ELSEVIER

Information & Management 37 (2000) 123–134

INFORMATION & MANAGEMENT

www.elsevier.com/locate/dsw

The content and design of web sites: an empirical study

Eelko K.R.E. Huizingh (Associate Professor)*

Faculty of Economics, University of Groningen, P.O. Box 800, 9700 AV Groningen, The Netherlands

Received 4 June 1998; received in revised form 20 April 1999; accepted 22 July 1999

Abstract

To support the emergence of a solid knowledge base for analyzing Web activity, we have developed a framework to analyze and categorize the capabilities of Web sites. This distinguishes content from design. *Content* refers to the information, features, or services that are offered in the Web site, *design* to the way the content is made available for Web visitors. Both concepts have been operationalized by means of objective and subjective measures to capture features as well as perceptions.

This framework has been applied to study how different groups of companies are using the Web for commercial purposes. We have compared Web sites based on their source, industry, and size. On average, larger Web sites seem to be 'richer' and more advanced. © 2000 Elsevier Science B.V. All rights reserved.

Keywords: Internet; Web sites; Web content; Web design

1. Introduction

For two decades, the Internet was used almost exclusively by scientific researchers. Although many big corporations had access to the Internet, their participation was limited to their research and engineering departments [14]. In a few years, the balance between commercial and non-commercial use of the Internet has changed dramatically. By the fall of 1994, the number of commercial Internet users exceeded the number of educational and research users [5]. A recent report [6] predicts that in 2002, US businesses will see between $360 and $480 billion in profits from Internet-based cost savings alone.

There is ample anecdotal evidence that suggests that the Internet is a fruitful tool for commercial purposes. However, there is not much known about how Web sites should be designed. As Hoffman and Novak [9] note, little is known about how to develop commercial Web sites to maximize profit. Berthon et al. [3] call for research to reveal the true nature and effectiveness of electronic commerce. Until now, companies tend to use traditional advertising concepts that do not take into account the interactive features of the Web. Lack of a good Web methodology is listed as extremely important [13].

The goal of this paper is to determine how different groups of companies are actually using the Web and their sites are structured. To provide focus to our study, we developed a framework with two essential aspects of Web sites: content and design. We have analyzed different groups of sites to gain understanding how the new tools are used by various companies. Web sites have been compared based upon their size and the industry. A third characteristic is the source of the Web

* Tel.: +31-503633779/31-503637065; fax: +31-503637207.
E-mail address: huizingh@eco.rug.nl (E.K.R.E. Huizingh).

0378-7206/00/$ – see front matter
PII: S0378-7206(99)00044-0

site. We have compared mainly North American sites (in Yahoo) with sites from companies based in the Netherlands (the Dutch Yellow Pages). Any company that builds a Web site in *marketspace* [23] can be considered to be a multinational. However, in terrestrial marketplaces, most companies in any country focus (often exclusively) on its domestic market. A new, global medium will not necessarily change this strategy. To understand the differences between countries we investigated Web sites from both sources.

2. Research framework

In our study we focused on two important characteristics of Web sites: content and design. Both characteristics were measured by means of features (objectively) and perceptions (subjectively). Fig. 1 provides an overview of our framework.

2.1. Content of web sites

Many studies describing how to use the Web stress the importance of content. 'Content is king' is a well known slogan [4,17]. Web sites can contain various features, including full-color virtual catalogues, on-screen order forms, and questionnaires to elicit customer feedback. Less ambitious companies may use the Web to present themselves to potential customers and provide general company information. Providing *information* is the basic goal of a Web site [2]. We have distinguished between three categories of information. The first is between commercial and non-commercial information. Commercial information can provide insight into the *background of the company* and may involve the mission statement, financial statements, a list of important customers, and an overview of completed projects. This kind of information is important to present the organization to (potential) customers or other stakeholders (shareholders, the public, politicians, etc.). A Web site can also support the sales function with *product information*, such as prices, specifications, terms of delivery, and descriptions or animations of the product in action. Recently, a growing number of companies try to strengthen the relationship with their customers [7,24,25]. Companies can use their Web sites to provide information or

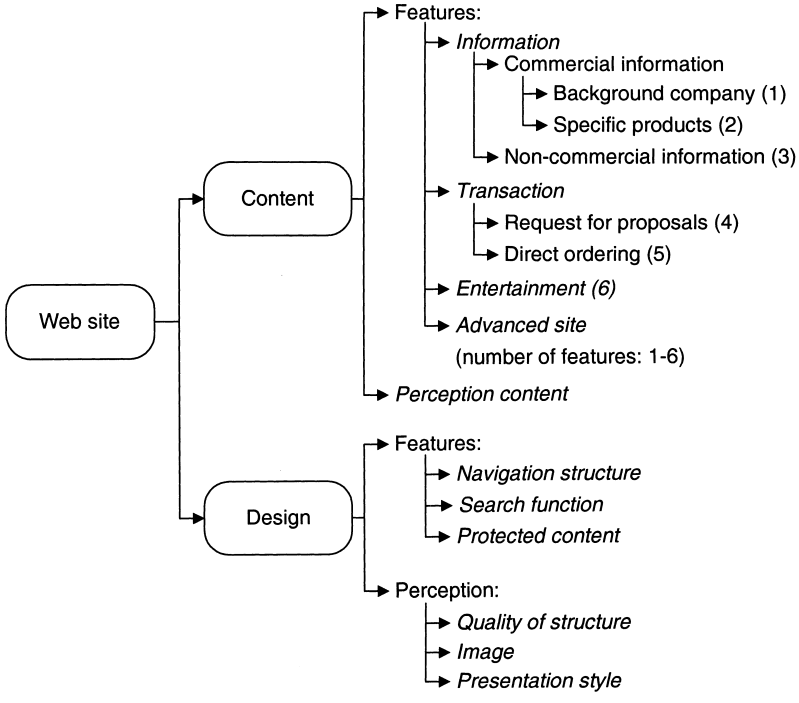

Fig. 1. The research framework that identifies the relevant aspects of the content and the design of Web sites.

links to information that is useful for their customers but that is not (directly) aimed at gaining new orders. *Non-commercial information* may relate to the company (sponsored events or new technologies), its industry, or its geographical location; e.g., a hotel may offer information on nearby parks or museums.

The second group of features comprises of *transaction* related features. Many studies have investigated the commercial potential of the Web based almost entirely on the direct sales function of Web sites. A transaction might be a simple purchase, but in more complex situations, the supplier and customer have to interact several times before all details of the order have been agreed. We distinguish between an *ordering feature* and a facility to *request for proposals*.

The third feature is *entertainment*. The Web as a communication medium is similar to a catalogue and also to television, with animation, and even full motion video and audio. Some observers assume that the TV will be the dominant Web vehicle of the future. The recent purchase of Web TV Networks by Microsoft illustrates the growing interest in the integration of Web and television technology. Like TV-commercials, Web sites will provide infotainment. Companies should both provide information and simultaneously entertain their Web visitors. Hoffman and Novak [10] distinguish between two categories of behavior of consumers on the Web: goal-directed and experiential. Entertainment can be used to support the experiential flow, with intrinsic motivation, enduring involvement, and hedonic benefits. We measured whether a site contains a form of entertainment and, if available, which entertainment elements the site contains. Possible entertainment elements include jokes, cartoons, pictures, games, and videoclips.

A Web site can concentrate on one function, e.g., providing information, enabling transactions, or entertainment. The more available functions, the more advanced the site is. The extent to which a Web site is *advanced* has been measured as the number of different features the site contains. We distinguished between six different features. These consist of three information elements (background of the company, specific products, and non-commercial), two transaction elements (request for proposal and direct ordering) and one entertainment element (whether or not the site contained specific entertainment features). All content characteristics are features that can be measured objectively. For example, direct ordering is possible, or it is not.

The final characteristic of the content of a Web site is subjective. The *perception of the content* can be measured as the degree to which the Web site is considered to be informative. Web sites should provide value to customers, and the extent to which a Web site is informative determines, among others, whether this promise is fulfilled. Based on this criterion Hoffman et al. [11] distinguish between image sites and information sites. In image sites, the customer appeal is emotional rather than rational, and information about the product, if any, is provided by the context in which the product is consumed or has meaning to the customer. An informative site provides detailed and specific information with respect to products (prices, specifications, delivery terms), the company (mission, balance sheet, number of employees), or other relevant topics. Hymers [12] describes a study of IRB in the UK among Internet users; it seems to indicate that Internet consumers prefer to be given facts. We used a seven points Likert scale to measure the extent to which a Web site can be considered informative.

2.2. Design of web sites

A Web site can be considered to be a number of content elements (pages) that are linked to each other. The essence of a hypertext environment is that it allows for deep, nonlinear searches initiated and controlled by customers [11]. Web designers can influence the search by imposing a structure: by adding or omitting links between pages. The resulting Web site structure can take on several forms. Two extremes are hierarchical and network structures. In a hierarchical or tree structure, a page is linked to one or more pages at the next level and to only one page at the previous level. In a full network structure a page is linked to all other pages in the Web site; this enables visitors to navigate through the available information as they wish, but at the cost of complexity. However, it often does not make sense to link any page to any other. The only way to guide customers is by imposing a structure that specifies a limited number of paths. Thus, the best structure will be somewhere between a tree and a full network. In our study, four different *navigation structure types* have been distinguished: a tree, a tree with a

return-to-home page button, a tree with a few horizontal links, and an extensive network.

If a Web site becomes larger, the hyperlink navigation structure might become cumbersome. To prevent customers from much clicking and jumping from page to page, a *search function* can lead customers directly to requested information. We determined whether a Web site contains a search function, that is a function that can be used to find specific information within that site (not the entire Web).

From a marketing point of view Web surfers can be classified, e.g., as customers, potential customers, suppliers, and competitors. All may be interested in the information. However, it is highly unlikely that a company wants to provide the same information to each group. Certain information may be protected by means of a password. Only visitors with a valid password and/or IP address are allowed to request this information [20]. The most common application of *protected content* are pages for customers only; e.g., carriers, such as Federal Express and UPS, assign a unique number to each package and allow customers to check the delivery status. In our study, we determined whether a Web site contained pages protected by means of a password.

The framework also includes three measures of the perception of the design. The first measure is the *quality of the navigation structure* as perceived by Web visitors. Maroney [15] argues that hypertext provides the opportunity to build content that is intelligible and interesting to people at different levels of understanding. Unfortunately, hypertext may also be confusing. Therefore, we needed a subjective measure to determine the extent to which the structure is perceived as logical (determined by using a seven points Likert scale). This measure reflects the extent to which a structure actually helps a user to find information.

The next element refers to the multi-media capabilities of the Web. A Web site can be rather sober, with plain information and focus on the content providing function, on the other hand it may contain many elements that are primarily used to make the content more attractive (e.g., graphics and animation). The *image* of a Web site is not just a matter of presentation. A unique characteristic of the Web is that it is the customer who is paying for the communication, while the marketer is only paying for the design of the communication messages. A more attractive content takes more time to download and thus increases the cost to the customer. Non-text elements also enhance communication by helping visitors find or interpret the information presented. Designers must find a balance between an attractive design and providing information (plain content); this is not always easy. The image of the Web site is measured on a seven points Likert scale by stating the extent to which the site can be considered sober.

The final design aspect is *the presentation style*, which should stimulate reading (listening and viewing) and facilitate the interpretation of the information. To stimulate flow, it is important that a similar presentation style is used for all pages within the Web site. The presentation style includes the lay out, colors, font style and size, mix of text and graphical information, and sort, shape, size, and placement of links. We measured the extent to which the presentation style within the various pages can be considered as uniform (using a seven points Likert scale).

3. Research design

Building a Web site implies doing business on a global level. Marshall Industries, an electronic distributor in El Monte, CA, has received queries from 52 countries, and receives as much requests between 5 p.m. and 8 a.m., as they do in their California daylight hours [1]. A global medium assumes global customers. Mehta and Sivadas [18] found that global (non-USA) and domestic (USA) Web users have similar attitudes, similar results were found in other studies [8,21]. Mehta et al. [19] studied users of USENET newsgroups and concluded that such users form a global segment. They suggest that these findings allow marketers to overcome problems associated with entering foreign markets: the problem of mode of entry and standardization of the marketing mix across countries. Although the global nature of the Web is often stressed [12,22], there is some evidence that Web visitors from different countries behave differently. McBride [16] reports that a UK based car-hire company found that Americans were more confident in booking services over the Internet, while Europeans preferred to use the fax or telephone.

As a sample frame we used two directories: Yahoo (www.yahoo.com) and the Dutch Yellow Pages (www.markt.nl). The directory 'Business and Economy – Companies' of the Yahoo Guide provides one of the best known international lists of links to commercial Web sites; the majority of its companies are from the US. The Dutch Yellow Pages contains companies from the Netherlands only. Altogether, we analyzed 651 Web sites, 501 from Yahoo and 150 from the Dutch Yellow Pages.

To select companies from both directories, we applied a sampling procedure that was a combination of quota sampling and proportionate stratified sampling. The first was used because of the dominance of computer (related) firms on the Internet. A random sample would mostly contain IT-companies. We selected 25 categories (or quota's) from the Yahoo directory, see Table 1. If possible we selected a minimum of 15 sites per category (10 for Dutch Yellow Pages). The remaining sites were selected as a proportion of the size of each category. To select sites within a category, the stepsize s was computed as

$$s = \frac{C}{(n+1)}$$

where C is the category size, the number of sites in a category, and n is the number of sites from that category that must be included in the sample. Within a category each sth site was selected, until n was reached.

Not all selected Web sites could be analyzed. Sites with non-commercial content and empty sites (e.g., only a home page with the company name in black and white capitals) were excluded. Due to the dynamic nature of the Web, some listed links appeared to be invalid. May be companies had changed their URL address without its being updated in the directory, or they may have left cyberspace. If it was not possible to

Table 1
The Web sites from the directories Yahoo and Dutch Yellow Pages (DYP) that have been analyzed

Group	Category	Number in		Group	Category	Number in	
		Yahoo	Sample			DYP	Sample
1	computers	8353	108	1	computers	147	39
2	audio	147	15	2	books	26	10
2	books	576	15	2	broadcast stations	21	10
2	communication & media services	431	15	2	music	29	10
2	information	211	15	2	newspapers & magazines	33	11
2	media	982	20	2	telecom	6	6
2	music	1299	23	3	financial business	20	10
2	telecommunication	654	16	4	consultancy	67	19
3	financial services	1426	25	4	PR, advertising, DM	40	13
3	insurance	218	15	4	travel	20	10
4	advertising	370	15	5	clothing	2	2
4	consulting	397	15	5	industry	18	10
4	corporate services	256	15				
4	travel	1061	21				
5	apparel	361	15				
5	automotive	870	18				
5	breweries	87	15				
5	electronics	341	15				
5	games	104	15				
5	gifts	410	15				
5	home & garden	559	15				
5	industrial suppliers	354	15				
5	jewelry	150	15				
5	photography	386	15				
5	toys	39	15				
	total	20042	501		total	429	150

analyze a particular Web site, the next site from the same category was selected.

Web sites were analyzed by a research team consisting of graduate students of the University of Groningen. After testing the questionnaire, they were told how to interpret and answer each question. A questionnaire was completed by two researchers working together. This procedure was used to decrease the chance of typographical errors and because several questions requested a subjective estimate. To increase the validity of the scores, both researchers had to agree upon these subjective estimates. The data were collected in the first quarter of 1996.

4. Categories of Web sites

4.1. Source

Five hundred and one Web sites were collected from the Yahoo directory and 150 from the Dutch Yellow Pages (DYP). Although several large corporations from the Netherlands are also listed in the Yahoo directory, there appeared to be no overlap between both samples.

4.2. Industry

Probably not all industries benefit from the Internet equally. For example, companies producing or selling information-based products may benefit more than manufactured products. We have used the categories from both directories to form five groups of industries. Table 2 provides an overview of the number of Web sites in each of the five industries. The largest group is 'Products', which consists of categories such as games, gifts, home & garden, and electronics. 'Computers' is the only single category industry, in both directories this is the by far largest category of companies.

Table 2
The distribution of the Web sites over five industries, based on the categories available in the directories Yahoo and DYP

Group	Industry	Number	%
1	computers	147	22.6
2	information	166	25.5
3	finance & insurance	50	7.7
4	services	108	16.6
5	products	180	27.6
	total	651	100.0

Table 3
The size of the Web site (measured as the number of pages)

No. of pages in Web site	Number	%
1 page	43	6.6
2–10	169	26.0
10–25	198	30.4
25–50	126	19.4
50–100	58	8.9
>100 pages	57	8.8
Total	651	100.0

4.3. Size

The size of the Web site has been measured by its number of pages. Because this number is hard to assess for very large Web sites, we have split the size into six categories, ranging from one page to over 100. Table 3 shows the distribution of the 651 Web sites. Most Web sites, over 75%, contain between 2 and 50, the median value is the category 10–25 pages.

The average size of the Yahoo and DYP sites are similar, with means of 3.3 (Yahoo) and 3.2 (DYP). The average size per Web site is not equal for all five industries. The largest sites are found in the industries Information and Computers, the smallest in Services and Finance & Insurance. The Kruskal–Wallis test confirmed that the observed differences are highly significant ($p < 0.000$).

5. Results: content

5.1. Information

The kind of information that is most often available in a Web site is specific product information. Of all analyzed Web sites, 81% contain this information (see Table 4). Information about the background of the company is generally available, but only a small number of the sites contain non-commercial information. When both sources, Yahoo and DYP, are compared only one significant difference is found. The

Table 4
Content features of the Web site versus the source of the site

Content features	Yahoo		Dutch Yellow Pages		χ^2-test		Total	
	N	%	*N*	%	χ^2	*p*	*N*	%
Informatio								
1. Background company	323	64.5	118	78.7	10.6	0.001	441	67.7
2. Specific products	411	82.0	115	77.2	1.8	0.185	526	80.9
3. Non-commercial	55	11.0	13	8.7	0.66	0.417	68	10.4
Transaction:								
4. Request for proposals	30	6.0	9	6.0	0.00	0.996	39	6.0
5. Ordering feature	188	37.9	42	28.0	4.9	0.026	230	35.6
Entertainment								
6. Entertainment available	148	29.5	49	32.7	0.53	0.465	197	30.3
Advanced site								
No. of features (mean)		2.30		2.32	*t*-test	0.907		2.31

DYP sites contain more often information about the background of the company.

There were several differences across industries (see Table 5). Information about the background of the company is most often found in the Services industry, and is least available in the Products industry. Specific product information is available in almost all Computers sites. Finally, non-commercial information is relatively often found in the Information industry but only sporadically in the Computers industry.

The size of the site is significantly related to both specific product information and non-commercial information (Mann–Whitney test, twice $p = 0.00$). Larger sites contain more specific product and non-commercial information.

5.2. Transaction

More than one-third of all sites contain an ordering feature, but only a small minority offer the facility to request for proposals. The ordering feature is significantly more often found in sites from Yahoo than DYP ($p = 0.026$). It is also quite often found in the Products industry, which has by far the highest score for

Table 5
Content features of the Web site versus the industry

Content features:	Computers		Information		Finance & Insurance		Services		Products		χ^2-test	
	N	%	*N*	%	*N*	%	*N*	%	*N*	%	χ^2	*p*
Information												
1. Background company	105	71.4	107	64.5	36	72.0	87	80.6	106	58.9	16.7	0.002
2. Specific products	134	91.8	126	75.9	38	76.0	82	75.9	146	81.1	16.4	0.003
3. Non-commercial	7	4.8	27	16.3	4	8.0	10	9.3	20	11.1	11.7	0.020
Transaction												
4. Request for proposals	10	6.8	5	3.0	5	10.0	12	11.1	7	3.9	10.7	0.031
5. Ordering feature	52	36.1	65	39.4	11	22.0	19	17.6	83	46.4	29.4	0.000
Entertainment												
6. Entertainment available	48	32.7	52	31.3	14	28.0	24	22.2	59	32.8	4.5	0.348
Advanced site												
No. of features (mean)		2.44		2.29		2.16		2.17		2.34	ANOVA	$p = 0.21$

ordering feature. The Services industry ranks first on request for proposals. Finally, the larger the site the more often it contains an ordering feature (Mann–Whitney test, $p = 0.00$).

5.3. Entertainment

About one-third of all sites contain some form of entertainment. Most popular were pictures, while videoclips were least available. The average number of different forms of entertainment is 0.44, with no more than 10% of sites offering two or more forms of entertainment. With respect to the presence of any form of entertainment, there are no significant differences between Yahoo and DYP sites. Also, the average number of different forms of entertainment is almost equal (Yahoo 0.44, DYP 0.49; t-test, $p = 0.52$).

With respect to industry, there were some differences but they were not significant. The average number of different forms of entertainment were also not significantly different.

Finally, there is a correlation between the size of a Web site and the presence of entertainment. Larger sites tend to contain more often (Mann–Whitney test, $p = 0.003$) and more different forms of entertainment (Spearman correlation 0.13, $p = 001$).

5.4. Advanced site

The average number of content features is 2.3, almost two-third of all sites contain two or less features (see Fig. 2). Only three sites (0.5%) contain all six content features. Both sources, Yahoo and DYP, score almost equal.

The Computers industry is the industry with the highest average number of features, while the industries Finance & Insurance and Services score lowest. However, these differences are not significant.

The size of the Web site correlates positively with the number of content features (Spearman correlation $= 0.18$, $p = 0.000$).

5.5. Perception of content

The DYP sites are considered to be more informative than the Yahoo sites. This difference is significant (t-test, $p = 0.08$). Products industry sites are the least informative ones, while the scores for the other indus-

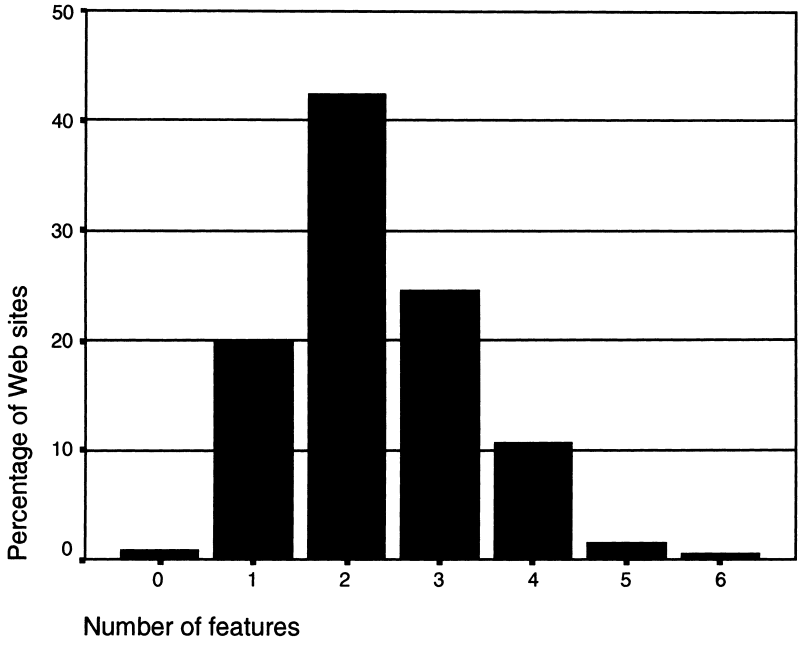

Fig. 2. The distribution of the extent to which the Web sites are advanced, measured as the available number of different content features.

Table 6
The structure of the Web site versus the source

Structure of the Web site	Yahoo		Dutch Yellow Pages		Total	
	N	%	*N*	%	*N*	%
Tree structure	104	21.5	37	25.2	141	22.3
+Back to Home Page	189	39.0	56	38.1	245	38.8
+Few horizontal links	109	22.5	30	20.4	139	22.0
Network structure	82	16.9	24	16.3	106	16.8

tries are almost equal. Due to the much lower score for Products, the analysis of variance showed a significant result ($p = 0.07$). Finally, we have found a positive and highly significant correlation between the size of the site and the degree to which the site is considered to be informative (Spearman correlation 0.30, $p < 0.000$). Larger sites are considered as more informative.

6. Results: design

6.1. Navigation structure

Overall, we found that the majority of Web sites have a quite simple structure, over 60% of the Web sites have a tree structure or a tree structure supplemented with a back to home page button (see Table 6). No more than 17% have an extensive network structure. There is no significant difference between both sources, Yahoo and DYP (Mann–Whitney test, $p = 0.43$).

Table 7 shows the distribution of the four different structures per industry. There is, on average, no large difference between the five industries (Kruskal–Wallis test, $p = 0.34$). The category 'tree structure with a return-to-home page button' is the largest, with declining percentages for the other categories. There is one exception: within Finance & Insurance the category 'network structure' is larger than 'tree structure with a few horizontal links'. Compared with the other industries, Finance & Insurance contain more sites with a network structure, while Services and Products have the least Web sites with such a complex structure. Overall, sites from the Information industry have a more complex structure.

There is a strong correlation between the size and the structure of the Web site. More complex structures are found in larger Web sites, while smaller sites have (and need) less complex structures (Spearman correlation 0.34, $p = 0.000$).

6.2. Search function

Only a small minority of the Web sites contained a search function. The Yahoo and DYP sites are almost equal with a χ^2-test of $p = 0.41$. The breakdown by industry uncovered highly significant differences. A relatively large percentage of sites from the Information industry contains a search function, whereas Services, Finance & Insurance, and Products score much lower. Computers scores almost equal to the overall average. These differences are highly significant (χ^2-test, $p < 0.000$). Finally, there is a strong correlation between the size of the Web site and whether or not a search function is available (Spearman correlation 0.33, $p < 0.000$).

6.3. Protected content

Only a small minority of the Web sites contained pages that are only accessible by means of a password

Table 7
The structure of the Web site versus the industry

Structure of the Web site	Computers		Information		Finance & Insurance		Services		Products	
	N	%	*N*	%	*N*	%	*N*	%	*N*	%
Tree structure	26	18.3	31	19.1	15	30.0	26	25.2	43	24.7
+ Back to Home Page	61	43.0	60	37.0	16	32.0	44	42.7	64	36.8
+ Few horizontal links	29	20.4	40	24.7	8	16.0	18	17.5	44	25.3
Network structure	26	18.3	31	19.1	11	22.0	15	14.6	23	13.2
Average structure	2.39		2.44		2.30		2.21		2.27	

(2.8%). For Yahoo this percentage is higher than for DYP, but this difference is not significant (χ^2-test, $p = 0.51$). Compared per industry, we have found significant differences. Content protection is most often found in Finance & Insurance and Computers sites. On the other hand none of the 108 sites from the Services industry contained protected information. Overall, these differences are significant (χ^2, $p = 0.05$). With respect to the third characteristic of a Web site, the size, we have found a significant difference: larger sites tend to be more likely to have pages that can only be requested by means of a password (Mann–Whitney test, $p = 0.03$).

6.4. Quality of the structure

To assess ease of navigation through the Web site we used a 7-points Likert scale reflecting the extent to which a Web site contains a logical structure. The overall mean is 5.12. On average, the structure of Yahoo sites is perceived to be somewhat better than the DYP sites, but this difference is not significant (t-test, $p = 0.17$).

By industry, there were remarkable differences. The Computers industry received the lowest average score on perceived quality of structure. Information, the industry with the most complex sites, and Products, with much less complex sites, have the highest scores. The differences between the various industries are significant (analysis of variance, $p = 0.02$). Finally, the size of the Web site was negatively correlated with the perceived quality of the navigation structure (Spearman correlation -0.12, $p = 0.002$).

6.5. Image

The extent to which a Web site is perceived as sober was measured using a 7-points Likert scale. On average, the Yahoo and DYP sites were judged to be equally sober (t-test, $p = 0.26$). By industry, the differences are much larger. The least sober sites are found in the Information industry, followed by Computers and Products. The differences per industry are significant (ANOVA, $p = 0.003$). Finally, the more pages a site contains, the less sober it is. We found a strong and negative correlation with the size of the site (Spearman correlation -0.33, $p < 0.000$).

6.6. Presentation style

On a 7-points scale the presentation style of most sites was considered to be quite uniform. Between both sources there are only minor differences, and also by industry the differences are not large. Finally, it turned out that the larger the Web site, the less uniform its presentation style (Spearman correlation -0.11, $p = 0.005$).

7. Summary

7.1. Source of the Web site

The source of the Web site (Yahoo versus Dutch Yellow Pages) was used to enlarge our understanding how the new global medium World Wide Web is applied in various countries. Overall, we have found only minor differences. With respect to content, the results are mixed, but for none of the six design characteristics we found significant differences. Thus, there is no evidence that the new medium the World Wide Web is used differently by companies that have registered themselves in Yahoo versus the Dutch Yellow Pages.

7.2. Industry

The overall differences per industry were significant for six of the eight aspects of content (the exceptions are entertainment and advanced site). Sites from the Computers industry contain more specific product information but less non-commercial information. In Information industry sites, non-commercial information is found most often, but mixed results have been found for transaction features: the lowest score for request for proposals but a high score for direct ordering. Finally, the industries Services and Products show remarkable results: for most content characteristics their extreme scores are each others opposite.

With respect to design, the overall differences per industry were significant for four of the six aspects of design (search function, protected content, perceived quality of the structure, and image). The Computers industry has more protected content and less sober sites, but received, on average, the lowest score on the

perceived quality of the structure. The Information industry scores, on average, very well on design issues. Sites from the Services industry have a relatively simple navigation structure, hardly ever contain a search function, none of them contain protected content and are considered to be the most sober sites.

7.3. Size

The third background characteristic is the number of pages. The size of the site was found to correlate with almost all content and design aspects (six out of eight content characteristics and all six design aspects). In summary, larger Web sites do not only contain more content but also more different content features (especially information about specific products, non-commercial information, direct ordering features, and entertainment), more different forms of entertainment and are considered as more informative. Larger Web sites have a more complex navigation structure and have more search functions and protected content. Nevertheless, we found a negative correlation with the perceived quality of the navigation structure. The more pages a site contains, the less sober it is, and the less uniform its presentation style.

8. Conclusions

The goal of this paper was to enlarge our understanding of how different groups of companies are using the Web for commercial purposes. Because a solid knowledge base for analyzing Web activity is missing, we developed a research framework to analyze and categorize the capabilities of Web sites. This framework distinguishes between content and design. We operationalized content and design by means of both objective and subjective measures. The results of our study show that this distinction is meaningful, which stresses the importance of including user perceptions when studying Web sites.

In our study we used three different classifications of Web sites. The first, the source of the Web site (Yahoo versus The Dutch Yellow Pages), showed hardly any differences. The second, based on industry, revealed more differences. The third was based upon the size of the site in terms of number of pages. It turned out that size was the most powerful descriptor of both content and design. Almost all content and design aspects were significantly related with the size of the site. On the Web 'larger' does not only mean 'more of the same', larger sites are also 'richer' sites.

Our study was limited in several ways. For example, we have analyzed the sites at one point in time while the Web is a highly dynamic medium. Similar studies at different times are likely to show different results. A second limitation concerns our subjective measures. When the researchers analyzed a Web site they almost always visited that particular site for the first time. It is possible that a site is aimed at regular customers and that subjective measures provided by repeat visitors would be different from those provided by first-time visitors. Thirdly, we compared two different directories to try to study differences between different countries. Methodologically, it is better to determine the country of origin before including a site in the sample. Finally, we have analyzed commercial Web sites assuming that the companies behind them had commercial goals in mind. Literature, however, suggests that many companies entered the Web in quite a rush, without clear goals and merely for experimental purposes.

While our study was a descriptive one, the Web also calls for normative studies. For example, it would be interesting to determine what kind of design is, according to the perception and behavior of visitors, optimal for an industry and the goals of the company. Further research could also concentrate on the development of a general framework to study Web sites. The academic community needs such a general framework to facilitate the comparison of the results of Web studies with different goals, among different categories of sites or user groups, or conducted at different times. The use of similar frameworks by different researchers can speed up the development of a solid Web knowledge base. Managers in practice are very eager to make use of such a knowledge base.

Acknowledgements

The author would like to thank the editor and three anonymous reviewers for their constructive comments and their insightful suggestions.

References

[1] D.A. Andelman, Betting on the Net, Sales & Marketing Management 147(6), 1995, pp. 47–60.

[2] A. Angehrn, Designing mature internet business strategies: the ICDT model, European Management Journal 15(4), 1997, pp. 361–369.

[3] P. Berthon, L.F. Pitt, R.T. Watson, The world wide web as an advertising medium, Journal of Advertising Research, January/February 1996, 43–54.

[4] L. Chase, MNE marketing in the world wide web of chaos, http://chaseonline.com/marketing/golden.html. (1996).

[5] M.J. Cronin, in: Doing more business on the Internet, Van Nostrand Reinhold, New York, NY, 1995.

[6] Giga Information Group, Global Annual Cost Savings From Electronic Commerce, August, www.gigaweb.com, 1999.

[7] C. Grönroos, From marketing mix to relationship marketing: towards a paradigm shift in marketing, Management Decision 32(2), 1994, pp. 4–20.

[8] S. Gupta, Consumer Survey of WWW Users, www.umich.edu/∼sgupta/hermes.html, 1995

[9] D.L. Hoffman, T.P. Novak, Marketing In Hypermedia Computer – Mediated Environments: Conceptual Foundations, Revised 11 July (www2000.ogsm.vanderbilt.edu/cmepaper.revision.july11.1995/cmepaper.html) (1995).

[10] D.L. Hoffman, T.P. Novak, Marketing in hypermedia computer – mediated environments: conceptual foundations, Journal of Marketing 60, 1996, pp. 50–68.

[11] D.L. Hoffman, T.P. Novak, P. Chatterjee, Commercial scenarios for the web: opportunities and challenges, Journal of Computer-Mediated Communication, Special Issue on Electronic Commerce, vol. 1(3) (1995).

[12] J. Hymers, Integrating the internet into marketing strategy, Journal of Targeting Measurement and Analysis for Marketing 4(4), 1996, pp. 363–371.

[13] F. Krijnen, Adverteerders moeten nog veel leren op de elektronische snelweg, Adformatie 36, 1997, pp. 48–50.

[14] Krol, E.E. Hoffman, FYI on What is the Internet? Network Working Group Request for Comments 1462 – FYI 20. (URL: gopher://ds1.internic.net/00/fyi/fyi20.txt) (1993).

[15] D. Maroney, In praise of hypertext, Journal of Advertising Research, March–April, 37(7–9) (1997).

[16] N. McBride, Business use of the internet: strategic decision or another bandwagon? European Management Journal 15(1), 1997, pp. 58–67.

[17] V. McCarthy, The web: open for business, Datamation 1, 1995, pp. 30–36.

[18] R. Mehta, E. Sivadas, Direct marketing on the internet: an empirical assessment of consumer attitudes, Journal of Direct Marketing 9(3), 1995, pp. 21–32.

[19] R. Mehta, R. Grewal, E. Sivadas, International direct marketing on the internet: do internet users form a global segment? Journal of Direct Marketing 10(1), 1996, pp. 45–58.

[20] N.J. Muller, Expanding the help desk through the world wide web, Information Systems Management, 13, 1996, pp. 37–44.

[21] J. Pitkow, M. Recker, Consumer Survey of WWW Users, www.cc.gatech.edu/gvu/user_surveys (1995).

[22] J.A. Quelch, L.R. Klein, The internet and international marketing, Sloan Management Review, 37, 1996, 60–75.

[23] J.E. Rayport, J.J. Sviokla, Exploiting the virtual value chain, Harvard Business Review, November–December, 1995, 75–85.

[24] F.F. Reichheld, in: The Loyalty Effect The Hidden Force Behind Growth, Profits and Lasting Value, Harvard Business School Press, Boston, 1996.

[25] R.B. Woodruff, Customer value: the next source of competitive advantage, Journal of the Academy of Marketing Science 25(2), 1997, pp. 139–153.

Eelko K.R.E. Huizingh is an Associate Professor of Marketing and Information at the Faculty of Economics of the University of Groningen, The Netherlands. His research focuses on marketing applications of information and communication technology. More specifically, he is involved in studies on e-commerce, Internet marketing, interactive marketing, the quality of customer databases, and marketing decision support systems. Dr. Huizingh has published papers in many journals in The Netherlands, and is (co-)author of various books on the use of ICT for marketing purposes. His papers have appeared in journals such as Decision Support Systems, Socio Economic Planning Sciences, Organizational Behavior and Human Decision Processes, the Journal of Market-Focused Management, Estudios de Administración, and the International Journal of Computer and Engineering Management. He holds a Ph.D. from the University of Groningen, The Netherlands.