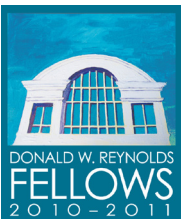


# U.S. Newspaper Reporters' Perceptions and Use of Government Data



DAVID HERZOG  
REYNOLDS JOURNALISM INSTITUTE  
2010-2011 FELLOW

OCTOBER 2011

# U.S. Newspaper Reporters' Perceptions and Use of Government Data

Ming Dai, David Herzog, and Ken Fleming

RJI

The 2011 Government Data Transparency Survey was conducted from February 14 to March 22, 2011. Six hundred and two reporters and editors in the United States took part in the survey. Journalists in Missouri (N=82) were oversampled for the survey. Five hundred and ninety-eight completed surveys were used for the data analysis below.

The results showed that reporters regularly turn to government websites looking for data that will help their reporting. In general, the reporters believe that different categories of data should be more accessible to them than the data actually are. Among the different categories of data, information about contracts and spending are the most accessible. Local and state governments are considered more transparent than the federal government. Local governments also are considered more accountable than state and federal governments.

The reporters did not seem to have difficulties in using and understanding the government data online. They were also neutral on whether they need to a third-party organization to help them understand the data. Spreadsheets are the tool most often used by the reporters to analyze data, with statistical software, such as SPSS and SAS, the least used.

The survey results are reported in the following two sections. The first section summarizes the results for the whole sample. The second section has the results for the journalists in Missouri. With each section, respondents' demographic information was reported first, followed by their perception of government data openness and accessibility, their attitudes toward different levels of government, and their use of government data.

## I. Results for all respondents

### • *Demographics*

The survey interviewed 598 newspaper reporters in the United States on the phone from late February to early March. Among the respondents, eight percent (N=50) work as both reporter and editor. Thirteen percent of the reporters (N=82) work for newspapers in Missouri, making it the most represented state in the survey. Nearly two thirds (N=381) are male and 36 percent (N=217) are female. The respondents' ages range from 21 to 78 years and the average age was 41. An overwhelming majority (N=583) of the respondents work at daily newspapers and less than two percent (N=10) of the reporters work at weekly newspapers. The experience of the reporters working at their current newspapers and as a paid journalist ranges from less than one year to nearly 50 years. On average, the reporters have worked as a paid journalist for about 17 years and have worked at their newspapers for nearly 10 years. Nearly 54 percent (N=324) of the reporters work in the print version of their newspapers and only two percent (N=12) work in the online version of their newspapers. Forty-four percent (N=262) of the reporters work in both the print and online versions. The variety of reporters' beats is diverse, including government affairs, business, environment and health. Very few reporters have science and technology as their beat. The size of the newspapers' staffs ranges from one person to nearly 800. Both the weekday circulation and the weekend circulation of the newspapers ranges from 25,000 to more than 200,000. The average size of weekday circulation of the newspapers is a little less than 50,000 and the average size of weekend circulation of the newspapers is a little more than 50,000. The respondents' political views tend to be neutral (M=4.42), though some reporters said they were very liberal or very conservative.

Table 1 Means and standard deviations for age, political ideology, work experience at current news organization, work experience as a paid journalist, weekday circulation, weekend circulation, and size of news organization

Variables	Mean	SD	N
Age (in years)	41.48	12.82	594
Political ideology*	4.42	1.24	547
Work experience at current news organization (in years)	10.34	9.78	598
Work experience as a paid journalist	16.75	11.89	597
Weekday circulation**	2.82	2.22	535
Weekend circulation**	3.29	2.42	484
Size of news organization***	56.10	85.13	534

\*Responses were coded 7 = very liberal, 6 = somewhat liberal, 5 = slightly liberal, 4 = neutral, 3 = slightly conservative, 2 = somewhat conservative, 1 = very conservative

\*\*1 = 25,000 or less; 2 = 25,000 to 50,000; 3 = 50,000 to 75,000; 4 = 75,000 to 100,000; 5 = 100,000 to 150,000; 6 = 150,000 to 200,000; 7 = 200,000 or more

\*\*\*Number of journalists, ranging from 1 to infinity

Table 2 Percentages for newspaper type, beat specializations, online and print responsibilities, job responsibility, gender, and location of news organization

<i>Variables</i>	<i>%</i>
Newspaper type	
Daily	97.5
Weekly	1.7
	100%
	(N = 593)
Beat specializations	
Business	5.7
Science	.8
Education	10.9
Health	1.8
Police	7.9
Municipal affairs/Government	29.3
Technology	.5
Sports	1.2
General assignment	13
Other	28.6
	100%
	(N=596)
Online vs print responsibilities	
Print	54.2
Online	2.0
Both	43.8
	100%
	(N=598)
Job responsibility	
Reporter	91.6
Reporter and Editor	8.4
	100%
	(N=598)
Gender	
Male	63.7
Female	36.3
	100%
	(N=598)

Location (state) of news organization

Alabama	2.8	Nebraska	2.2
Alaska	.7	Nevada	.5
Arizona	2.0	New Hampshire	.3
Arkansas	1.3	New Jersey	1.0
California	5.7	New Mexico	1.3
Colorado	1.5	New York	2.7
Connecticut	.8	N. Carolina	2.0
Delaware	.3	N. Dakota	1.2
District of Columbia	.2	Ohio	4.0
Florida	2.8	Oklahoma	2.0
Georgia	1.2	Oregon	.8
Hawaii	.3	Pennsylvania	3.8
Idaho	.8	Rhode Island	.3
Illinois	3.5	S. Carolina	1.2
Indiana	5.0	S. Dakota	.3
Iowa	1.8	Tennessee	1.7
Kansas	2.8	Texas	5.7
Kentucky	1.2	Utah	.3
Louisiana	2.0	Vermont	.7
Maine	.3	Virginia	1.0
Maryland	.7	Washington	2.2
Massachusetts	1.3	Wisconsin	2.8
Michigan	3.5	Wyoming	.5
Minnesota	2.3	Outside U.S./Other	.7
Mississippi	1.2		
Missouri	13.7		100%
Montana	.5		(N=597)

• *Perception of government data openness and accessibility*

On average, government data are quite relevant to the reporters' daily work. A number of questions were asked about reporters' presumption of openness for different types of data. The reporters, on average, felt that government data are open to them, with campaign finance and lobbying information presumed the most open category (M=4.88) and farming/agricultural activities the least (M=4.23). In terms of the accessibility of the data categories, the vital records such as births, deaths, and marriages were the most accessible to the reporters (M=3.83) and the information about workplace safety inspections was the least accessible (M=3.11). Overall, the reporters say the data is less available to them than they presume it should be.

**Table 3** Means and standard deviations for respondents' perception of government data openness

Variables*	Mean	SD	N
Relevance of government data to daily work	4.29	.96	598
Government spending and contracts	4.84	.50	598
Identities of government employees and their salaries	4.57	.73	598
Environmental monitoring and enforcement	4.76	.57	593
Campaign finance and lobbying	4.88	.44	597
Public schools, education and teachers	4.72	.54	591
Vital records such as deaths, births and marriages	4.40	.87	593
Public safety such as crimes, etc.	4.81	.50	595
Motor vehicle accident reports	4.55	.77	596
Workplace safety inspections	4.64	.72	592
Regulation of licensed professionals such as doctors and lawyers	4.64	.72	592
Information about farming/agricultural activities	4.23	.89	568

\*Responses were coded 5 = very much; 4 = somewhat; 3 = neutral; 2 = not; 1 = not at all

**Table 4** Means and standard deviations for respondents' perception of government data accessibility

Variables*	Mean	SD	N
Government spending and contracts	3.55	.95	563
Identities of government employees and their salaries	3.78	1.04	569
Environmental monitoring and enforcement	3.34	.93	519
Campaign finance and lobbying	3.82	1.04	556
Public schools, education and teachers	3.52	.99	541
Vital records such as deaths, births and marriages	3.83	1.23	559

Public safety such as crimes, etc.	3.48	1.08	586
Motor vehicle accident reports	3.57	1.16	567
Workplace safety inspections	3.11	1.06	438
Regulation of licensed professionals such as doctors and lawyers	3.15	1.14	499
Information about farming/agricultural activities	3.30	1.08	397

\*Responses were coded 5 = very much; 4 = somewhat; 3 = neutral; 2 = not; 1 = not at all

• *Attitudes toward different government levels*

Regarding the accuracy of government data, the reporters felt state government data (M=3.74) were most accurate, followed by federal (M=3.73) and local government (M=3.58) data. The one-way ANOVA test showed a significant difference among the three levels of government (F = 7.46, p < .01). A follow-up Tukey test showed both federal government and state government were more accurate than local government, though no difference was shown between federal government and state government. In terms of transparency, the reporters said the local government (M=3.20) was most transparent followed by state government (M=3.19) and federal government (M=2.99). The ANOVA test showed a significant difference among the three levels of government (F = 9.28, p < .001). The Tukey test showed both state government and local government were more transparent than federal government. No difference was observed between state government and local government. Regarding accountability of government, the reporters also said the local government (M=3.45) was most accountable followed by state government (M=3.11) and federal government (M=2.84). The ANOVA test showed a significant difference among the three levels of government (F = 57.33, p < .001). The Tukey test showed that local government was the most accountable, followed by state government. Federal government was considered the least accountable.

Table 5 Means and standard deviations for respondents' attitudes toward different levels of government

Variables*	Mean	SD	N
Accuracy of federal government data	3.73	.75	533
Accuracy of state government data	3.74	.75	570
Accuracy of local government data	3.58	.88	571
Transparency of federal government	2.99	.88	549
Transparency of state government	3.19	.81	582
Transparency of local government	3.20	1.01	582
Accountability of federal government	2.84	1.00	560
Accountability of state government	3.11	.89	583
Accountability of local government	3.45	.98	583

\*Responses were coded 5 = very much; 4 = somewhat; 3 = neutral; 2 = not; 1 = not at all

Table 6 One-way analyses of variance for respondents' attitude toward federal, state and local governments

Variables*	Governments			F	df	Significance
	Federal mean (SD)	State mean (SD)	Local mean (SD)			
Accuracy	3.73 (.75)	3.74 (.75)	3.58 (.88)	7.46	2	p < .01
Transparency	2.99 (.88)	3.19 (.81)	3.20 (1.01)	9.28	2	p < .001
Accountability	2.84 (1.0)	3.11 (.89)	3.45 (.98)	57.33	2	p < .001

\*Responses were coded 5 = very much; 4 = somewhat; 3 = neutral; 2 = not; 1 = not at all

• *Use of government data*

The reporters, on average, spend three to four days a week (M=3.62) using government websites for reporting purposes. They reported being generally successful (M=3.89) in finding data on the websites. The reporters were neutral in their attitude about the ease in searching for (M=3.11), using (M=3.20) and understanding government data online (M=3.30). The reporters had a medium level of need (M=3.10) for a third party to help them understand government data. When the reporters analyzed government data, they used a spreadsheet (M=2.84) most often, followed by geographic information system (M=2.27) and database manager (M=1.97). The statistical programs such as SPSS and SAS (M=1.42) were the least-used tools used by the reporters for data analysis.

Table 7 Means and standard deviations for respondents' use of government data

Variables	Mean	SD	N
Time in a week spent using a government website to obtain data for reporting purposes (in days)	3.62	1.66	598
Frequency of success in finding needed information on a government website*	3.89	.73	580
Difficulty in searching for government data online**	3.11	.71	573
Ease of using online government data online**	3.20	.81	571
Understandability of government data online***	3.30	.79	576
Necessity to have a third-party organization to help understand government data****	3.10	1.12	590
Frequency of using spreadsheet to analyze government data*	2.84	1.32	597
Frequency of using database manager to analyze government data*	1.97	1.11	587
Frequency of using geographic information system to analyze government data*	2.27	1.19	589
Frequency of using statistical software such as SPSS and SAS*	1.42	.83	586

\*Responses were coded 5 = very often; 4 = often; 3 = sometimes; 2 = rarely; 1 = never

\*\*Responses were coded 5 = very easy; 4 = easy; 3 = neutral; 2 = difficulty; 1 = very difficult

\*\*\*Responses were coded 5 = very understandable; 4 = understandable; 3 = neutral; 2 = not understandable; 1 = not at all understandable

\*\*\*\*Responses were coded 5 = very necessary; 4 = necessary; 3 = neutral; 2 = not necessary; 1 = not at all necessary

## II. Results for Missouri journalists

### • Demographics

A total of 82 newspaper reporters in Missouri took part in the phone survey. Similar to the composition of the sample of all reporters, seven percent (N=8) work as both reporter and editor and the rest work as reporters only. Nearly two thirds (N=53) are male and 35 percent (N=29) are female. The respondents' ages range from 22 to 70 years and the average age is 44. An overwhelming majority (N=80) of the respondents work at daily newspapers and only one reporter works at a weekly newspaper. The work experience of the reporters at their current newspapers and as paid journalists ranges from less than one year to nearly 50 years. On average, the reporters have worked as a paid journalist for about 19 years and have worked at their newspapers for nearly 11 years. The Missouri journalists' experience as a paid journalist and at their current newspapers are longer than the average years of experience of all the journalists in the national sample. Nearly 55 percent (N=45) of the reporters work in the print version of their newspapers and 37 percent (N=37) work in the online version of their newspapers. Compared to the national sample, the Missouri sample has more reporters working for newspapers' online version. The reporters' beats include government affairs, business, environment and health. There were very few reporters in science and sports. The size of the newspapers' staffs ranges from three people to nearly 800. The weekday circulation and the weekend circulation of the newspapers ranges from 25,000 to more than 200,000. The average size of weekday circulation of the newspapers is between 50,000 and 75,000. The average size of weekend circulation of the newspapers is slightly more than 75,000. Like their national counterparts, the Missouri reporters' political views tend to be neutral (M=4.32), though some reporters said they were very liberal and very conservative.

**Table 8** Means and standard deviations for age, political ideology, work experience at current news organization, work experience as a paid journalist, weekday circulation, weekend circulation, and size of news organization

Variables	Mean	SD	N
Age (in years)	43.65	12.91	81
Political ideology*	4.32	1.28	74
Work experience at current news organization (in years)	11.15	10.72	82
Work experience as a paid journalist	18.63	12.20	82
Weekday circulation**	3.67	2.75	70
Weekend circulation**	4.17	2.71	65
Size of news organization***	83.61	115.32	72

\*Responses were coded 7 = very liberal, 6 = somewhat liberal, 5 = slightly liberal, 4 = neutral, 3 = slightly conservative, 2 = somewhat conservative, 1 = very conservative

\*\*1 = 25,000 or less; 2 = 25,000 to 50,000; 3 = 50,000 to 75,000; 4 = 75,000 to 100,000; 5 = 100,000 to 150,000; 6 = 150,000 to 200,000; 7 = 200,000 or more

\*\*\*Number of journalists, ranging from 1 to infinity

Table 9 Percentages for newspaper type, beat specializations, online and print responsibilities, job responsibility and gender

Variables	%
Newspaper type	
Daily	97.6
Weekly	1.2
	100%
	(N = 82)
Beat specializations	
Business	14.6
Science	1.2
Education	9.8
Health	4.9
Police	8.5
Municipal affairs/Government	20.7
Technology	0
Sports	3.7
General assignment	18.3
Other	18.3
	100%
	(N=82)
Online vs print responsibilities	
Print	54.9
Print and online	45.1
	100%
	(N=82)
Job responsibility	
Reporter	92.7
Reporter and Editor	7.3
	100%
	(N=82)
Gender	
Male	64.6
Female	35.4
	100%
	(N=82)

• *Perception of government data openness and accessibility*

On average, government data are quite relevant ( $M=4.18$ ) to the reporters' daily work. Questions from 2a to 2k asked about reporters' opinions of how open government data categories should be to them. Similar to the national journalists, the reporters, on average, felt the government data should be open to them with campaign finance and lobbying information with the highest presumptions of openness ( $M=4.84$ ) and farming/agricultural activities the least open ( $M=4.29$ ). In terms of the accessibility of the data categories, the information about government employees and their salaries was the most accessible to the reporters ( $M=3.97$ ) and the information about the regulation of licensed professions was the least accessible ( $M=2.71$ ).

Table 10 Means and standard deviations for respondents' perception of government data openness

Variables*	Mean	SD	N
Relevance of government data to daily work	4.18	1.07	82
Government spending and contracts	4.79	.58	82
Identities of government employees and their salaries	4.57	.77	82
Environmental monitoring and enforcement	4.72	.68	81
Campaign finance and lobbying	4.84	.62	81
Public schools, education and teachers	4.69	.65	80
Vital records such as deaths, births and marriages	4.43	.96	81
Public safety such as crimes, etc.	4.74	.58	82
Motor vehicle accident reports	4.61	.66	82
Workplace safety inspections	4.60	.78	81
Regulation of licensed professionals such as doctors and lawyers	4.65	.79	81
Information about farming/agricultural activities	4.29	.95	76

\*Responses were coded 5 = very much; 4 = somewhat; 3 = neutral; 2 = not; 1 = not at all

Table 11 Means and standard deviations for respondents' perception of government data accessibility

Variables*	Mean	SD	N
Government spending and contracts	3.61	.91	76
Identities of government employees and their salaries	3.97	.97	75
Environmental monitoring and enforcement	3.21	.96	70
Campaign finance and lobbying	3.67	1.02	73
Public schools, education and teachers	3.50	.98	72
Vital records such as deaths, births and marriages	3.89	1.12	75
Public safety such as crimes, etc.	3.40	1.13	80
Motor vehicle accident reports	3.89	1.00	79
Workplace safety inspections	3.16	1.08	63
Regulation of licensed professionals such as doctors and lawyers	2.71	1.18	76
Information about farming/agricultural activities	3.36	1.16	59

\*Responses were coded 5 = very much; 4 = somewhat; 3 = neutral; 2 = not; 1 = not at all

• *Attitudes toward different government levels*

Different from the national journalists, the Missouri reporters said federal government data (M=3.73) were most accurate, followed by state (M=3.61) and local government (M=3.32) data. The one-way ANOVA test showed a significant difference among the three levels of government (F = 4.94, p < .01). A follow-up Tukey test showed that the respondents considered federal government more accurate than local government. No difference was shown between federal government and state government. No difference was shown between state government and local government. Although the reporters said the state government (M=3.15) was most transparent followed by local government (M=3.05) and federal government (M=3.01), the ANOVA test did not show a statistically significant difference among the three means (F = .46, p < .63). As to accountability of government, the reporters said the local government (M=3.30) was most accountable, followed by state government (M=3.19) and federal government (M=3.00). The ANOVA test didn't show difference among the three levels of government (F = 2.14, p = .12).

**Table 12** Means and standard deviations for respondents' attitudes toward different levels of government

Variables*	Mean	SD	N
Accuracy of federal government data	3.73	.71	74
Accuracy of state government data	3.61	.76	77
Accuracy of local government data	3.32	.95	77
Transparency of federal government	3.01	.92	76
Transparency of state government	3.15	.88	79
Transparency of local government	3.05	1.02	80
Accountability of federal government	3.00	.95	78
Accountability of state government	3.19	.78	80
Accountability of local government	3.30	1.01	80

\*Responses were coded 5 = very much; 4 = somewhat; 3 = neutral; 2 = not; 1 = not at all

**Table 13** One-way analyses of variance for respondents' attitudes toward federal, state and local governments

Variables*	Governments			F	df	Significance
	Federal mean (SD)	State mean (SD)	Local mean (SD)			
Accuracy	3.73 (.71)	3.61 (.76)	3.32 (.95)	4.94	2	p < .01
Transparency	3.01 (.88)	3.15 (.81)	3.05 (1.01)	.46	2	n.s.
Accountability	3.00 (.95)	3.19 (.78)	3.30 (1.01)	2.14	2	n.s.

\*Responses were coded 5 = very much; 4 = somewhat; 3 = neutral; 2 = not; 1 = not at all

• *Use of government data*

The reporters, on average, spend three to four days a week (M=3.63) using government websites for reporting purposes. They were generally successful (M=4.01) in finding data on the website. The reporters were neutral in their attitude about the ease in searching for (M=3.21), using (M=3.12) and understanding government data online (M=3.29). The reporters had a medium level of need (M=3.28) for a third party to help them understand government data. When the reporters analyzed government data, they used a spreadsheet (M=2.95) most often, followed by a database manager (M=2.20) and geographic information system (M=2.15). The statistical programs, such as SPSS and SAS (M=1.42), were the least used tools used by the reporters for data analysis.

Table 14 Means and standard deviations for respondents' use of government data

Variables	Mean	SD	N
Time in a week spent using a government website to obtain data for reporting purposes (in days)	3.63	1.73	82
Frequency of success in finding needed information on a government website*	4.01	.75	78
Difficulty in searching for government data online**	3.21	.63	78
Ease of using online government data online**	3.12	.81	78
Understandability of government data online***	3.29	.76	78
Necessity to have a third-party organization to help understand government data****	3.28	1.17	80
Frequency of using spreadsheet to analyze government data*	2.95	1.33	82
Frequency of using database manager to analyze government data*	2.20	1.09	81
Frequency of using geographic information system to analyze government data*	2.15	1.20	80
Frequency of using statistical software such as SPSS and SAS*	1.42	.83	79

\*Responses were coded 5 = very often; 4 = often; 3 = sometimes; 2 = rarely; 1 = never

\*\*Responses were coded 5 = very easy; 4 = easy; 3 = neutral; 2 = difficulty; 1 = very difficult

\*\*\*Responses were coded 5 = very understandable; 4 = understandable; 3 = neutral; 2 = not understandable; 1 = not at all understandable

\*\*\*\*Responses were coded 5 = very necessary; 4 = necessary; 3 = neutral; 2 = not necessary; 1 = not at all necessary