Best of the Web

comScore analysis of best practices among government websites

comScore

Prepared for GSA
REID JOHNSON Analyst, comScore
BHASKAR BULUSU Senior Analyst, comScore
MATT ARCHER Senior Director, comScore

Purpose and Scope

President Barack Obama signed into effect Executive Order 13571, "Streamlining Service Delivery and Improving Customer Service" on April 27, 2011, requiring federal agencies to take specific steps to strengthen customer service and improve government websites' delivery of services and information to the public. This effort promises to help eliminate outdated websites and reduce redundancy caused by the rapid proliferation of government websites over the past decade. Between June and August of 2011, the Office of Management and Budget froze the issuance of new .gov domain names, established a Reform Task Force to recommend updates to site policies and guidelines, and delineated a process for identifying sites subject to elimination, consolidation, or streamlining.

In light of this Executive Order, the General Services Administration's (GSA) Office of Citizen Services commissioned comScore to develop a "Best of the Web" analysis, highlighting best practices of .gov websites in order to improve all government websites and better serve their respective audiences. Sites were evaluated quantitatively and qualitatively in order to understand whether each site is attracting visitors, engaging those visitors, and maintaining an interface that is easy to use, has relevant content, and is reaching its target audience.

75 government sites are featured in this analysis, selected based on overall visitation and engagement or other specific qualities of importance, such as high relevance to niche populations among the general public.

comScore established broad "use categories" to segment government sites according to their functionality and purpose. These sites then underwent rigorous quantitative analysis according to a variety of metrics including overall visitation trends, time spent on the site, pages viewed, duration of visits, and frequency of visitation. This was followed by a qualitative analysis leveraging comScore's industry-leading expertise in the digital media landscape to evaluate ease of use and access, aesthetics, and more.

The analysis presented in this white paper is intended to empower GSA to structure the government-wide effort of measuring and analyzing website performance, guide continued improvements to government sites, and potentially identify cost savings opportunities in accordance with the directives of the Executive Order.

Preface & Editor's Note

comScore is a global leader in measuring the digital world and preferred source of digital business analytics. Through a powerful combination of behavioral and survey insights, comScore enables clients to better understand, leverage, and profit from the rapidly evolving worldwide web and mobile arena.

comScore provides syndicated and custom solutions in online audience measurement, e-commerce, advertising, search, video and mobile, and offers dedicated analysts with digital marketing and vertical-specific industry expertise. Advertising agencies, publishers, marketers and financial analysts turn to comScore for the industry-leading solutions needed to craft successful digital, marketing, sales, product development and trading strategies.

For this report, comScore leveraged its digital media and behavioral research expertise to develop a process of measuring and ranking top government sites approved by the General Services Administration's Office of Citizen Services. comScore's Government team spent six months investigating the sites included in this analysis and determining optimal metrics for both the quantitative and qualitative aspects of this research.

No measurement perspective is exactly identical, nor can clear preferences among different measurement approaches be established in all cases. As such, the analysis provided hereinafter must be understood as **comScore's own perspective**, provided as a guideline to assist government agencies in improving their respective online presences. comScore recognizes that the variety inherent to government websites poses serious obstacles to the kind of broad comparisons this study draws, and remains open to comments, questions, and suggestions on how to best segment and measure sites for future analyses.

FOR FURTHER INFORMATION, PLEASE CONTACT:

Reid Johnson
comScore, Inc.
+1 703 234 8633
rjohnson@comscore.com

Bhaskar Bulusu
comScore, Inc.
+1 703 234 9724
bbulusu@comscore.com

Executive Summary

RESEARCH OVERVIEW

- The top 75 government websites were evaluated on certain quantitative and qualitative dimensions in order to determine areas of relative strength as well as potential opportunities for improvement.
- Government websites were divided into three categories depending on their purpose: interactive, directional, or research.
- Websites were then individually rated:
  - quantitatively, according to "per visit" metrics.
  - qualitatively, according to navigability, aesthetics, and several other category-specific metrics.

HIGH LEVEL RESULTS

comScore's analysis revealed a fairly broad distribution of government website performance along the selected metrics, with the vast majority performing close to average (81% of sites scored between 2 and 4). The top-performing sites overall and within each category are as follows:

TOP 5 GOVERNMENT SITES				
1 State.gov	2 DHS.gov	3 Medicare.gov	4 OPM.gov	5 HUD.gov

TOP INTERACTIVE SITE	TOP DIRECTIONAL SITE	TOP RESEARCH SITES	
House.gov	Medicare.gov	State.gov	DHS.gov

The top-scoring sites were not only well-designed and aesthetically appealing, as determined by the qualitative analysis, but were also better able to engage their audience than the average site within their category. Sites scoring in the middle of the pack tended to have average visitor engagement compared to other sites in their category and at times exhibited minor design issues. The lowest-scoring sites tended to have comparatively low visitor engagement metrics combined with lower design quality.

The majority of sites scored similarly in both the quantitative and qualitative analyses, providing evidence that better site design is associated with improved visitor engagement. However, this trend is more evident for interactive sites than research sites, as visitors to interactive sites tend to perform tasks that might otherwise be confusing or tedious.

Features of the top-scoring sites may be used to inform the development of best practices for government sites; however, the limitations of the study should be kept in consideration when determining whether a low-scoring site is in genuine need of improvement. While underperformance for a particular site according to this analysis may indicate a genuine need for improvement in visitor engagement, it may also reflect characteristics of the site which produce patterns of visitor interaction which differ from the expected norm for the relevant use category. As such, these results are best employed as a starting point in the effort to improve the digital presence of federal agencies.

Qualitative and quantitative investigations of top government sites were segmented, based on the presumed intent of a typical visitor.

Methodology

OVERALL APPROACH

comScore's Best of the Web investigation began with identifying 75 key government sites, 65 of which were provided by GSA at the project's inception, while 10 were included at the suggestion of comScore based on high observed visitation.

These sites were then divided into three "use categories" based on the presumed intent and behavior of a typical visitor: "Interactive Engagement" (interactive), "Directional Engagement" (directional), and "Research Engagement" (research).

TOP 75 FEDERAL GOVERNMENT SITES

Interactive Engagement	Directional Engagement:	Research Engagement:
Visitors come to the site with a distinct purpose and the need to interact directly with the agency or committee of the site.	The sites content serves as a portal for its visitors linking them to other sites and agencies for more information.	Site serves as a learning portal providing all the information relative to its agency. A one stop shop.
ED.gov IRS	USA.gov Challenge.gov	Social Security U.S. Department of State
18 SITES	12 SITES	45 SITES

Interactive sites primarily enable visitors to accomplish a specific task or set of tasks. These sites facilitate the provisioning of government services that would otherwise need to be delivered to the public through different forms of correspondence. Examples of interactive websites include the Social Security Administration website (www.ssa.gov), the Department of Education's Federal Student Aid website (www.studentloans.gov), and the Office of Personnel Management's "USA Jobs" recruiting website (www.usajobs.gov).

Directional sites function largely as portals, or directories, to other content. Though they may provide visitors with some unique content, their primary role is to direct visitors to other government domains for more information. Examples of directional websites include GSA's USA.gov portal, the Library of Congress portal (www.loc.gov), and the House of Representatives main page (www.house.gov).

Segmenting the top 75 government websites allowed for comparisons between similarly-purposed websites based on a tailored set of metrics.

Research sites are primarily learning portals, communicating all information related to the agency to visitors, as well as, in some cases, facilitating the dissemination of government data to the public. Research sites include the National Oceanographic and Atmospheric Administration (NOAA) website (www.noaa.gov), the Census Bureau homepage (www.census.gov), and the Bureau of Labor Statistics website (www.bls.gov).

Sites often perform multiple functions; job seekers, for instance, could either use OPM's USAjobs.gov website as a research tool to examine the federal employment landscape, or they could interact with the site to apply for posted government positions. comScore's segmentation therefore relied on the primary purpose of the website to assign an appropriate use category, determined subjectively on a case-by-case basis.

Segmenting the top 75 government websites in this way allowed for comparisons between similarly-purposed websites based on a tailored set of metrics. While segmenting allows differences between use categories to be controlled, it is not sensitive to differences within each category. To the extent that a website differs from the prototypical use pattern of its assigned category, the metrics used in its evaluation may not be ideal. While it is impossible to completely eradicate this issue while also maintaining the ability to meaningfully compare very different websites, comScore determined that the balance provided by this approach was reasonable and appropriate to the purpose of the research.

Research	March 2012 Unique Visitors (000)
NIH.GOV	11,850
WEATHER.GOV	3,388
STATE.GOV	3,322
NOAA.GOV	3,271
NASA.GOV	3,220
OPM.GOV	3,142
CDC.GOV	2,888
USDA.GOV	2,425
NPS.GOV	2,026
WHITEHOUSE.GOV	1,927
USGS.GOV	1,586
CENSUS.GOV	1,417
BLS.GOV	1,372
USCOURTS.GOV	1,161
HUD.GOV	1,085
DHS.GOV	961
ARCHIVES.GOV	935
SENATE.GOV	927
FDA.GOV	912
DOL.GOV	824
JUSTICE.GOV	814
CIA.GOV	775
FTC.GOV	775

Research	March 2012 Unique Visitors (000)
DOT.GOV	754
FBI.GOV	751
OSHA.GOV	677
DFAS.MIL	649
USEMBASSY.GOV	599
CANCER.GOV	563
CMS.GOV	541
GPO.GOV	526
TSA.GOV	511
TSP.GOV	442
WOMENSHEALTH.GOV	392
DEFENSE.GOV	367
TREASURYDIRECT.GOV	350
CBP.GOV	340
CPSC.GOV	334
SAMHSA.GOV	308
MAKINGHOMEAFFORDABLE.GOV	282
NIST.GOV	238
NEH.GOV	219
NSA.GOV	186
HRSA.GOV	149
FDIC.GOV	144

Interactive	March 2012 Unique Visitors (000)
IRS.GOV	13,788
ED.GOV	8,724
SSA.GOV	3,854
USAJOBS.GOV	2,759
USCIS.GOV	2,414
VA.GOV	1,850
ARMY.MIL	1,608
NAVY.MIL	1,202
AF.MIL	1,013
EFTPS.GOV	922
STUDENTLOANS.GOV	803
MEDLINEPLUS.GOV	727
MEDICARE.GOV	594
BOP.GOV	457
GRANTS.GOV	405
MARINES.MIL	327
FBO.GOV	206
BENEFITS.GOV	204

Directional	March 2012 Unique Visitors (000)
USA.GOV	6,793
Search USA.gov	6,620
LOC.GOV	1,463
FUELECONOMY.GOV	1,438
HHS.GOV	1,256
HOUSE.GOV	1,142
FEMA.GOV	784
RECREATION.GOV	572
GSA.GOV	399
ENERGY.GOV	368
Gobierno USA.gov	130
Answers USA.gov	66

Within each use category, comScore conducted quantitative and qualitative analyses, described below, and assigned a score from 1-5 for each website for each analysis. The quantitative and qualitative scores were averaged to generate a final, aggregate score. No compelling reason to prefer one form of analysis to another could be established, so they were assessed as equally important and informative.

We assigned key metrics independently to each use category, based on a combination of theoretical and empirical bases.

Quantitative Methodology

OVERVIEW

comScore's quantitative analysis sought to establish an empirical basis by which the usability and efficacy of websites within each use category could be compared. Important to this effort was isolating total visitation from the analysis; certain government sites will have much higher traffic than others simply due to a larger audience for the content they provide, but these differences are not necessarily indicative of differences in quality or efficacy. comScore therefore focused on "per visit" metrics – such as pages per visit, minutes per visit, and visits per unique visitor – as proxy measurements of effective site design, with these metrics being independent of overall site visitation.

In addition, because typical usage is fundamentally different between the three categories, we assigned key metrics independently to each use category, based on a combination of theoretical and empirical bases. Data for the selected metrics were then collected for all sites in each category, and a normalization process was applied to generate a score from 0% - 100%. Ratings 1-5 were then applied based on the quintiles of these normalized scores, adjusting for whether the metric had a direct or an inverse relationship with site effectiveness. Finally, weights were applied to each metric rating within a category and combined to generate an aggregate rank.

METRIC SELECTION

Interactive Engagement

For Interactive websites, those whose primary purpose was to allow the visitor to accomplish a specific task, the average visits per visitor metric and the average minutes per visit metric were assessed as the strongest measures of effectiveness.

- Visits per visitor has an inverse relationship with efficiency; the fewer visits required to accomplish a task the better.
- Minutes per visit has a direct relationship with efficiency; More time spent during a visit to an interactive site indicates both that the site is facilitating the visitor's inquiry, and that the interactive process is sufficiently clear to prevent the visitor from breaking their inquiry into multiple sessions.

Directional Engagement

Average minutes per visit, average pages per visit, and average visits per visitor were used to gauge effectiveness of directional websites.

- Minutes per visit and pages per visit are inversely related to effectiveness; the quicker the site leads users to their intended destination the better.
- Visits per visitor is directly related to effectiveness, as this serves to indicate users' inclination to use the government directional site to locate information on a frequent basis.

Research Engagement

Key metrics: average minutes per visit and average pages per visit

- Both metrics are directly related to effectiveness, as indications of higher engagement with research content

NORMALIZATION

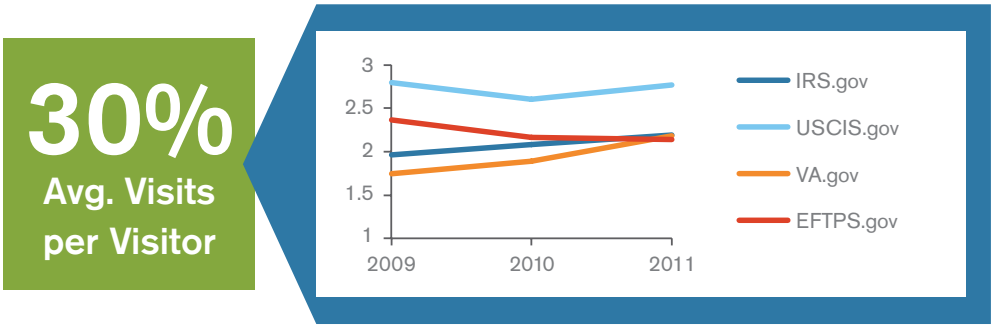
The normalization process used the distribution within each metric to calculate a cumulative probability value for each data point. Using the mean and standard deviation of the sample, the calculated p-value for each site represents the proportion of all sites it outperforms, projecting the observed distribution to a hypothetical, normally distributed population.

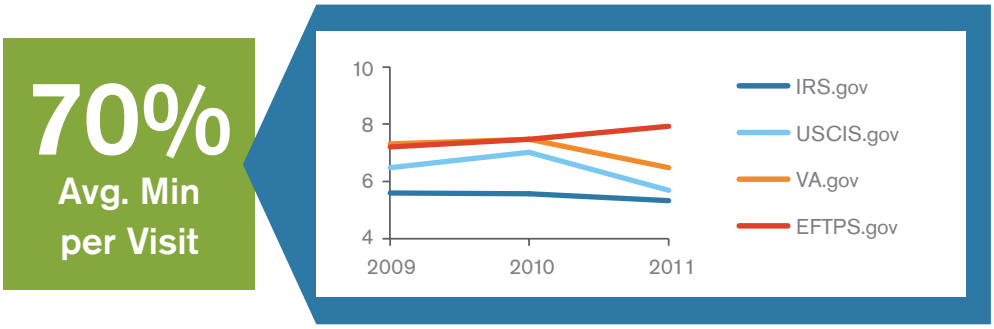
Ranks were then assigned to each metric based on quintile distribution. For metrics with a direct relationship to effectiveness, the score of 5 was assigned to the top quintile (those sites performing better than 80% of all others). For metrics with an inverse relationship to effectiveness, sites with the highest metrics were assigned a score of 1 (e.g., in cases where minutes per visit was a negative indicator, sites whose average minutes per visit were in the top 20% received a score of 1). Applying scores in this manner ensured that high performance on any given metric would always result in a high score.

WEIGHTING BY METRIC

The relative weights for each metric were informed by comScore's analysis of the historical relationship between increasing visitation to the top-performing government websites and the directionality of the changes among these metrics.

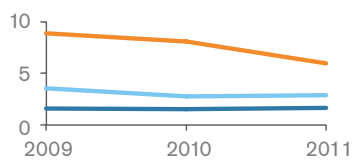
Interactive Engagement



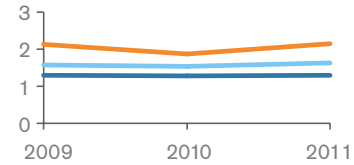


Directional Engagement

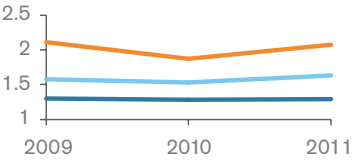
40%
Avg. Min per Visit



40%
Avg. Pages per Visit

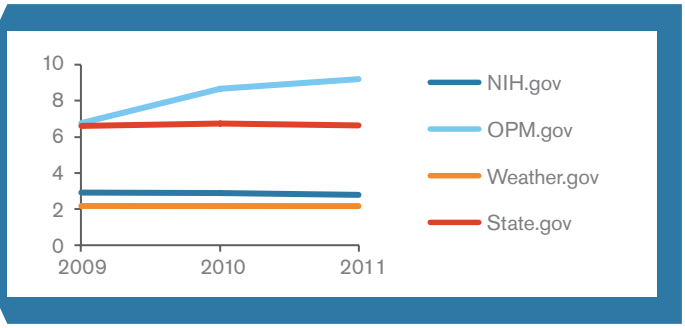


20%
Avg. Visits per Visitor

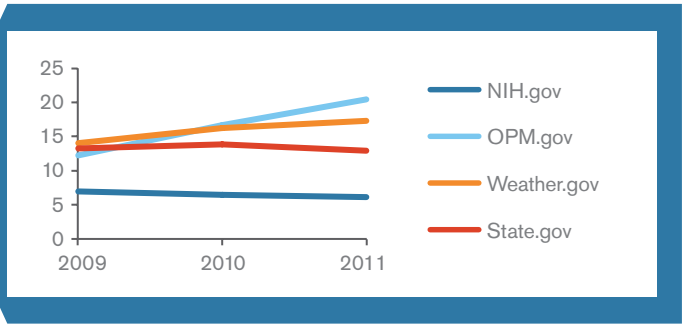


Research Engagement

70%
Avg. Minutes per Visit



30%
Avg. Pages per Visit



AGGREGATE QUANTITATIVE RATING

For each site within a category, the aggregate rating was established simply by summing the weighted ratings across each metric.

Appropriate categorization of sections dramatically improves the efficiency of any site with significant amounts of content.

Qualitative Methodology

comScore's qualitative analysis was intended to evaluate government sites based on more subjective criteria that are difficult or impossible to assess quantitatively. A number of criteria were established to guide this effort, some of which were specific to a single use category.

UNIVERSAL METRICS

Ease of Navigation plays a major role in the quality of a visitor's online experience. Ease of navigation was evaluated according to several metrics, including:

- **Location of items**: most of the important content is located above the fold (where it is viewable in the browser without scrolling), ensuring that the content is seen.

Figure 1 USA.gov main page (all relevant content is in-view without scrolling)

- **Categorization hierarchy**: appropriate categorization of sections dramatically improves the efficiency of any site with significant amounts of content. When the site's content is well-organized in layered categories, finding specific content becomes less daunting, and fewer clicks are required to arrive at the target destination. Drop-down menus are significantly more adept at organizing information than large, static lists.

Figure 2 NIH.gov (clear, attractive, and visual drop-downs help organize content)

Aesthetic aspects of a website greatly impact its ability to convey the target message.

- **Return navigability**: the ability to "step back" in a navigation process, or return to the home page, can be a very helpful feature in navigating sites with significant amounts of information.

Figure 3 State.gov (navigation banner tells visitors how they got to where they are)

- **Task-finding**: the website enables completion of a pre-selected task through an uncomplicated sequence of actions. This metric was only evaluated when a relatively straightforward purpose could be identified that was relevant to a member of the general public.

Aesthetic aspects of a website greatly impact its ability to convey the target message. Websites were assessed based on the following metrics:

- **Simplicity and minimalism**: well-organized content complements the message better than general clutter.
- **Effectiveness of images**: Do the images help convey relevant content? Are they visually appealing, or distracting? Effective use of images helps to quickly communicate content and increase visitor engagement. Ineffective use of images causes distraction, can decrease engagement, and prevents the inclusion of information more relevant to the visitor's purpose.

The searchability of a directional website is a key determinant of its ability to direct visitors to non-obvious or unfamiliar content.

Figure 4 House.gov (The image of the U.S. Capitol Building, in addition to being relevant to the site's contents, helps reinforce the "Find Your Representative" feature in its caption)

CATEGORY-SPECIFIC METRICS

Directional

The searchability of a directional website is a key determinant of its ability to direct visitors to non-obvious or unfamiliar content. While well-organized dropdown menus are an excellent way to organize content with which the visitor has some understanding, search functionality helps to direct less guided inquiries. The searchability of directional websites was established based on the following criteria:

- Does the search bar's placement on the page make it accessible? Or is it difficult to find?

Figure 5 USA.gov (prominent search bar placement)

- How accurately do search results match the content of search terms? How relevant are the results to what the searcher is trying to see? When appropriate, the availability of advanced search functionality (narrowing the search based on parameters other than key words) was considered an important aspect of searchability.

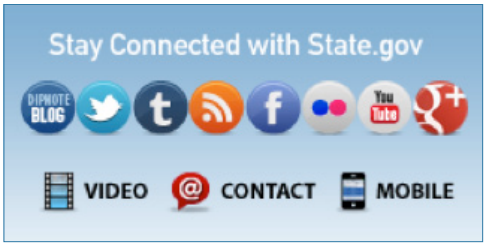
An extremely useful site feature is the ability to save information for later viewing and analysis.

Figure 6 USA.gov (top search results are highly relevant to the inquiry)

Research

In conducting research, an extremely useful site feature is the ability to save information for later viewing and analysis. In general, the most effective way to accomplish this in an online environment is to enable visitors to share content, either via email or social networking. The presence of these features for communicating content was therefore taken into strong consideration for research sites. An additional consideration for research sites whose agencies regularly publish content was whether one could subscribe to directly receive these periodic updates.

Figure 7 State.gov (options to leverage social media)



Average pages per visit was surprisingly high for directional sites.

Research Results

QUANTITATIVE

Table 1 Summary Statistics of Quantitative Metrics

	Minutes/Visit			Pages/Visit			Visits/Visitor		
	mean	median	st. dev	mean	median	st.dev	mean	median	st.dev
Interactive	6.0	6.0	2.9				2.0	1.8	0.7
Directional	3.6	3.2	2.0	6.4	5.2	6.0	1.6	1.5	0.3
Research	3.9	3.3	1.9	5.0	4.1	2.8			

Across the different categories of sites analyzed, the average minutes per visit metric behaved roughly as expected: it was highest for interactive engagement sites, where it represents the ability of the site to sustain user attention as they complete a task, and lowest for directional sites, where it represents potential time spent searching for linked content. Research sites, on average, had lower minutes per visit than interactive sites, which is somewhat surprising given that research tasks might be expected to take a similar amount of time as interactive tasks.

Average pages per visit was surprisingly high for directional sites; on average, visitors clicked through more than six pages before being sent to their intended destination. However, the large standard deviation reflects some of the significant differences observed between directional sites on this metric: pages per visit on recreation.gov were more than sixteen times higher than Gobierno USA.gov (23.8 pages per visit vs. 1.4 pages per visit, respectively).

After applying the normalization, quintile rating, and weighted aggregation described in the Methodology section, *supra*, the best- and worst- performing sites according to comScore's quantitative analysis are as follows:

Table 2 Final Quantitative Scores

Site	Use Category	Score
DHS.GOV	Research	5.0
STATE.GOV	Research	5.0
OPM.GOV	Research	5.0
HUD.GOV	Research	5.0
USCOURTS.GOV	Research	5.0
TREASURYDIRECT.GOV	Research	5.0
STUDENTLOANS.GOV	Interactive	4.7
SSA.GOV	Interactive	4.4
ED.GOV	Interactive	4.4
TSP.GOV	Research	4.3
USAJOBS.GOV	Interactive	4.1
HHS.GOV	Directional	4.1
MEDICARE.GOV	Interactive	4.0
USDA.GOV	Research	4.0
DOT.GOV	Research	4.0
CENSUS.GOV	Research	4.0
FEMA.GOV	Directional	3.8
GSA.GOV	Directional	3.7
CMS.GOV	Research	3.7
EFTPS.GOV	Interactive	3.7
OSHA.GOV	Research	3.7
MAKINGHOMEAFFORDABLE.GOV	Research	3.4
VA.GOV	Interactive	3.4
ARCHIVES.GOV	Research	3.3
GRANTS.GOV	Interactive	3.3
RECREATION.GOV	Directional	3.3
HRSA.GOV	Research	3.3
SAMHSA.GOV	Research	3.3
DOL.GOV	Research	3.3
FBO.GOV	Interactive	3.1
AnswersUSA.gov	Directional	3.0
FDA.GOV	Research	3.0
CDC.GOV	Research	3.0

Site	Use Category	Score
BENEFITS.GOV	Interactive	2.9
FUELECONOMY.GOV	Directional	2.8
WHITEHOUSE.GOV	Research	2.7
SENATE.GOV	Research	2.7
CIA.GOV	Research	2.7
HOUSE.GOV	Directional	2.6
LOC.GOV	Directional	2.6
ENERGY.GOV	Directional	2.6
BOP.GOV	Interactive	2.6
USA.GOV	Directional	2.4
SearchUSA.gov	Directional	2.4
USCIS.GOV	Interactive	2.4
NSA.GOV	Research	2.3
FDIC.GOV	Research	2.3
MEDLINEPLUS.GOV	Interactive	2.2
GobiernoUSA.gov	Directional	2.2
NOAA.GOV	Research	2.0
NIH.GOV	Research	2.0
CBP.GOV	Research	2.0
NPS.GOV	Research	2.0
FBI.GOV	Research	2.0
GPO.GOV	Research	2.0
IRS.GOV	Interactive	2.0
BLS.GOV	Research	2.0
CANCER.GOV	Research	2.0
TSA.GOV	Research	2.0
USGS.GOV	Research	2.0
NASA.GOV	Research	2.0
USEMBASSY.GOV	Research	2.0
NIST.GOV	Research	2.0
AF.MIL	Interactive	1.9
NAVY.MIL	Interactive	1.9
ARMY.MIL	Interactive	1.6
MARINES.MIL	Interactive	1.6
JUSTICE.GOV	Research	1.3
WEATHER.GOV	Research	1.3
FTC.GOV	Research	1.0
DFAS.MIL	Research	1.0
DEFENSE.GOV	Research	1.0
CPSC.GOV	Research	1.0
WOMENSHEALTH.GOV	Research	1.0
NEH.GOV	Research	1.0

Six sites received a "perfect score" of 5, all of which were categorized as research sites.

Highest-rated sites:

- Interactive: Studentloans.gov (4.7), SSA.gov (4.4)
- Directional: HHS.gov (4.1), FEMA.gov (3.8)
- Research: DHS.gov, State.gov and four others (5)

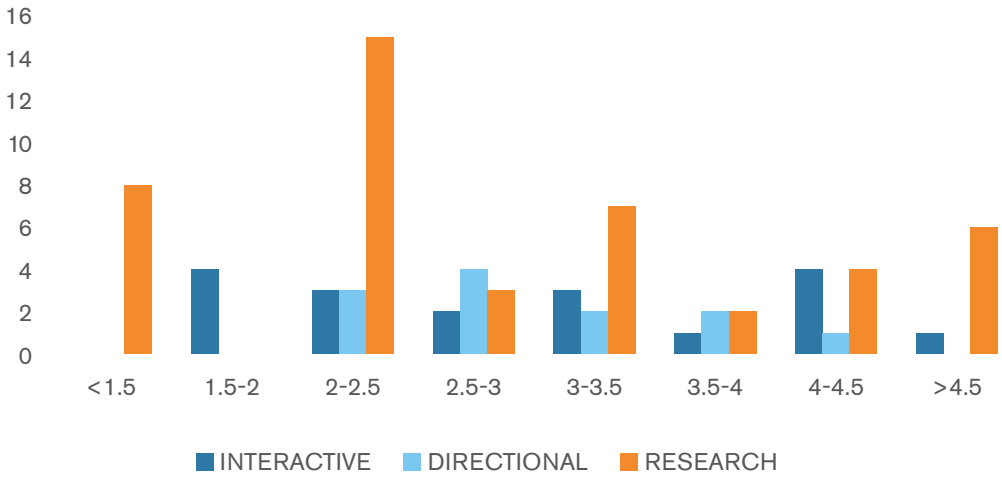
Lowest-rated sites:

- Interactive: Army.mil and Marines.mil (1.6)
- Directional: Gobierno USA.gov (2.2), SearchUSA.gov (2.4)
- Research – NEH.gov, Womenshealth.gov, and four others (1)

Six sites received a "perfect score" of 5, all of which were categorized as research sites. However, the research category as a whole realized the lowest average quantitative score (2.75, versus 3.01 for interactive and 2.96 for directional engagement sites) due to a number of very low-rated sites, which brought down the category average.

In general, greater variation in scores was observed with research sites than with interactive or directional sites. This is somewhat surprising given that the inclusion of a larger number of websites would, all else equal, tend to decrease volatility within the sample. While this is a strong indication that certain research sites genuinely perform better than others, it may also reflect a range of visitation patterns that differ from the category prototype and an opportunity for those underperforming sites to improve.

Table 3 Distribution of Quantitative Scores



QUALITATIVE

Table 4 Qualitative Scores

Site	Use Category	Qualitative Score
HOUSE.GOV	Directional	5.0
MEDICARE.GOV	Interactive	5.0
DHS.GOV	Research	5.0
STATE.GOV	Research	5.0
ARCHIVES.GOV	Research	5.0
WHITEHOUSE.GOV	Research	5.0
NOAA.GOV	Research	5.0
NIH.GOV	Research	5.0
JUSTICE.GOV	Research	5.0
RECREATION.GOV	Directional	4.0
LOC.GOV	Directional	4.0
USA.GOV	Directional	4.0
GRANTS.GOV	Interactive	4.0
SSA.GOV	Interactive	4.0
BOP.GOV	Interactive	4.0
USAJOBS.GOV	Interactive	4.0
OPM.GOV	Research	4.0
HUD.GOV	Research	4.0
USDA.GOV	Research	4.0
HRSA.GOV	Research	4.0
SAMHSA.GOV	Research	4.0
FDA.GOV	Research	4.0
SENATE.GOV	Research	4.0
CIA.GOV	Research	4.0
CBP.GOV	Research	4.0
NPS.GOV	Research	4.0
FBI.GOV	Research	4.0
GPO.GOV	Research	4.0
WEATHER.GOV	Research	4.0
FTC.GOV	Research	4.0
DFAS.MIL	Research	4.0
HHS.GOV	Directional	3.0
FEMA.GOV	Directional	3.0
GSA.GOV	Directional	3.0
ENERGY.GOV	Directional	3.0
SearchUSA.gov	Directional	3.0
GobiernoUSA.gov	Directional	3.0
STUDENTLOANS.GOV	Interactive	3.0
MEDLINEPLUS.GOV	Interactive	3.0
ED.GOV	Interactive	3.0
EFTPS.GOV	Interactive	3.0
AF.MIL	Interactive	3.0
ARMY.MIL	Interactive	3.0
IRS.GOV	Interactive	3.0
USCIS.GOV	Interactive	3.0
USCOURTS.GOV	Research	3.0
DOT.GOV	Research	3.0
CENSUS.GOV	Research	3.0
CMS.GOV	Research	3.0
CDC.GOV	Research	3.0
NSA.GOV	Research	3.0
FDIC.GOV	Research	3.0
BLS.GOV	Research	3.0
CANCER.GOV	Research	3.0
TSA.GOV	Research	3.0
DEFENSE.GOV	Research	3.0
CPSC.GOV	Research	3.0
FUELECONOMY.GOV	Directional	2.0
BENEFITS.GOV	Interactive	2.0
NAVY.MIL	Interactive	2.0
VA.GOV	Interactive	2.0
MARINES.MIL	Interactive	2.0
TREASURYDIRECT.GOV	Research	2.0
TSP.GOV	Research	2.0
OSHA.GOV	Research	2.0
MAKINGHOMEAFFORDABLE.GOV	Research	2.0
USGS.GOV	Research	2.0
NASA.GOV	Research	2.0
USEMBASSY.GOV	Research	2.0
NIST.GOV	Research	2.0
WOMENSHEALTH.GOV	Research	2.0
NEH.GOV	Research	2.0
AnswersUSA.gov	Directional	1.0
FBO.GOV	Interactive	1.0
DOL.GOV	Research	1.0

House.gov: All relevant information was located above the fold and was easily accessible throughout the website.

NIH.gov: Search results returned relevant and substantive information in a timely manner.

Key sites receiving high ratings:

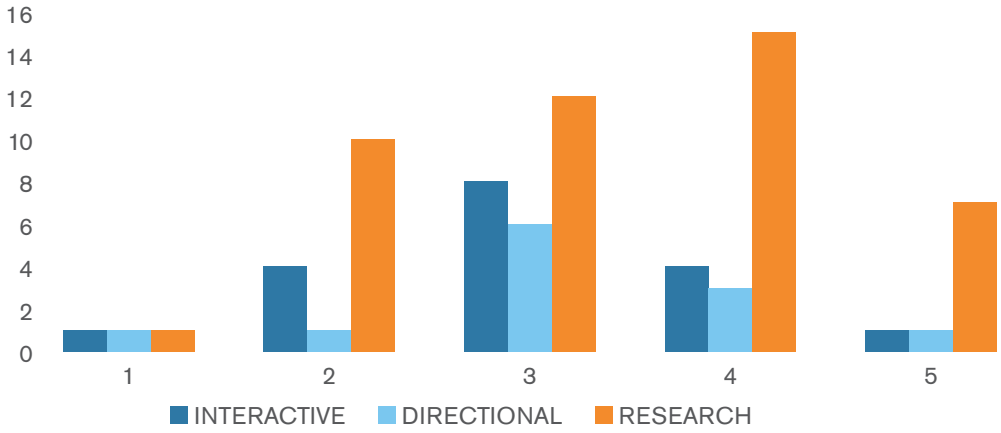
HOUSE.GOV

House.gov linked visitors to other sites and agencies for more information in an efficient manner with well-placed search bars. Search results were obtained quickly and were accurate and included the option to further refine searches. In addition, all relevant information was located above the fold and was easily accessible throughout the website.

NIH.GOV

Researching specific topics and general information was convenient and effective on NIH.gov because the site was well-organized, with most information located above the fold and drop-down menus categorized according to topic. Search results returned relevant and substantive information in a timely manner. In addition, the website included a variety of social media-sharing options.

Table 5 Distribution of Qualitative Scores



AGGREGATE RATINGS

Overall, DHS.gov and State.gov were the best performing websites, receiving overall scores of 5 (excellent). Eight other sites scored at or above 4 (good). While no sites received the lowest possible score, 11 received an overall score of 2 or less. The full list of aggregate scores is presented in Table 6.

Table 6 Aggregate Scores of Government Websites

Site	Use Category	Score
DHS.GOV	Research	5.0
STATE.GOV	Research	5.0
MEDICARE.GOV	Interactive	4.5
OPM.GOV	Research	4.5
HUD.GOV	Research	4.5
SSA.GOV	Interactive	4.2
ARCHIVES.GOV	Research	4.2
USAJOBS.GOV	Interactive	4.1
USCOURTS.GOV	Research	4.0
USDA.GOV	Research	4.0
STUDENTLOANS.GOV	Interactive	3.9
WHITEHOUSE.GOV	Research	3.9
HOUSE.GOV	Directional	3.8
ED.GOV	Interactive	3.7
GRANTS.GOV	Interactive	3.7
RECREATION.GOV	Directional	3.7
HRSA.GOV	Research	3.7
SAMHSA.GOV	Research	3.7
HHS.GOV	Directional	3.6
NOAA.GOV	Research	3.5
NIH.GOV	Research	3.5
TREASURYDIRECT.GOV	Research	3.5
FDA.GOV	Research	3.5
DOT.GOV	Research	3.5
CENSUS.GOV	Research	3.5
FEMA.GOV	Directional	3.4
GSA.GOV	Directional	3.4
SENATE.GOV	Research	3.4
CIA.GOV	Research	3.4
CMS.GOV	Research	3.4
EFTPS.GOV	Interactive	3.4
BOP.GOV	Interactive	3.3
LOC.GOV	Directional	3.3
USA.GOV	Directional	3.2
JUSTICE.GOV	Research	3.2
TSP.GOV	Research	3.2
CBP.GOV	Research	3.0
NPS.GOV	Research	3.0
FBI.GOV	Research	3.0
GPO.GOV	Research	3.0
CDC.GOV	Research	3.0
OSHA.GOV	Research	2.9
ENERGY.GOV	Directional	2.8
MAKINGHOMEAFFORDABLE.GOV	Research	2.7
SearchUSA.gov	Directional	2.7
VA.GOV	Interactive	2.7
USCIS.GOV	Interactive	2.7
WEATHER.GOV	Research	2.7
NSA.GOV	Research	2.7
FDIC.GOV	Research	2.7
MEDLINEPLUS.GOV	Interactive	2.6
GobiernoUSA.gov	Directional	2.6
FTC.GOV	Research	2.5
DFAS.MIL	Research	2.5
IRS.GOV	Interactive	2.5
BLS.GOV	Research	2.5
CANCER.GOV	Research	2.5
TSA.GOV	Research	2.5
BENEFITS.GOV	Interactive	2.5
AF.MIL	Interactive	2.5
FUELECONOMY.GOV	Directional	2.4
ARMY.MIL	Interactive	2.3
DOL.GOV	Research	2.2
FBO.GOV	Interactive	2.1
AnswersUSA.gov	Directional	2.0
DEFENSE.GOV	Research	2.0
CPSC.GOV	Research	2.0
USGS.GOV	Research	2.0
NASA.GOV	Research	2.0
USEMBASSY.GOV	Research	2.0
NIST.GOV	Research	2.0
NAVY.MIL	Interactive	2.0
MARINES.MIL	Interactive	1.8
WOMENSHEALTH.GOV	Research	1.5
NEH.GOV	Research	1.5

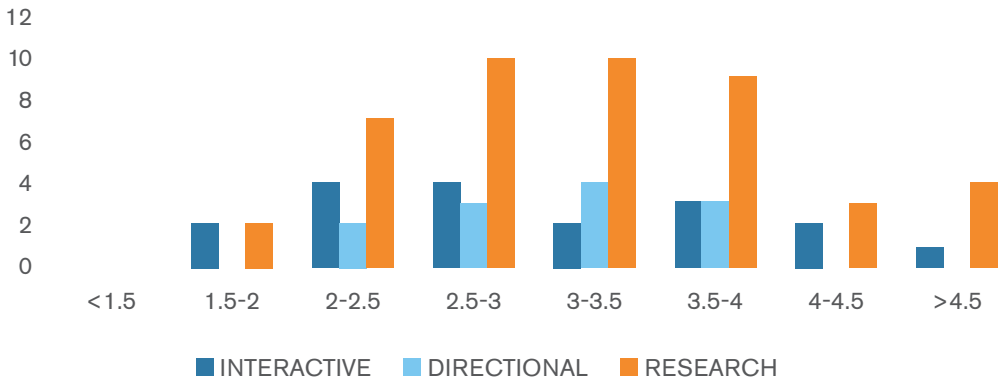
DISTRIBUTION OF AGGREGATE SCORES

As expected, most government websites cluster around the average rating of 3. 44% of sites scored between 2.5 and 3.5, while 81% of sites scored between 2 and 4. In addition, nearly the same number of sites scored greater than 3 (36) as scored lower than 3 (34). Five sites overall were rated exactly a score of 3.

Table 7 Distribution of Aggregate Scores

Category	<1.5	1.5-2	2-2.5	2.5-3	3-3.5	3.5-4	4-4.5	>4.5
Interactive	0	2	4	4	2	3	2	1
Directional	0	0	2	3	4	3	0	0
Research	0	2	7	10	10	9	3	4
(% of Total)	0.0%	5.3%	17.3%	22.7%	21.3%	20.0%	6.7%	6.7%

Figure 8 Frequencies of Aggregate Scores



A close match between a site's quantitative and qualitative score is evidence that typical users are engaged with the site as expected based on the quality of the site's design.

Research Sites are Stronger Outliers

A number of patterns emerge after combining the quantitative and qualitative scores into an aggregate rating for the various websites. Research sites tended to feature prominently at the highest end of the rating spectrum, representing 7 out of the 10 sites scoring 4 or above, and 4 out of 5 sites with ratings of 4.5 or above. On the other hand, poor performing research sites were more numerous than in other categories, accounting for 8 of 11 sites (73%) with scores of 2 or below. The frequencies of both high- and low-scoring research sites are greater than the overall prevalence of research sites in the sample.

Directional sites scored far closer to the mean, with all aggregate scores for this category falling between 2 and 4. Scores for interactive sites followed a roughly similar distribution to research sites, but appear slightly more frequently on the low end of the distribution (under 2.5) than on the high end.

Differences Between Quantitative and Qualitative Scores

A close match between a site's quantitative and qualitative score is evidence that typical users are engaged with the site as expected based on the quality of the site's design. However, differences between these two scores emerged quite often: nearly 60% of sites received a higher qualitative score than quantitative score, though many of these differences were minimal. In total, 34 out of 75 sites received quantitative and qualitative scores with differences greater than 1. Within this group, qualitative scores exceeded quantitative scores in 22 of the 34 instances. In addition, the average difference when qualitative scores exceeded quantitative scores (2.1) was greater than sites for which the inverse was true (1.9).

Differences between quantitative and qualitative scores are expected to some degree, but wide divergence (>1) between the two may indicate any/all of: the site's visitors do not exhibit prototypical behavior for the category; the overall quality of the site's design has a relatively modest impact on the site's engagement effectiveness; or the qualitative measures selected were insufficient barometers of engagement effectiveness.

Deviations from prototypical behavior likely explain the discrepancy between the quantitative score for weather.gov (1.3) and its qualitative score (4). Weather.gov is very well-designed to provide visitors with detailed information on weather forecasts and accommodates this kind of inquiry with a great deal of economy. After landing on the homepage, a typical visitor is relatively unlikely to spend a significant amount of time on the website viewing numerous pages; more likely, he will view two or three pages and spend a relatively short period of time on the site (on average, weather.gov received 2.0 minutes per visit and 2.7 pages per visit). The economy of the site's design, however, is not well-translated into the quantitative scoring.

A disconnect between site design and engagement effectiveness likely explains the observed difference in quantitative (4.3) and qualitative (2.0) scores for the Thrift Savings Plan website (www.tsp.gov). The site's design is relatively unsophisticated, but visitors evidently accomplish their research objectives with a greater degree of success than an average research site's visitor.

Conclusions

In general, and in line with expectations, the results of comScore's Best of the Web analysis indicate that a number of government websites are performing very well, a few are relative underperformers, and most fall somewhere in the middle.

The highest-performing sites can serve as a model for effective site design and communication with their respective audiences. Sites in the middle of the pack may be able to improve their effectiveness significantly by implementing features of higher-scoring sites – and for certain sites, these improvements may require minimal investment and be easy to implement. Sites that scored poorly may be either ineffectively reaching their target audience, or have such poor design elements that this audience is uninterested in fully engaging with the site.

The analysis also reveals that site design is potentially more important for sites where a typical visitor is trying to accomplish some specific (and, potentially, otherwise-tedious or confusing) interaction with the government agency than those for which a typical visitor is seeking information or conducting research.

However, in all cases the limitations of the methodology used here must be taken into consideration. In particular, sites assessed as underperforming may be designed for a specific purpose not fully addressed by this methodology. Therefore, these results are best employed as a starting point for further investigation by GSA, OPM, and the relevant agencies.

Case study: CDC

Overall, the Centers for Disease Control website (www.cdc.gov) is relatively well-laid-out and effective at engaging its audience. With an already-effective visual style, the replacement of the "A-Z Index" with category-based drop-down menus (of which the index could easily be a component) would help simplify the site's appearance, increase the concentration of relevant information above-the-fold, and help visitors navigate to content with which they may lack familiarity.

Frequently Asked Questions

COMSCORE'S BEST OF THE WEB ANALYSIS

Why wasn't our site included? How were websites selected?
Incorporating feedback and suggestions from GSA, comScore selected government sites for analysis based on overall web presence and visitation, importance among particular target audiences, and/or relevance to key initiatives by federal agencies.

How was the categorization of websites determined? Our website falls into more than one category (e.g., it has both research and directional features) – how is this reflected in the analysis?
Sites were assigned a use category based on comScore's estimation of how a typical visitor likely engages with the site and why they might be visiting in the first place. Sites that provide more than one of the basic usage engagement avenues were categorized based on comScore's qualitative interpretation of their primary engagement purpose.

What was the time period considered when calculating the quantitative metrics?
All quantitative metrics represent monthly data, averaged over a one-year period from March 2011 – March 2012 in order to capture seasonal variance of visitation and engagement patterns. Thus, the "visits per visitor" metric represents average monthly visits per visitor.

How were weights for each quantitative metric selected?
The weighting of each quantitative metric was determined by comScore's assessment of the metric's relationship to a prototypical site visit within each use category – this assessment was informed by the relationship between the metric's performance among sites whose overall visitation grew from March 2011 – March 2012. A close relationship between increasing visitation and improving performance of the metric was taken as evidence of relatively higher importance to visitor engagement.

What does it mean for a quantitative metric to be inversely related to engagement, and how does that influence the calculation of scores?
Scoring was inverse only in certain circumstances, such as with the use of the "visits per visitor" metric in evaluating Interactive Engagement sites. In such cases, we determined that a higher number of visits per visitor actually indicated a lower ability of users to effectively engage with the site (in this example, requiring multiple visits to accomplish an interaction with the agency).

Regardless of whether a metric was directly or inversely related to engagement, sites with the highest average pages per visit, minutes per visit, and visit per visitor will appear in the top quintile of the normalized rankings. Ordinarily, sites falling within the top quintile received a score of 5 for that metric, but for metrics with an inverse relationship to engagement, the top quintile received a score of 1.

Was a weighting system used to combine the various qualitative metrics into an aggregate score, similar to how the quantitative score was calculated?
The importance of qualitative metrics was determined on a case-by-case basis; they were points of consideration in establishing an overall qualitative score, but not direct determinants. A website would not, for instance, have its score reduced from 4 to 3 simply because it lacks a Twitter link. The goal of the qualitative analysis was to provide a holistic, context-sensitive score which could also be compared across different sites within a use category; comScore determined that using predefined criteria for each use category, but retaining flexibility in assessing the importance of each metric, provided the best means to achieve this objective.

comScore