
Executive Summary	1
Key Findings.....	3
Detailed Results Table	5
Methodology – what are the metrics?	7
About VisibleThread	8

The VisibleThread Website Index, US Government – Q4, 2011

Executive Summary

In October, 2010, US president Barack Obama signed the [Plain Writing Act of 2010](#). This mandates clear communications for citizens. The law requires that federal agencies use "clear Government communication that the public can understand and use." The implementation deadline is October 2011, after which all issued communications need to be in plain writing.

Two reasons underpin this act:

Reason 1 - Improve Customer Engagement: Clarity in documents and websites means citizens are more likely to achieve their goals. Whether you are filing a tax return or trying to understanding labor rights; clear instructions mean we have a higher chance of completing an action. When people know what the government asks of them, there's a better chance they'll comply. For example, the state of Washington Department of Revenue agency estimates it collected \$800k in extra revenue after rewriting one explanatory tax collection letter¹.

Reason 2 - Reduce Cost: If federal agencies can make it easier for people to engage, it means lower cost. It means lower cost for government due to less follow-up calls, letters and less time spent chasing. It cuts down on possible litigation due to misunderstood terms and conditions.

This inaugural plain writing Index ranks **30 US federal agency web sites** for clarity of written content. We scanned a 100 page sample from each agency web site analysed. We measured across these four dimensions:

- Readability – How readable is the content?
- Passive Language – Active Language communicates clearly. What proportion of sentences is passive?
- Long Sentences – What proportion of all sentences are considered too long?
- Word Complexity Density – Complex words make web pages hard to understand.

The samples were analysed on October 24th, 2011.

The findings show that the clearest web site is the [Bureau of Labor Statistics](#) followed by the [Centers for Disease Control and Prevention](#). Good readability (58), low complexity density (0.83) and a low proportion of

The VisibleThread Website Index, US Government – Q4, 2011



passive sentences (5.9%) helped make the bureau the leader. In each of these categories, the bureau ranked between 2 and 3. Long sentences (12.7%) ranking 8, is an area for improvement.

Notable mentions go to [NASA \(National Aerospace and Space Administration\)](#) and [Veteran Affairs](#). Despite a complex subject matter, NASA shows how clarity of communication is possible ranking joint 7th.

Veteran Affairs scored very well on 2 dimensions; ranking 2 for long sentences (8.8%) and 4 for readability (55). Passive language and complex density dragged their position to mid-table.

Sites at the bottom of the index were; the [Federal Railroad Administration](#), with the [National Highway Traffic Safety Administration \(Transportation\)](#) and the [Pension Benefit Guaranty Corporation](#) ranking joint 28th. These contained complex language, overly long sentences with high levels of passive language.

The full index is:

		US Government – Q4, 2011	Plain Writing Index
1	Bureau of Labor Statistics (Labor)	www.bls.gov	3.75
2	Centers for Disease Control (Health and Human Services)	www.cdc.gov	6
3	Bureau of the Census (Commerce)	www.census.gov	6.25
4	U.S. Geological Survey (the Interior)	www.usgs.gov	7
5	National Cancer Institute (Health and Human Services)	www.cancer.gov	7.75
6	Smithsonian Institution	www.si.edu	9
7	National Archives and Records Administration (NARA)	www.archives.gov	11.75
7	National Aeronautics and Space Administration	www.nasa.gov	11.75
9	State	www.state.gov	12.25
10	Federal Aviation Administration (Transportation)	www.faa.gov	13
10	Veteran Affairs	www.vba.va.gov	13
12	Natural Resources Conservation Service (Agriculture)	www.nrcs.usda.gov	13.75
13	the Interior	www.doi.gov	14
14	U.S. Mint (the Treasury)	www.usmint.gov	14.5
15	Social Security administration (Health and Human Services)	www.ssa.gov	15
16	Health Resources and Services Administration (Health and Human Services)	www.hrsa.gov	16.25
17	Federal Bureau of Investigation (Justice)	www.fbi.gov	16.5
17	U.S. Citizenship and Immigration Services (Justice)	www.uscis.gov	16.5
19	Substance Abuse and Mental Health Services Administration (Health and Human Services)	www.samhsa.gov	16.75
20	General Services Administration	www.gsa.gov	17
21	Community Oriented Policing Services (Justice)	www.cops.usdoj.gov	17.5
22	Small Business Administration (Labor)	www.sba.gov	18
23	Homeland Security	www.dhs.gov	18.5
24	Federal Deposit Insurance Corporation	www.fdic.gov	18.75
25	Federal Highway Administration (Transportation)	www.fhwa.dot.gov	22.75
25	Program Support Center (Health and Human Services)	www.psc.gov	22.75
27	Transportation	www.dot.gov	23.75
28	National Highway Traffic Safety Administration (Transportation)	www.nhtsa.gov	24.5
28	Pension Benefit Guaranty Corporation	www.pbpc.gov	24.5
30	Federal Railroad Administration (Transportation)	www.fra.dot.gov	28.5

We will publish the next index in Q1, 2012 and each quarter from there on. We will analyze trends as part of that index report. Additional agency sites may be added to that index.

Key Findings

- Leaders

The top ranking web sites were:

			Plain Writing		Readability		Passive		Long		Complex	
			Index	Score	Rank	Score	Rank	Score	Rank	Score	Rank	
1	Bureau of Labor Statistics (Labor)	www.bls.gov	3.75	58	3	5.9%	2	12.7%	8	0.83	2	
2	Centers for Disease Control (Health and Human Services)	www.cdc.gov	6	52	6	7.0%	7	9.4%	3	1.19	8	
3	Bureau of the Census (Commerce)	www.census.gov	6.25	85	1	8.6%	10	13.5%	11	0.85	3	
4	U.S. Geological Survey (the Interior)	www.usgs.gov	7	44	15	5.2%	1	11.3%	6	0.98	6	
5	National Cancer Institute (Health and Human Services)	www.cancer.gov	7.75	49	9	10.2%	17	10.3%	4	0.70	1	

The [Bureau of Labor Statistics](#) (rank 1) and [Centers for Disease Control and Prevention](#) (rank 2) are the leaders. Low levels of long sentences, high levels of active language, combined with easy to understand terminology were the winning traits. Readability levels are impacted directly by these factors. A readability level of close to 60 (out of a top possible score of 100) rivals that of the best commercial websites.

Both the [US Geological Survey](#) (rank 4) and the [National Cancer Institute](#) (rank 5) scored well but were let down by a single metric, readability (44) in the case of the Geological Survey and passive language (10.2%) in the case of the National Cancer Institute.

- Notable Mentions:

- [NASA \(National Aerospace and Space Administration\)](#) (rank joint 7th) deserves special mention. Scores for readability (50, rank 7), passive language (8.7%, rank 12) and complex language (1.39, Rank 12) were impressive, given the technical nature of NASA's remit. Despite a complex subject matter, NASA shows how clarity of communication is possible. A reduction in long sentences (14.9%, Rank 17) will improve NASAs overall position.
- [Veteran Affairs](#) (rank Joint 10th). Veteran Affairs scored very well in two of four measures. Long sentences (8.8%, rank 2) and readability (55, rank 4) were impressive results. Passive language coming in at 30th position (11.3%, rank 25) and complex density () pulled the overall ranking down. If Veteran Affairs can adopt more active & less complex language, it will climb significantly.

- Room for Improvement:

The bottom ranking web sites were:

25	Federal Highway Administration (Transportation)	www.fhwa.dot.gov	22.75	39	23	10.3%	19	18.6%	23	2.05	26
25	Program Support Center (Health and Human Services)	www.psc.gov	22.75	36	25	10.7%	24	13.5%	13	2.38	29
27	Transportation	www.dot.gov	23.75	35	28	10.7%	23	18.7%	24	1.92	20
28	National Highway Traffic Safety Administration (Transportation)	www.nhtsa.gov	24.5	40	21	12.2%	27	24.3%	28	1.95	22
28	Pension Benefit Guaranty Corporation	www.pbgc.gov	24.5	46	12	12.7%	30	22.3%	26	2.43	30
30	Federal Railroad Administration (Transportation)	www.fra.dot.gov	28.5	29	30	12.6%	29	25.7%	30	2.00	25

Coming in at the bottom of the index was the [Federal Railroad Administration](#). The [National Highway Traffic Safety Administration \(Transportation\)](#) and the [Pension Benefit Guaranty Corporation](#) ranked joint 28th.

Factors making these the worst performers were: high proportion of long sentences (between 22% and 25%), high levels of passive language (between 12% and 13%), high degrees of complex language (density between 1.95 and 2.43), and poor readability ratings (between 29 and 46).

- **Takeaways:**

1. **Wide Variability between best and worst performers**

There is very wide variability between the best and lowest scoring websites. This applies across all metrics.

For example, readability ranges from 85 (out of 100) for 3rd place [Bureau of the Census](#) contrast with a very low 29 (out of 100) for bottom placed Federal Railroad Administration. This means that someone reading the [Federal Railroad Administration](#) site would need a higher level degree, while a 5th grader would be able to understand the [Bureau of the Census](#).

2. **Agencies can dramatically improve clarity by focussing on certain metrics**

Many agencies scored well in certain areas but were let down by one or two metrics.

For instance, the [State Department](#) site ranks 9th overall, however it ranks a much higher 3rd and 5th on passive and complexity. The score for long sentences (24.8%) ranks 29th. So here, we will see possibility for immediate improvement by reducing the number of long sentences.

3. **Complex Content can be clear**

[NASA](#) is an excellent example of a clear writing in a website which has a relatively complex subject matter. It demonstrates that sites can be clear even when dealing with technical topics.

4. **Sites catering to non-native English speakers have room for improvement**

The [US Citizen and Immigration Services](#) ranked in the bottom half at joint 17th. Increasing clarity here by reducing complex and passive language levels will result in better citizen engagement.

5. **Agencies who engage with the public to complete tasks online can reduce costs by improving clarity**

Agencies where people can complete tasks online without recourse to phone will save money by improving clarity. For instance, the [Social Security Administration](#) can improve clarity by reducing the number of long sentences (16.3%) and passive language (10.2%).

Detailed Results Table

The full detailed table is shown below.

The VisibleThread Website Index, US Government – Q4, 2011



VisibleThread™ US Government – Q4, 2011			Plain Writing		Readability		Passive		Long		Complex		
			Index	Score	Rank	Score	Rank	Score	Rank	Score	Rank	Score	Rank
1	Bureau of Labor Statistics (Labor)	www.bls.gov	3.75	58	3	5.9%	2	12.7%	8	0.83	2		
2	Centers for Disease Control (Health and Human Services)	www.cdc.gov	6	52	6	7.0%	7	9.4%	3	1.19	8		
3	Bureau of the Census (Commerce)	www.census.gov	6.25	85	1	8.6%	10	13.5%	11	0.85	3		
4	U.S. Geological Survey (the Interior)	www.usgs.gov	7	44	15	5.2%	1	11.3%	6	0.98	6		
5	National Cancer Institute (Health and Human Services)	www.cancer.gov	7.75	49	9	10.2%	17	10.3%	4	0.70	1		
6	Smithsonian Institution	www.si.edu	9	43	18	6.8%	4	11.2%	5	1.24	9		
7	National Archives and Records Administration (NARA)	www.archives.gov	11.75	54	5	7.5%	8	12.1%	7	2.07	27		
7	National Aeronautics and Space Administration	www.nasa.gov	11.75	50	7	8.7%	11	14.5%	17	1.39	12		
9	State	www.state.gov	12.25	46	12	6.7%	3	24.8%	29	0.86	5		
10	Federal Aviation Administration (Transportation)	www.faa.gov	13	46	12	12.3%	28	6.1%	1	1.37	11		
10	Veteran Affairs	www.vba.va.gov	13	55	4	11.3%	25	8.8%	2	1.94	21		
12	Natural Resources Conservation Service (Agriculture)	www.nrcs.usda.gov	13.75	36	25	6.9%	6	12.8%	9	1.72	15		
13	the Interior	www.doi.gov	14	60	2	9.2%	13	23.7%	27	1.65	14		
14	U.S. Mint (the Treasury)	www.usmint.gov	14.5	50	7	11.5%	26	17.3%	21	0.85	4		
15	Social Security Administration (Health and Human Services)	www.ssa.gov	15	47	10	10.2%	18	16.3%	19	1.64	13		
16	Health Resources and Services Administration (Health and)	www.hrsa.gov	16.25	40	21	9.2%	12	14.0%	15	1.80	17		
17	Federal Bureau of Investigation (Justice)	www.fbi.gov	16.5	41	20	10.1%	16	16.8%	20	1.36	10		
17	U.S. Citizenship and Immigration Services (Justice)	www.uscis.gov	16.5	47	10	10.5%	21	13.5%	12	1.96	23		
19	Substance Abuse and Mental Health Services Administration	www.samhsa.gov	16.75	43	18	7.6%	9	14.5%	16	1.98	24		
20	General Services Administration	www.gsa.gov	17	36	25	9.6%	15	13.3%	10	1.83	18		
21	Community Oriented Policing Services (Justice)	www.cops.usdoj.gov	17.5	39	23	6.9%	5	13.8%	14	2.11	28		
22	Small Business Administration (Labor)	www.sba.gov	18	44	15	10.4%	20	15.3%	18	1.92	19		
23	Homeland Security	www.dhs.gov	18.5	35	28	9.4%	14	22.1%	25	1.09	7		
24	Federal Deposit Insurance Corporation	www.fdic.gov	18.75	44	15	10.6%	22	17.5%	22	1.75	16		
25	Federal Highway Administration (Transportation)	www.fhwa.dot.gov	22.75	39	23	10.3%	19	18.6%	23	2.05	26		
25	Program Support Center (Health and Human Services)	www.psc.gov	22.75	36	25	10.7%	24	13.5%	13	2.38	29		
27	Transportation	www.dot.gov	23.75	35	28	10.7%	23	18.7%	24	1.92	20		
28	National Highway Traffic Safety Administration (Transportation)	www.nhtsa.gov	24.5	40	21	12.2%	27	24.3%	28	1.95	22		
28	Pension Benefit Guaranty Corporation	www.pbgc.gov	24.5	46	12	12.7%	30	22.3%	26	2.43	30		
30	Federal Railroad Administration (Transportation)	www.fra.dot.gov	28.5	29	30	12.6%	29	25.7%	30	2.00	25		

Each score in the index is color-coded green to red. Green indicates best, red indicates worst.

Color-coding helps us to understand sites where one or two specific scores may be dragging down the overall ranking. Flagging specific areas, for instance passive language pinpoints areas for improvement.

For instance, the [State Department](#) site ranks 9th overall, however it ranks a much higher 3rd and 5th on passive and complexity. The score for long sentences (24.8%) ranks 29th. So here, we will see possibility for immediate improvement by reducing the number of long sentences.

Methodology – what are the metrics?

- We scan the websites between 8 – 12PM EST.
- The scan is conducted on a business day.
- We scan 100 pages of content as a result of using automated crawling techniques.
- We crawl starting from the publically available URL.
- Certain pages within the sample of 100 contain non-textual content (eg: videos). We omit these pages from our scan.

We calculate the Index based on 4 metrics. Each metric contributes equally to the final score. The metrics are:

Metric	Formula
--------	---------

1. Readability

Readability ranges from 1 to 100. 100 is the top mark. If communicating with citizens, aim for at least 50.

This is based on the Flesch Reading Ease index.

$(206.835 - (1.015 \times \text{Average Sentence Length}) - (84.6 \times \text{Average Syllables per Word}))$

2. Passive Language

The % rating is the proportion of sentences with a passive construct. Passive Language is where the subject acted upon appears before the verb.

"Quality is monitored" vs. "We monitor quality"

If you use active voice, you will increase clarity & strength. You will also flush out the 'actor', i.e. who did the action?

$(\text{Passive Sentences} / \text{Total Sentences} * 100)$

3. Long Sentences

The % rating is the proportion of sentences that are long. Long sentences are greater than 25 words. Long sentences mask multiple concepts. Splitting them, results in a clearer message.

$(\text{Long Sentences} / \text{Total Sentences} * 100)$

4. Complex Word Density

The density rating is the proportion of complex word hits relative to the total word count. This scan looks for complex words/phrases based on Federal Guidelines. See <http://www.plainlanguage.gov/howto/wordsuggestions/implewords.cfm> for the list scanned. Replacing complex words with simpler words lets your readers concentrate on your content.

(Complex Words/Total Words * 100)

About VisibleThread

VisibleThread provides document content analysis solutions and web site reports. We flag poor quality and complex language for documents and websites. This eliminates tedious manual review cycles, saving you time and cost.

Our solutions & reports allow customers to:

- Scan docs and websites in minutes
- Identify risky and complex language with objective metrics
- Benchmark web sites against sector peers

For Proposal Managers, Government Agencies and IT delivery teams', our solutions increase the likelihood of winning bids, drive plain language audits and make review teams 40% more efficient.

Please direct inquiries as follows:

- For a specific agency or bureau index, email: info@visiblethread.com
- For questions on the metrics or methodology, email: info@visiblethread.com

ⁱ http://www.usatoday.com/news/nation/2006-12-10-washington-plain-talk_x.htm & <http://www.governor.wa.gov/priorities/plaintalk/default.asp>