

December 2002

HOMEPAGE GENRE DIMENSIONALITY

Terry Ryan
Claremont Graduate University

Richard Field
University of Alberta

Lorne Olfman
Claremont Graduate University

Follow this and additional works at: <http://aisel.aisnet.org/amcis2002>

Recommended Citation

Ryan, Terry; Field, Richard; and Olfman, Lorne, "HOMEPAGE GENRE DIMENSIONALITY" (2002). *AMCIS 2002 Proceedings*. 154.
<http://aisel.aisnet.org/amcis2002/154>

This material is brought to you by the Americas Conference on Information Systems (AMCIS) at AIS Electronic Library (AISeL). It has been accepted for inclusion in AMCIS 2002 Proceedings by an authorized administrator of AIS Electronic Library (AISeL). For more information, please contact elibrary@aisnet.org.

HOMEPAGE GENRE DIMENSIONALITY

Terry Ryan
Claremont Graduate University
Terry.Ryan@cgu.edu

Richard H. G. Field
University of Alberta
Richard.Field@ualberta.ca

Lorne Olfman
Claremont Graduate University
Lorne.Olfman@cgu.edu

Abstract

Prior theory stipulates that three dimensions can explain the appearance of electronic media genres: content, form, and function. This paper reports an analysis of subjective groupings of U.S. state government homepages to determine if the three dimensions do in fact underlie the visual appearance of this genre. Homepage images were collected—one per state per year for all available states for 1998-2001—from a publicly available archive of Web sites. Subjects sorted the images for each year into groups according to perceived similarities. Multidimensional scaling was used to identify dimensions that account for the groupings. The empirically derived dimensions were examined in light of theoretical expectations. A four-dimensional solution was found to fit the data, the first three dimensions of which correspond to content, form, and function. Directions for future research are discussed.

Keywords: Homepage, genre, dimensionality, Web, archive

Introduction

To communicate in any medium, writers and readers need to know what works, what to expect, what to include, and how content should be presented. Genres are patterns of communication in media that serve to guide writers and readers (Bazerman 1994). Each genre defines an appearance that materials in a medium can be expected to take (Agre 1998). Yates, Orlikowski, and Okamura (1999) found that genres serve as institutionalized templates for social interaction and organizing structures for communications. Genres are shared between readers and writers, who produce and reproduce them together (Orlikowski et al. 1995).

Yates and Orlikowski (1992) coined the term ‘cybergenre’ to refer to genres of electronic communications in organizations. A great deal of research has been done on cybergenres since that time. A key finding is that as a result of ongoing usage, genres change over time. Genres can change as media capabilities change. They can change as communicators adopt innovative techniques for media use. They can change due to the migration of existing genres from other media (Yates et al. 1999). Genres may disappear through a process of selection, as other new genres become preferred. As one example, Yates (1989) described how the memo and report communication systems in American businesses evolved from 1850 to 1920. Erickson (1999) found that new conventions were developed when a participatory community moved to using on-line genres. Web homepages have been described as a cybergenre (Dillon and Gushrowski 2000; Roberts 1998).

The empirical study of homepages from a genre perspective is made more difficult by the ephemeral nature of the pages on the Web. Web pages seem to change and disappear unpredictably. Combating this, there have been two major efforts to collect pages from the web before they change and are gone forever. One is the “Ghostsites” collection (<http://www.disobey.com/ghostsites/>), which had 1,082 screenshots from approximately 700 Web projects between 1998 and 2002 as of March 3, 2002. The other is the Internet Archive (Wayback Machine -- <http://web.archive.org>), which has been collecting snapshots of the entire web since 1996 (Boyle 1997). It is possible to access the Internet Archive’s servers and retrieve pages, including homepages, from 1996 to the present. These pages span the entire Internet.

Given this resource, a number of questions can be asked that were not answerable before about the way web genres have changed over time and within populations of web pages. Moreover, extrapolations can be made for the future of web genres based on the historical evidence of what has happened so far. In order to accomplish these goals, it is first necessary to clearly define the dimensions of cybergenres.

At present, there is no complete agreement over cybergenre dimensionality. Yates and Orlikowski (1992) and Orlikowski and Yates (1994) proposed two cybergenre dimensions: substance (or purpose) and form. These ideas represent answers to the questions “What is the page about?” and “How is it presented?” Watters and Shepherd (1997) and Shepherd and Watters (1999) added interactivity (or functionality) to the list of cybergenre dimensions, that is, “How is the digital document accessed and used?” In addition to these authors, many others have discussed cybergenre dimensionality, including Crowston and Williams (2000), Eriksen and Ihlström (2000), and McDonnell, Koehler, and Carroll (2000). Although there is not total agreement by all authors, it seems that at least three dimensions characterize cybergenres: content (what the page means), form (how the page is laid out), and function (how the page can be used).

The questions asked by this research are: To what extent are the three dimensions of content, form, and function found in organizational homepages? To what extent do these findings change over time?

Method

Population Studied

Given the massive size of the World Wide Web—even the limited version that is stored on the Internet Archive—it was necessary to choose a population to study. One consideration was that we wished to study homepages, rather than entire sites. Although there is much to be said for studying entire sites, as Nielsen and Tahir (2002) state, “the homepage is the most important page on any website, getting more page views than any other page” (p 1). The population chosen for study in this research was US State homepages having URL’s of the form ‘www.state.xx.us’ where ‘xx’ is the two-digit state code.

One reason for choosing this group of homepages was that it is static in size. Another was that the sites themselves are unlikely to disappear. Yet another was that the domain address form of ‘www.state.xx.us’ allows the researcher to be certain that the top-level state page—the homepage for the state—is being accessed. A final reason was that choosing US state government homepages provided an entire population to study, not merely a sample (making the assumption that each state homepage for the time under consideration would be retrievable through the Internet Archive). As an additional consideration, a population size of 50 was deemed large enough to guarantee variation among homepages, yet small enough to be manageable.

State homepage images were gathered for the year 2001 from the live Web sites for each state (not from the Internet Archive site) using MS Internet Explorer during a two-day period, November 13 and 14, 2001. Collecting the 2001 homepage images live meant that all could be collected at the same time and ensured that all 50 states would be represented in the sample for that year.

The Internet Archive site was used to gather state homepage images for the years 1996 to 2000. When a state address is entered in the “Wayback Machine,” the archive provides a listing of all pages gathered for that site by year. An example for the state of Alaska is provided in Figure 1.

Initially, it was planned that each homepage image collected would be the last listing for its state for each year from 1996 to 2000. To illustrate, using the data in Figure 1 the Alaska homepages to be gathered would be: December 23, 1996; December 11, 1997; December 12, 1998; April 29, 1999; and December 07, 2000. The intent was to capture homepages that are approximately one year apart, given a starting date of November 2001. In practice, however, it was not possible to gather the very last state homepage for every state for every year. Sometimes the homepage shown in the Internet Archive was actually from a different site. At other times, the homepage shown was not available in the Archive. On other occasions, a state homepage would only partially load from the archive. The goal in gathering homepages became to obtain a full representation of the homepage as close in time as possible to the last homepage listed in the Archive for that state for that year. In the Alaska case, the actual homepages gathered were December 23, 1996; December 11, 1997; June 10, 1998; April 27, 1999; and September 17, 2000. The time intervals between captured homepages for a given state are therefore not exactly one year. For this study, state homepages from 1998 through 2001 were used.



Figure 2. Alaska to Illinois Official State Home Pages, 1999



Figure 3. Alaska Home Pages, 1996 to 2001

Subjects

Subjects were mostly undergraduate students (and some MBA students) at a large public university in Western Canada. Subjects were recruited by poster, e-mail, and word-of-mouth. Subjects were paid \$10 CDN for their participation. The second author collected all of the data.

Each subject (n=120, 30 per year for 1998 to 2001) was given a set of state homepages for a particular year. Each subject was told that the research concerned the design of homepages and how they change over time. They were asked to create groupings of homepages where each grouping has homepages that are similar to each other, but different from pages in other groups. They were told that there is no correct number of groups.

Data were coded into an agreement matrix for each subject and aggregated over the subjects within a year as per Hair, Anderson, Tatham, and Black (1998). The SPSS Multi Dimensional Scaling (MDS) program was used to determine the dimensions of the data. Euclidian distances were created from the data using counts and ordinal measurement. A minimum of two and maximum of six dimensions were examined. Because we had been led by theory to expect at least three dimensions, we decided to limit the maximum number of dimensions in a MDS solution six. This would allow for the discovery of some additional dimensions, if such were present in the data. We did not believe that analysis would reveal more than twice as many dimensions as we had anticipated based on theory.

Results

Descriptive Data

Subjects created homepage groupings that on average had nearly seven groups, a largest group of about 16 members and a smallest group of about three members (Table 1). No subjects reported difficulties in sorting the state homepages into groups, and most completed the task within one half hour. Debriefing notes were taken to capture each subject's reasoning for the groupings that that person had created. However, these are not reported in this paper.

Table 1. Descriptive Data

Descriptive Data Means				
Year	# of groups	largest group	smallest group	# of states
1996	4.80	9.40	2.80	27
1997	6.80	14.53	2.50	46
1998	6.43	15.57	3.13	49
1999	6.80	15.10	2.90	49
2000	5.40	15.97	3.70	49
2001	5.93	17.27	3.13	50
Grossed Up to 50 States				
Year	# of groups	largest group	smallest group	# of states
1996	8.89	17.41	5.19	50
1997	7.39	15.80	2.72	50
1998	6.56	15.88	3.20	50
1999	6.94	15.41	2.96	50
2000	5.51	16.29	3.78	50
2001	5.94	17.28	3.14	50
Mean	6.86	16.00	3.46	

A plot of the number of groups by year shows a trend of fewer groupings over time (Figure 4). This data is suggestive that the number of groupings found by subjects in the data is declining, from nearly nine in 1996 to six by the year 2001. It is interesting to speculate that a better-defined and smaller genre repertoire is developing over time in this population of state homepages. This process might be explicit (Yates et al. 1999) as state webmasters read books by opinion leaders on website design (e.g., Nielsen

1993, 2000), read magazines directly targeted to e-government (such as Electronic Government and NetGov), and consider the results of award competitions (such as that run by <http://www.govtech.net/>). It might also be an implicit process as the state webmasters go to technology and e-government conferences, and compare their own state's website to those of the other states.

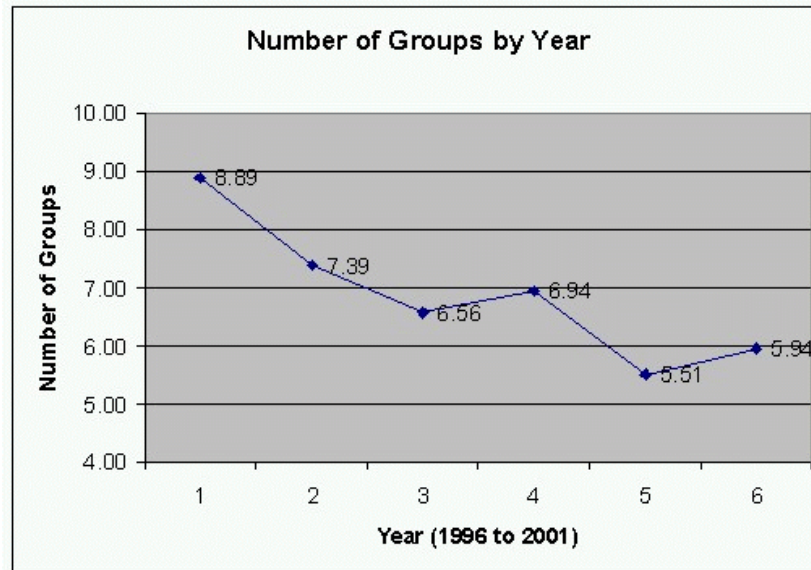


Figure 4. Average Number of Subject Groupings by Year

Dimensionality Results

Researchers use MDS to understand the distances between objects in terms of underlying dimensions. For distances between every pair of N objects in a set, MDS can be used to “find a representation of the items in a few dimensions such that the interitem proximities ‘nearly match’ the original similarities (or distances).” (Johnson and Wichern 1998)

Guidelines for deciding on the dimensionality of our MDS analysis of the homepage subjective grouping task are provided by Davison (1983). He states that the analyst should obtain several solutions in different dimensionalities and then choose between them based on three criteria. They are interpretability, fit to the data, and reproducibility. This last criterion can be used when there are two or more sub-samples that come from the same population. The idea is to select the number of dimensions that emerge consistently in the separate sub-samples and then stick with that answer.

A key metric of MDS is stress, which is a measure of the closeness of the match between the dimensions derived and the distances between original data. Johnson and Wichern (1998) provide a fit guideline for MDS based on the stress metric. They advise that when stress is .20, goodness of fit is ‘Poor’; at .10, fit is ‘Fair’; at .05, fit is ‘Good’; and at .025, fit is ‘Excellent’. The final criterion for deciding the best dimensionality in the results is that the user should look for the best fit with the smallest number of dimensions. A scree plot can be used to look for substantial improvement moving from n to $n+1$ dimensions (Hair et al. 1998).

Examining the dimensionality data in Table 2, the first question is whether the minimum solution of two dimensions provides a good fit to the data. The answer to this question is negative since stress is between ‘Fair’ and ‘Poor’ for the years 1999, 2000, and 2001, and ‘Fair’ for 1998. For three dimensions, fit improves to ‘Fair’; stress is less than .10 for all four years. Further, the literature indicates that it requires at least three dimensions (viz., content, form, and function) to understand cybergenres. Therefore, a minimum of three dimensions is required to fit the data.

Next, the task shifts to choosing the best solution from those having from three to six dimensions. For three dimensions, stress is ‘Fair’; for four dimensions, stress is almost ‘Good’; and for five dimensions, it is ‘Good’. The six-dimensional solution can be ruled out, as the simpler five-dimensional solution provides a good fit (Hair et al. 1998).

Table 2. MDS Dimension Results

Year	Measure	Number of Dimensions				
		6	5	4	3	2
2001	S-Stress	0.03039	0.03791	0.04647	0.05847	0.07804
	Stress	0.03546	0.04756	0.06084	0.08173	0.11618
	RSQ	0.99278	0.98817	0.98276	0.97308	0.95656
2000	S-Stress	0.03494	0.04105	0.04738	0.08061	0.09824
	Stress	0.03592	0.04422	0.05372	0.08656	0.11088
	RSQ	0.99079	0.98715	0.98284	0.96434	0.95224
1999	S-Stress	0.03083	0.03890	0.04797	0.05777	0.06450
	Stress	0.03607	0.04830	0.06734	0.09147	0.11185
	RSQ	0.99285	0.98855	0.98140	0.97176	0.96471
1998	S-Stress	0.02734	0.03158	0.03974	0.05238	0.06300
	Stress	0.03385	0.04113	0.05524	0.07741	0.09419
	RSQ	0.99337	0.99101	0.98567	0.97568	0.96970

Scree results of gain in stress scores are shown in Table 3. The improvement (lowering) in stress values when moving from three to four dimensions is meaningful, averaging .025. When moving from four to five dimensions the gain drops to .014. Therefore, four dimensions seem to fit the data best.

Table 3. Scree Results of Dimensionality

Year	Scree Results		
	Stress Gain in Moving from N to N+1 Dimensions		
	3 --> 4 Dim. Gain	4 --> 5 Dim. Gain	4 --> 5 to 3 --> 4
2001	0.02089	0.01328	64%
2000	0.03284	0.0095	29%
1999	0.02413	0.01904	79%
1998	0.02217	0.01411	64%
Mean	0.02501	0.01398	59%

Dimensionality Interpretation

The goal of this paper was to determine to what extent the theoretical cybergenre dimensions content, form, and function could be detected from subjective analyses of US State homepages, and whether these dimensions varied over time. The analysis proceeded at this point by taking the four dimensional solution for each year of the data and preparing sheets of pictures of the homepages falling at either end of each of the dimensions. For the year 1998, Figure 5 shows the six state homepages falling at one end of Dimension 1 (on the left of the figure) and the three state homepages falling at the other end of that dimension. We viewed homepages as ‘falling’ at the end of a dimension based on the absolute value of their dimensional scores for that dimension.

Examining the Dimension 1 endpoint exemplars in Figure 5, it was necessary to decide what theoretical label best describes the dimension. Does the first dimension—the one that explains the most variation in the data—represent the content of the homepages, their form, or their functionality? A case could be made for each. States on the left of the figure have lower content. There is less information on those pages. Their form (that is, their visual appearance and look and feel of the page) is that of a header and left menu bar. Their functionality, the user interface or means of access used on the page, is that of hierarchical menus versus the link lists shown on the right of the figure.



Figure 5. 1998 Dimension 1 End Point Exemplars

With some uncertainty about the meaning of the first dimension, we moved on to Dimension 2 (see Figure 6). This dimension does not represent content, as there are examples of both high and low content on each end of the dimension. It also does not appear to represent functionality. Each side of the figure has examples of many versus few links and categories versus lists of links. The second dimension seems to represent form. States on the left of Figure 6 all have a header bar and a left menu bar whereas those on the right have a scattering of links and categories without apparent structure. Dimension 3 (Figure 7) can be interpreted as content, with states of high content on the right of the figure and those of low content on the left.



Figure 6. 1998 Dimension 2 End Point Exemplars

By elimination, dimension 1 of the 1998 data can best be interpreted as functionality. The user interface of the states on the left of Figure 5 allows navigation by categories (buttons) as opposed to the text links used by the states on the right of the figure.

Because theory does not provide a fourth cybergenre dimension, Figure 8 must be examined and an interpretation proposed. One possibility is that the fourth dimension is full-color background. Those homepages with white backgrounds are on the left of the

figure, those with full color backgrounds on the right. This type of analysis was repeated for the years 1999 to 2001, with the results shown in Table 4.¹

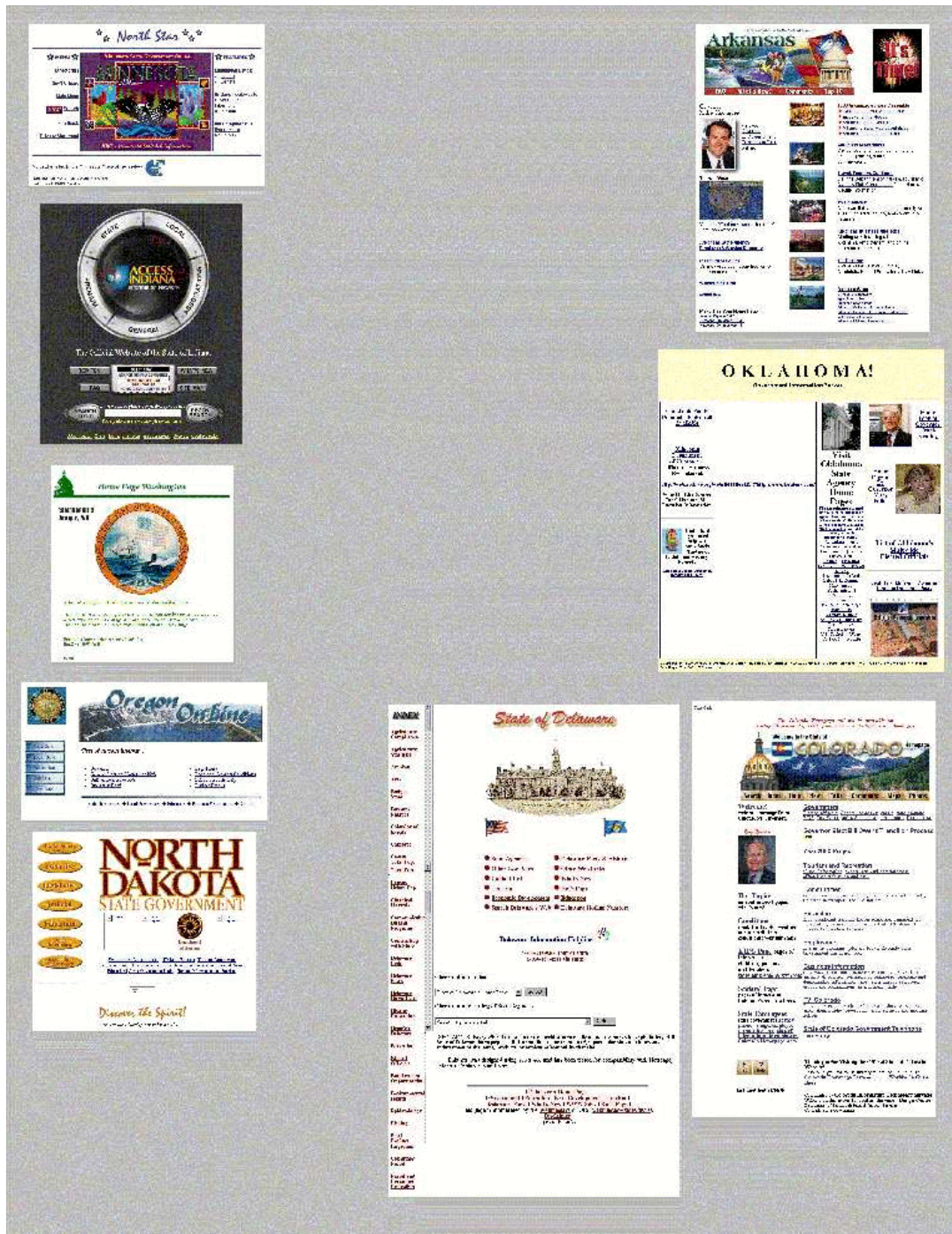


Figure 7. 1998 Dimension 3 End Point Exemplars

¹Due to limitations on the length of papers, all homepage images cannot be shown here, but the interested reader can find them at <http://www.bus.ualberta.ca/rfield/Research6.htm> under the heading Web Genre Dimensionality.

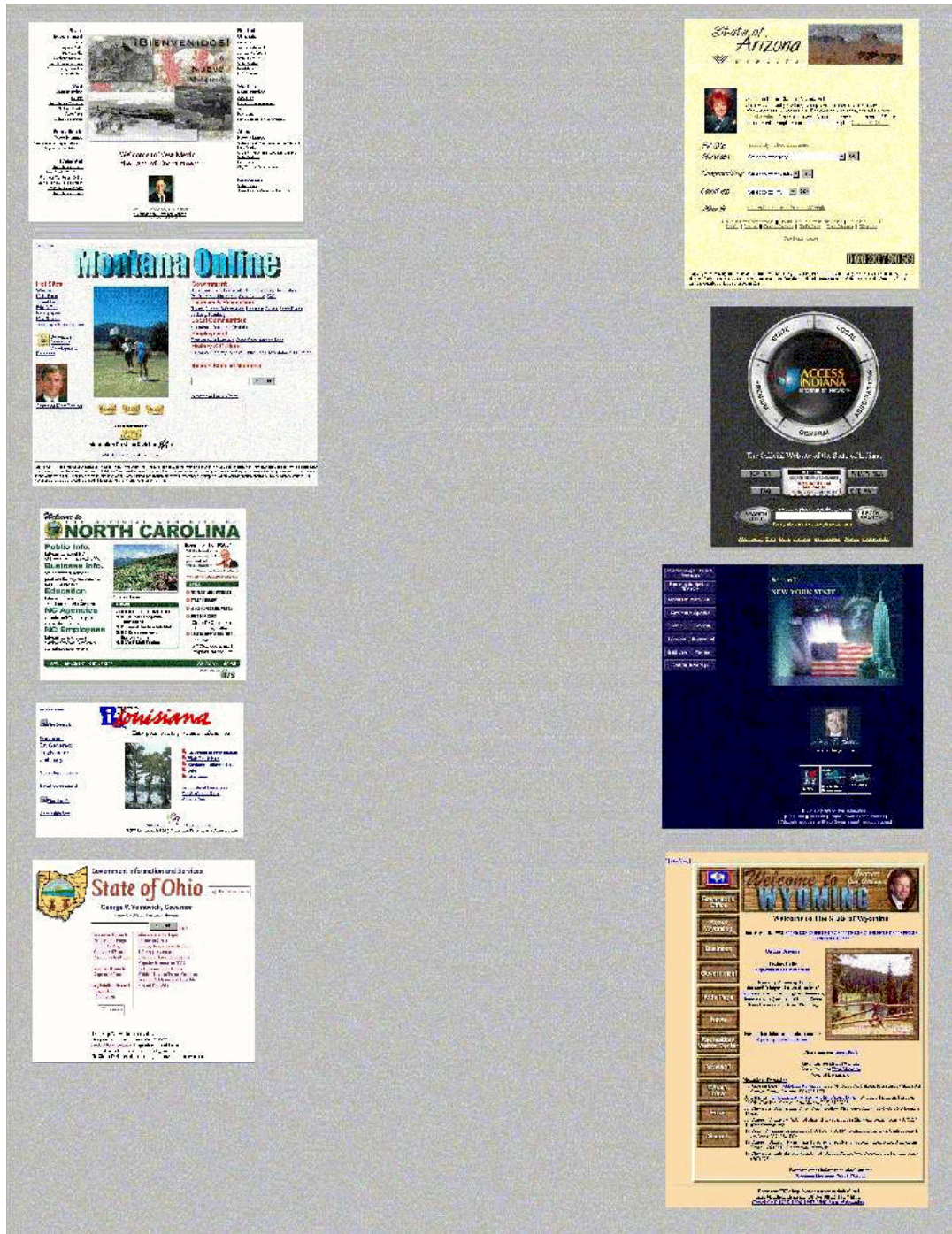


Figure 8. 1998 Dimension 4 End Point Exemplars

Dimensions over Time

In MDS analysis, the first dimension is the one that accounts for the most variance in the groupings made by the subjects. As the web has evolved over the years of 1998 to 2001, our subjects found more variation in content. From the third-most important dimension in 1998, content became second-most important in 1999 and the most important dimension in 2000 and 2001. We can

speculate that the web is polarizing into high content ‘portal’ homepages and low content ‘entry page’ homepages. We can check this later with our homepage materials by coding for page attributes and mapping them over time.

Another finding is that early variation in functionality found in the 1998 results became less pronounced as time passed. Web designers, in the process of shaping and being shaped by the web genres of the homepage, may have applied the same functionality lessons. Homepages that differed dramatically in 1998 on functionality have become more similar over time.

Table 4. Dimensions by Year

Year	1	2	3	4
1998	Functionality	Form	Content	Color
1999	Form	Content	Functionality	Color
2000	Content	Form	Functionality	Color
2001	Content	Form	Functionality	? (Not color)

Differences in form seem to fall into the middle range. The look and feel in a homepage may depend on recognizable choices that are available to designers. There may be no general agreement that one particular look and feel is the best but, rather, there may be a number of different available options. Genre theory states that the writers of communications in a genre have to decide what forms are best for their audiences and communicative situations. States may make similar choices about their audiences and communicative intents. Some might choose their current citizens as their audience and ‘e-government’ as their communicative purpose. Other states, especially those focused on tourism, might take a different, more outward-looking, focus.

The fourth dimension of a full-color background seems to have disappeared by 2001. An inspection of the 2001 state homepages shows that color is now prominent in many homepages and that full color backgrounds are disappearing in favor of white ones. This seems to be an example of how genres evolve incrementally and go through a process of natural selection (Agre 1998). Full color backgrounds were tried, but failed to survive.

Conclusion

Electronic communications of organizations can be classified into types, or genres, which may be viewed as institutionalized templates or organizing structures for communication (Yates and Orlikowski 1992; Yates et al. 1999). Some writers (e.g., Agre 1998) have suggested that genres evolve incrementally, changing over time due to technological development, competition, and natural selection. Others (e.g., Yates et al. 1999) have suggested that genres change due to non-evolutionary influences, both explicit—including planned modifications—and implicit—including migration of genres from other media.

This study, using multidimensional scaling, identified four dimensions that explain perceived similarities among US state government WWW homepage images. The dimensions identified map well to the determinants of genres specified in theory: form, functionality, and content. One contribution of this study therefore is the empirical documentation of theoretically proposed cybergenre dimensions. The other contribution of this study is the finding that the variation explained by these three dimensions has changed over time. Content has exhibited greater variation as the web has evolved, while functionality has shown decreased variation.

While the results of the study have value for both theory and practice, the study does have some limitations. There may be critical differences between US state government homepages and the homepages of non-government entities or other government entities. This is an issue that could be addressed in later studies. In addition, the use of college students as judges of homepage similarity could be faulted to the extent that they are systematically different from other possible judges. This could be addressed in further research. Future studies might also seek to identify genres of state government homepages using clustering analysis on the dimensional scores from this study. In addition to identifying genres using dimensions based on ‘subjective’ similarity grouping data, the homepage images will be coded for a variety of attributes related to form, functionality, and content. This data can be used to examine how the genres vary in terms of these ‘objective’ attributes.

Our planned program of research will consider the nature of the genres that have appeared in state government homepages over a six-year period—1996 to 2001—and how these genres have changed over time. It will address a number of research questions: What state government homepage genres exist now? What are the key attributes of each genre? Has genre change been more

nearly evolutionary or revolutionary? It is important that states decide which genre they want to be in. Since genres are organizing structures for communications, each state has to decide which template to choose for the types of interactions it wants with its web users, such as state citizens, potential visitors, and business users.

References

- Agre, P. E. "Designing Genres for New Media: Social, Economic, and Political Contexts," in *CyberSociety 2.0: Revisiting CMC and Community*, Steve Jones (ed.), Sage, Thousand Oaks, CA, 1998.
- Bazerman, C. "Systems of Genres and the Enactment of Social Intentions," in *Genre and the New Rhetoric*, Aviva Freedman and Peter Medway (eds.), Taylor and Francis, Bristol, PA, 1994, pp. 79-101.
- Boyle, A. "Archiving the Internet for Posterity," 1997, <http://web.archive.org/web/19970810233147/www.msnbc.com/ews/60659.asp>.
- Crowston, K., and Williams, M. "Reproduced and Emergent Genres of Communication on the World Wide Web," *The Information Society*, (16:3), 2000, pp. 201-215.
- Davison, M. L. *Multidimensional Scaling*, Wiley, New York, 1983.
- Dillon, A., and Gushrowski, B. A. "Genres and the Web: Is the Personal Homepage the First Uniquely Digital Genre?" *Journal of the American Society of Information Science* (51:2), January 2000, pp. 202-205.
- Erickson, T. "Rhyme and Punishment: The Creation and Enforcement of Conventions in an On-Line Participatory Limerick Genre," in *Proceedings of the Thirty-Second Annual Hawaii International Conference on System Sciences, IEEE Computer Society, Los Alamitos, CA, 1999*.
- Eriksen, L. B., and Ihlström, C. "Evolution of the Web News Genre – The Slow Move Beyond the Print Metaphor," in *Proceedings of the Thirty-Third Annual Hawaii International Conference on System Sciences, IEEE Computer Society, Los Alamitos, CA, 2000*.
- Hair, Jr., J. F., Anderson, R. E., Tatham, R. L., and Black, W. C. *Multivariate Data Analysis* (5th ed.), Prentice-Hall, Upper Saddle River, NJ, 1998.
- Johnson, R. A., and Wichern, D. W. *Applied Multivariate Statistical Analysis* (4th ed.), Prentice Hall, Englewood Cliffs, NJ, 1998.
- McDonnell, J., Koehler, W., and Carroll, B. "Cataloging Challenges in an Area Studies Virtual Library Catalog (ASVLC): Results of a Case Study," *Journal of Internet Cataloging* (2:2), 2000, pp. 15-42.
- Nielsen, J. *Usability Engineering*, Morgan Kaufmann, San Diego, 1993.
- Nielsen, J. *Designing Web Usability: The Practice of Simplicity*, New Riders Publishing, Indianapolis, 2000.
- Nielsen, J., and Tahir, M. *Homepage Usability: 50 Websites Deconstructed*, New Riders Publishing, Indianapolis, 2002.
- Orlikowski, W. J., and Yates, J. "Genre Repertoire: The Structuring of Communicative Practices in Organizations," *Administrative Science Quarterly* (39:4), December 1994, pp. 542-574.
- Orlikowski, W. J., Yates, J., Okamura, K., and Fujimoto, M. "Shaping Electronic Communication: The Metastructuring of Technology in the Context of Use," *Organization Science* (6:4), July-August 1995, pp. 423-444.
- Roberts, G. F. "The Home Page as Genre: A Narrative Approach," in *Proceedings of the Thirty-First Annual Hawaii International Conference on System Sciences, IEEE Computer Society, Los Alamitos, CA, 1998*.
- Shepherd, M. A., and Watters, C. "The Functionality Attribute of Cybergenres," in *Proceedings of the Thirty-Second Annual Hawaii International Conference on System Sciences, IEEE Computer Society, Los Alamitos, CA, 1999*.
- Watters, C., and Shepherd, M. A. "The Role of Genre in the Evolution of Interfaces for the Internet," 1997, <http://web.archive.org/web/19990202201230/http://net97.dal.ca/970326-03/>
- Yates, J. *Control through Communication: The Rise of System in American Management*, John Hopkins University Press, Baltimore, 1989.
- Yates, J., and Orlikowski, W. J. "Genres of Organizational Communication: A Structural Approach to Studying Communication and Media," *Academy of Management Review* (17:2) 1992, pp. 299-326.
- Yates, J., Orlikowski, W. J., and Okamura, K. "Explicit and Implicit Structuring of Genres in Electronic Communication: Reinforcement and Change of Social Interaction," *Organization Science* (10:1), January-February 1999, pp. 83-103.