

# **IBM Research Report**

## **Feeling Blue**

**Daniel F. Gruhl**  
IBM Research Division  
Almaden Research Center  
650 Harry Road  
San Jose, CA 95120-6099






**Research Division**  
**Almaden - Austin - Beijing - Haifa - India - T. J. Watson - Tokyo - Zurich**

# Feeling Blue

Daniel Gruhl

IBM ARC

06/08/2004

<i>Avg</i>	<i>Color</i>	<i>Hue</i>
		

What color is the Web? Well, if you look at all the pages that set a bgcolor and average it, the answer is a slightly blue-greenish light gray. The light gray is easy to understand; most pages are white (67.4%), some are black (6.09%) or gray (6.06%). However, this leaves 20.4% that are colored.

## Blue-green?

But why blue-green? There aren't that many blue-greenish web pages out there... well, borrowing an explanation from the January 2002 announcement that the universe is mint green, it's because there are an awful lot of red, yellow and blue pages. On average this is blue-green.



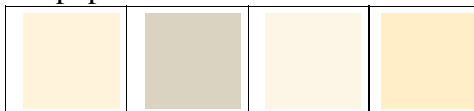
*Illustration 1* Linear Histogram of Hue



*Illustration 2* Log histogram of hue

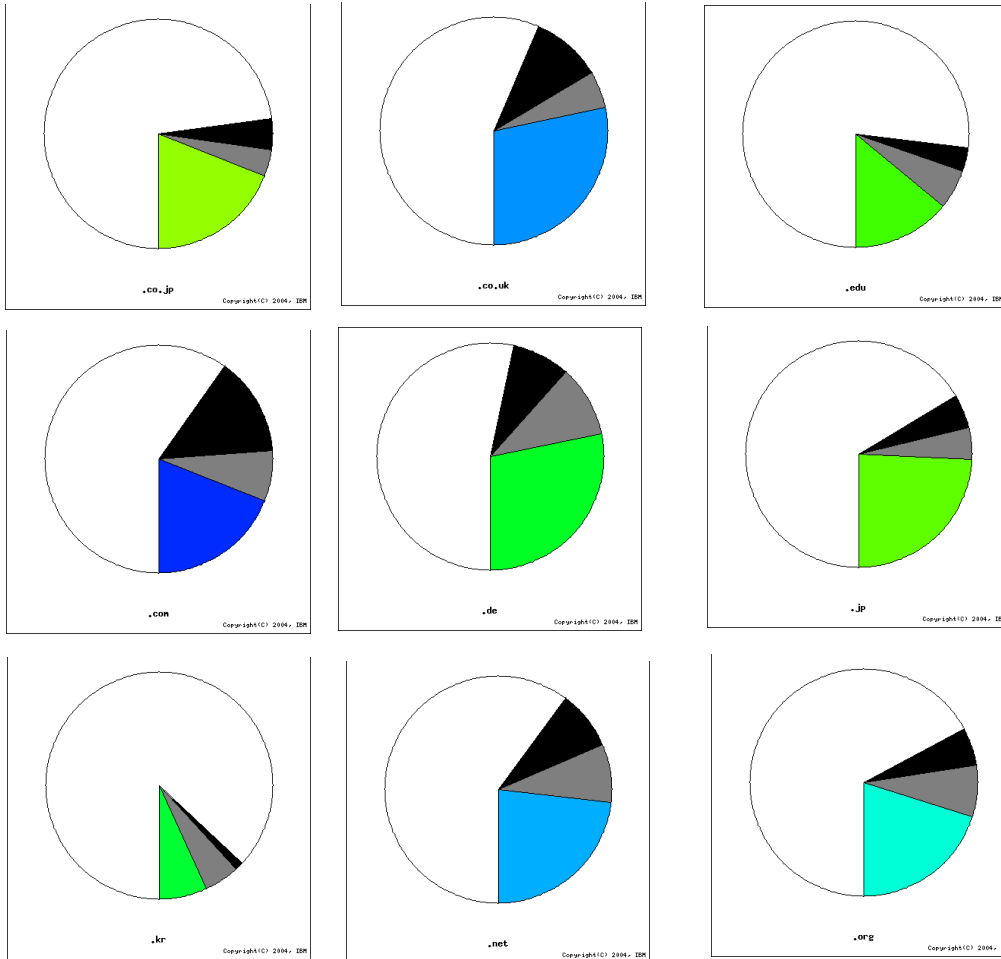
## What's the yellow spike?

The most prominent color on the web turns out to be yellow. Wait a second, you say, yellow pages? Well, many pages that are “off white” are actually a shade of yellow (albeit a very pale one). Here are the most popular 4:



## Color use by web region

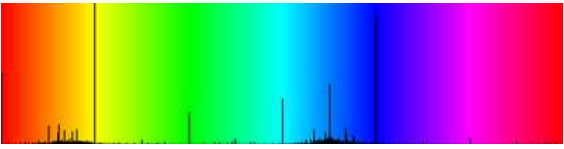
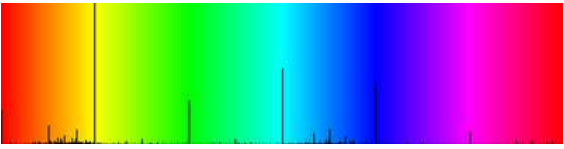
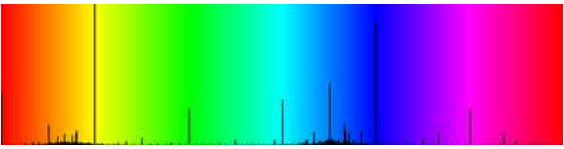
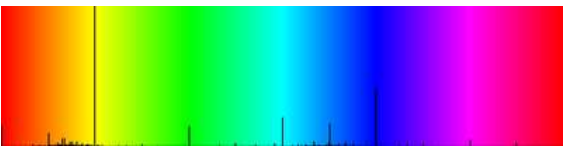
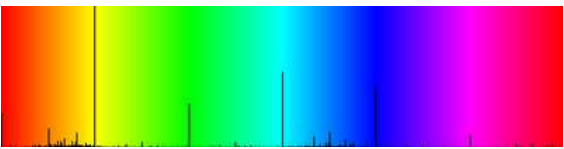
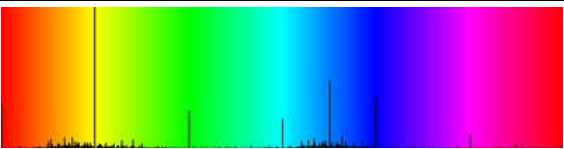
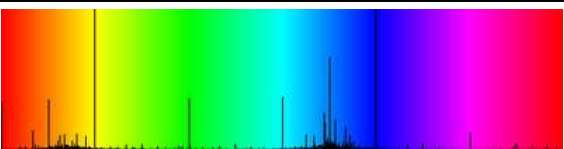
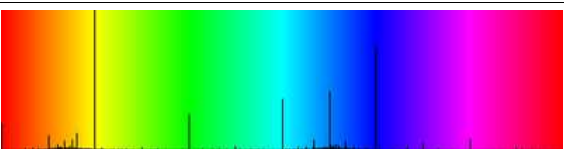
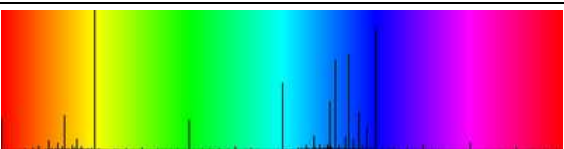
Different parts of the web use different colors, and more strikingly have different affinities for black, white, gray or a chroma. The below pie charts show this ratio for different parts of the web. The color in the “colored section” is the average color in that part of the web.



The full “spectra” appear below. Some interesting observations can be made, from the prevalence of blue in the UK commercial space, to the relative starkness of “.org” space.

### Why?

Why do a study like this? Mostly just for fun (this was run during the holiday break at the end of 2003) – and to illustrate and test some of the features of the WebFountain SDK for research. But it also is an example of a different way to look at the web.

<i>Color Use</i>	<i>Region</i>
	.de
	.co.jp
	.com
	.edu
	.jp
	.kr
	.net
	.org
	.co.uk

**Dataset**

Pseudo random selection from the WebFountain business crawl. Only pages that set a bgcolor explicitly were considered (i.e., no CSS set pages, no default pages). This was a set of 142,042,644 pages. Colored domains were represented by randomly selecting domains of more than a million examples (although most were around 8 million or so).