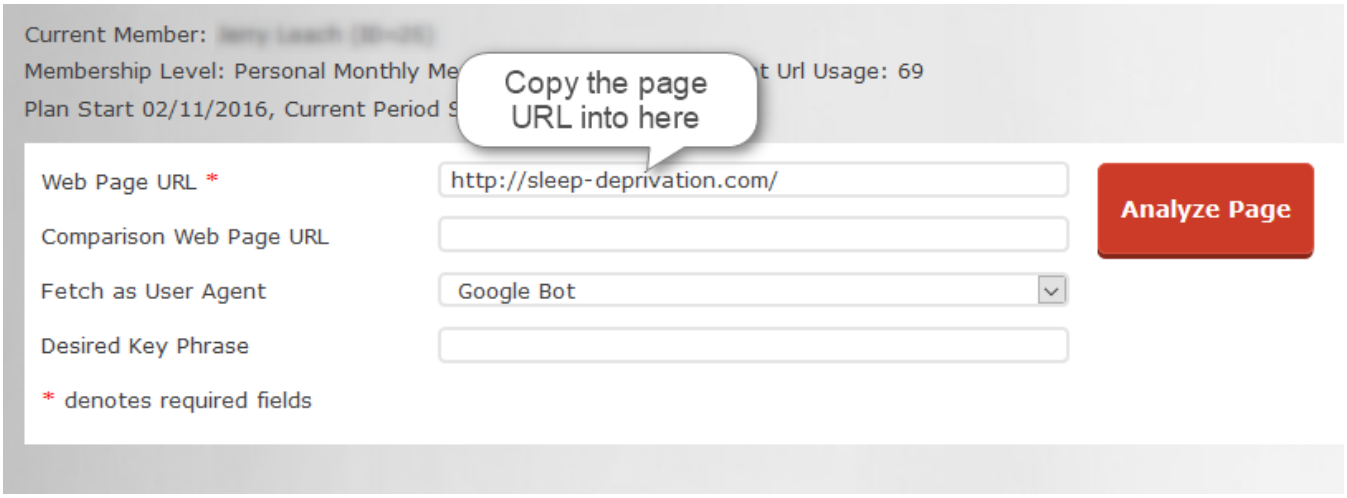
[secure search]	5,499	+46%
sexsomnia	47	+124%
sexsomnia cure	37	+825%
sexsomnia causes	26	+999%
sexsomnia treatment	21	+17%
Sexsomnia Symptoms	21	-34%
sleep deprivation	13	-75%
sleep sex disorder symptoms	6	-45%
what causes sexsomnia	6	-33%
sleep deprivation treatment	6	0%

The focus keyword for the home page currently is ‘sleep deprivation’.

Keyword rankings on Google.com as at 17th November 2016 show that 'sleep deprivation' is on page 2 (position 11) and some of the best performing keywords are around a subset of Sleep Deprivation (sexsomnia).

	Keyword	Google Rank	Ranking page(s)
1	sexsomnia causes	1	http://sleep-deprivation.com/sleep-disorder-symptoms/sleep-sex-se
2	sexsomnia symptoms	1	http://sleep-deprivation.com/sleep-disorder-symptoms/sleep-sex-se
3	sleep-deprivation.com	1	http://sleep-deprivation.com/
4	sexsomnia treatment	2	http://sleep-deprivation.com/sleep-disorder-symptoms/sleep-sex-se
5	sexsomnia cure	3	http://sleep-deprivation.com/sleep-disorder-symptoms/sleep-sex-se
6	sleep deprivation treatment	7	http://sleep-deprivation.com/sleep-deprivation-treatment/
7	home sleep deprivation	10	http://sleep-deprivation.com/sleep-deprivation-treatment/
8	sleep deprivation	11	http://sleep-deprivation.com/
9	sexsomnia	14	http://sleep-deprivation.com/sleep-disorder-symptoms/sleep-sex-se
10	sleep disorder causes	33	http://sleep-deprivation.com/
11	sleeping disorders	38	http://sleep-deprivation.com/

Let's now use the Web Text Analyzer (WTA) tool to analyze the words on the home page. Enter the URL into WTA and click 'Analyze'.



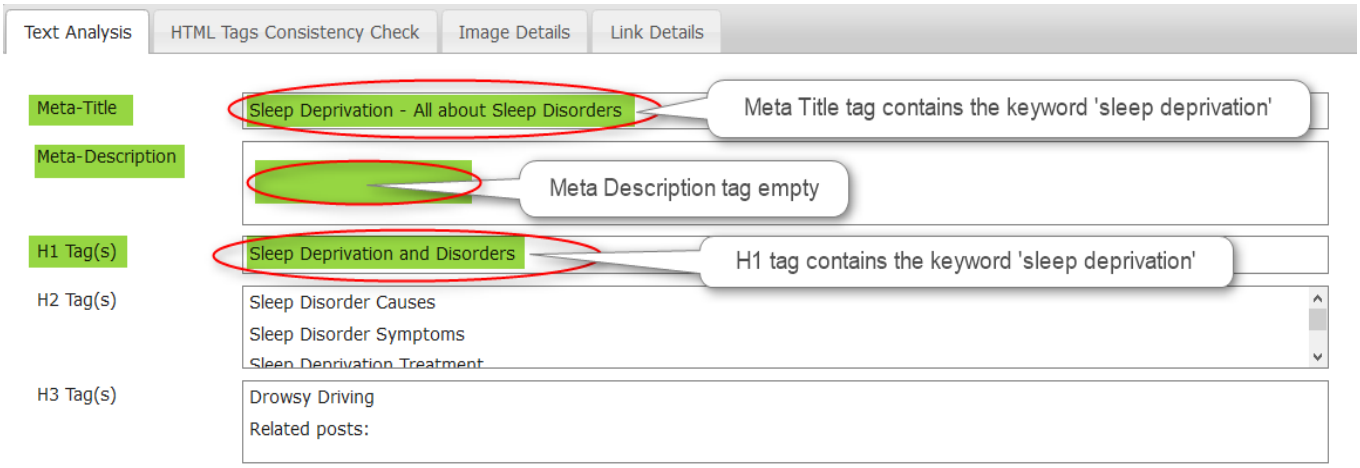
At this stage we just want to have a look at the text on the page. The default 'Bot' is Google Bot.

Running Analyze brings up the results sections. Let's now look at each of these in turn.

Text Analysis

This is in 4 sections; the first section displays the ‘meta tags’ which are essential for helping search engines in understanding what the keyword for the page is. Meta Title, Meta Description and H1 Tag(s) should never be empty and should contain the keyword for the page. The H2 and H3 tags are less important but can be used to re-inforce a keyword.

So for the sleep-deprivation site, we have –



Straight away I can see there is no Meta Description so Google will try to pick some relevant text on the page to make one. This will probably create a poor message for attracting click-through’s. When writing this text, it should contain the keywords ‘Sleep Deprivation’ and the message should make people want to click on the link.

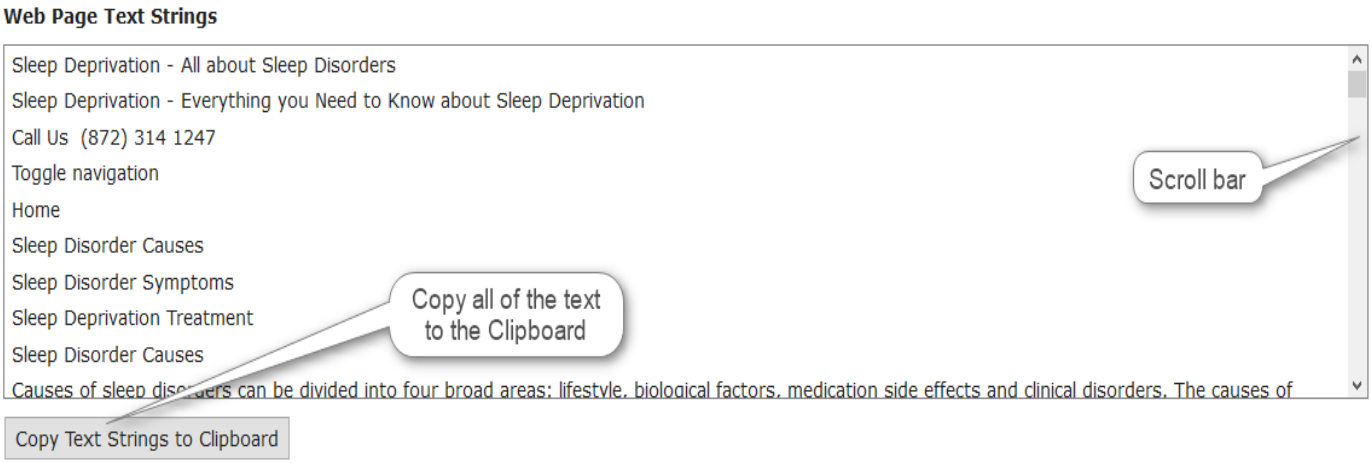
I would also change the hyphen in the Meta Title after ‘Sleep Deprivation’ to a pipe ‘|’ which clarifies this term sits on its own –

‘Sleep Deprivation | All about Sleep Disorders’.

The H1 tag is good for the term ‘Sleep Deprivation’.

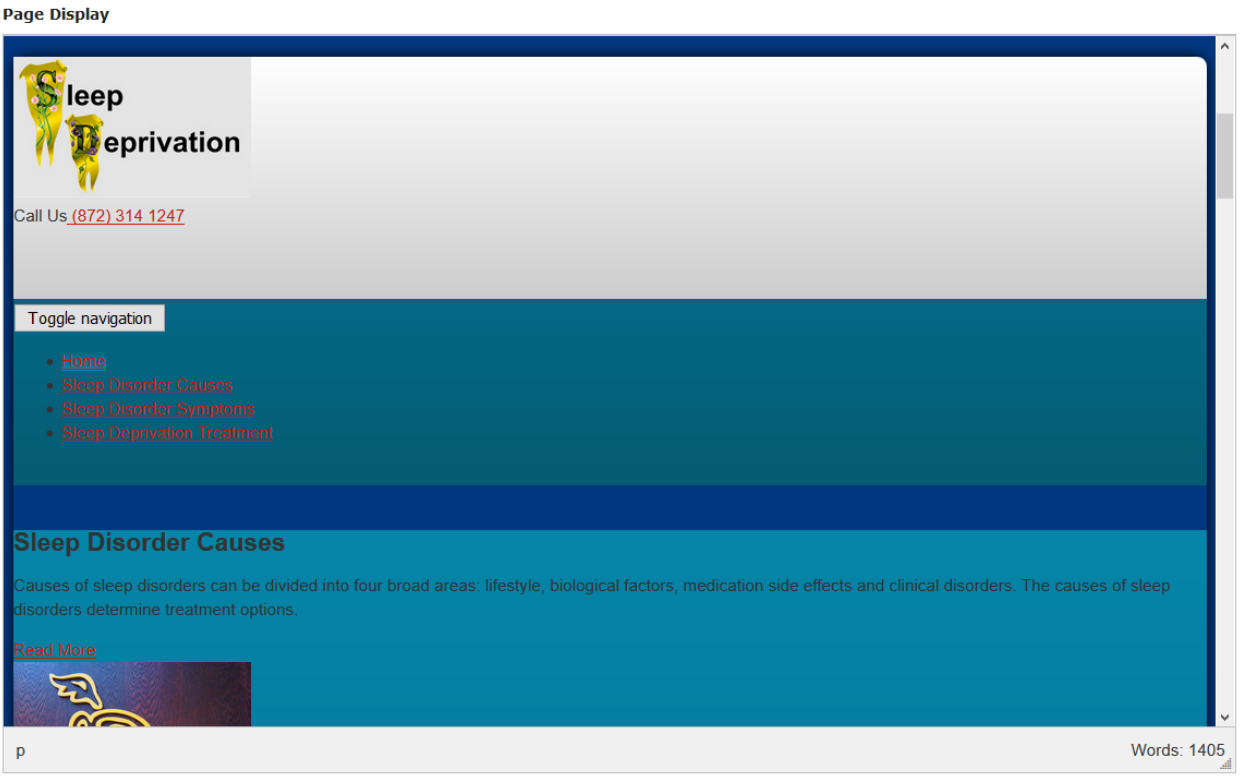
Web Page Text Strings

This section (scrollable) shows ALL of the text on the page. This includes all the text in the header and footer of the page plus all the menu text. It is displayed as a list starting from the very top of the web page. This makes it very easy to find issues such as spelling and grammar and duplicate lines of text and text that is not displayed on the web page but is visible to the search engines.



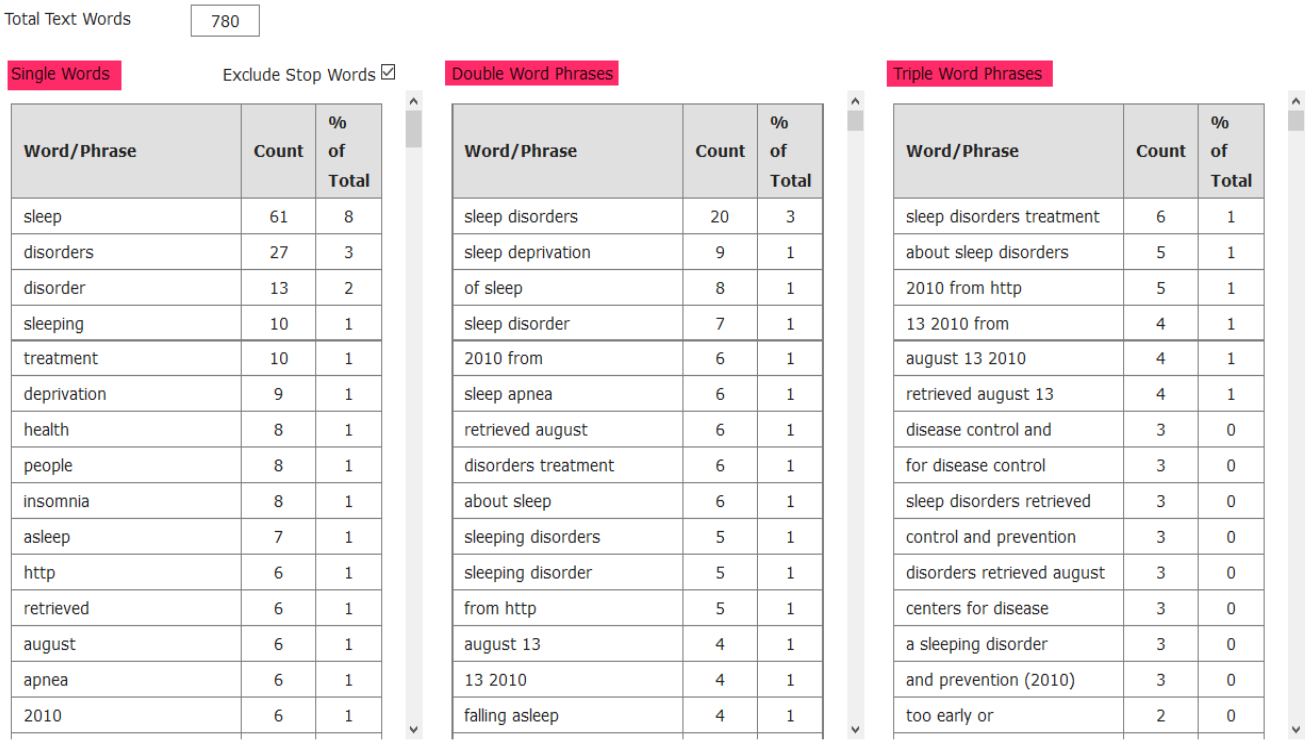
Page Display

The page displayed as an HTML page.



Total Text Words

These tables show a breakdown of every single word, two word and three word phrases on the page listed by the number found in descending order. There is also a total word count for the page. There is an option to exclude ‘stop words’ such as ‘a’, ‘the’, ‘and’ etc. which is turned on by default. These words would normally be ignored by search engines.



Total Text Words 780

Single Words | Exclude Stop Words ☑ | Double Word Phrases | Triple Word Phrases

Word/Phrase	Count	% of Total
sleep	61	8
disorders	27	3
disorder	13	2
sleeping	10	1
treatment	10	1
deprivation	9	1
health	8	1
people	8	1
insomnia	8	1
asleep	7	1
http	6	1
retrieved	6	1
august	6	1
apnea	6	1
2010	6	1

Word/Phrase	Count	% of Total
sleep disorders	20	3
sleep deprivation	9	1
of sleep	8	1
sleep disorder	7	1
2010 from	6	1
sleep apnea	6	1
retrieved august	6	1
disorders treatment	6	1
about sleep	6	1
sleeping disorders	5	1
sleeping disorder	5	1
from http	5	1
august 13	4	1
13 2010	4	1
falling asleep	4	1

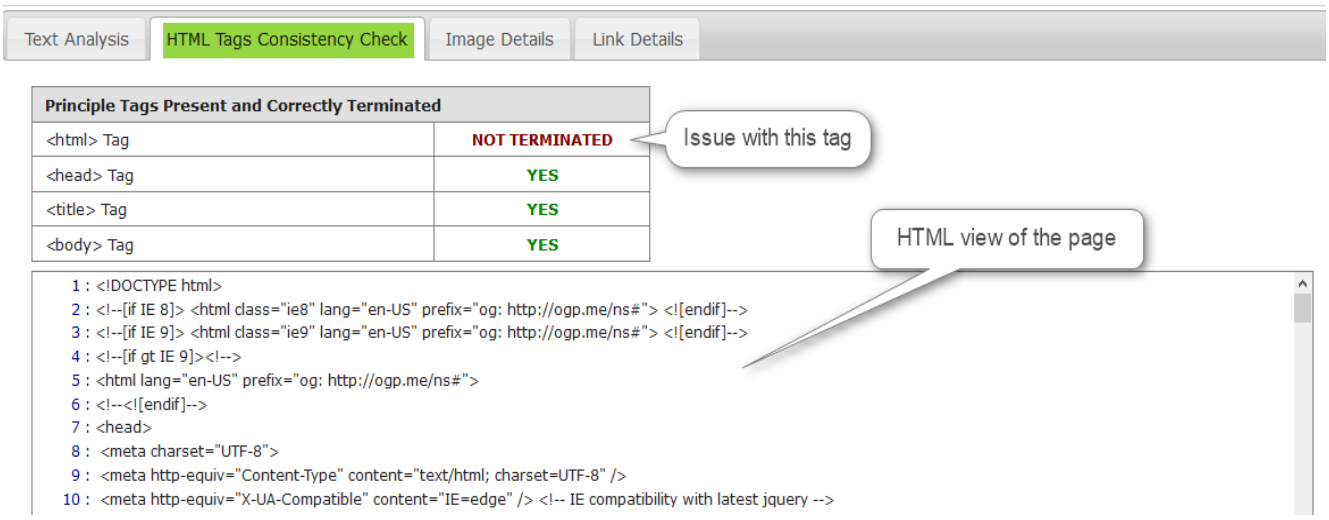
Word/Phrase	Count	% of Total
sleep disorders treatment	6	1
about sleep disorders	5	1
2010 from http	5	1
13 2010 from	4	1
august 13 2010	4	1
retrieved august 13	4	1
disease control and	3	0
for disease control	3	0
sleep disorders retrieved	3	0
control and prevention	3	0
disorders retrieved august	3	0
centers for disease	3	0
a sleeping disorder	3	0
and prevention (2010)	3	0
too early or	2	0

This shows that both the Single word and the Double Word phrases for this page are dominated by the terms ‘sleep’, ‘disorders’ and ‘sleep disorders’. The keyword term for the page ‘sleep deprivation’ has less than half the number of occurrences.

HTML Tags Consistency Check

This displays the full page in a scrollable window as HTML. The main tags are checked to ensure if they are opened the equivalent tag is closed. We will be adding more tag checks to this shortly such as checking tags.

The figure below shows there is a big problem on the sleep-deprivation page for the <html> tag which is not being closed.



Scrolling to the bottom of the code displayed shows there is no </html> tag –

This is a simple fix.

The tags check which we will add shortly is important for text; non-closure of a div can cause the text to be ignored.

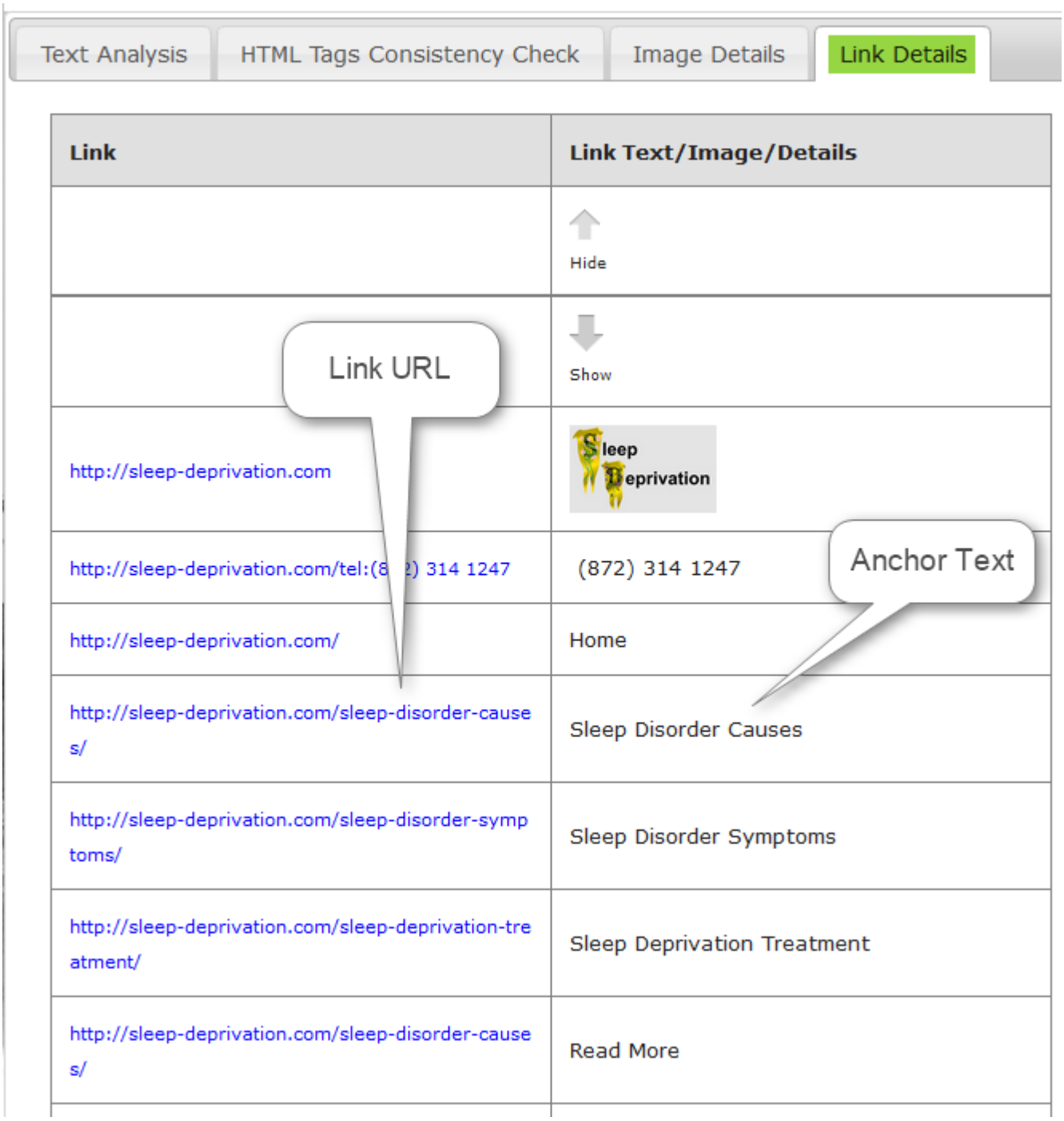
Image Details

This displays all of the images on the page along with their URL (where they are stored) and the ALT text for each of those images.

This makes it really simple to find the key images on the page and make sure they have relevant ALT tags and ensuring the ALT tags are not empty.

Link Details

This displays all of the links on the page and the Anchor text for those links.



Using this information enables you to manage the page links anchor text to be relevant to the keyword for that page.