

# **A step-by-step approach to web classification design**

**Learn how you can effectively design a robust,  
reader-friendly web classification**

**October 2002**

**By Gerry McGovern**

**[www.gerrymcgovern.com](http://www.gerrymcgovern.com)**

**[info@gerrymcgovern.com](mailto:info@gerrymcgovern.com)**

**Copyright © 2002 Gerry McGovern**

# A step-by-step approach to web classification design

## Table of Contents

1.	Principles of classification design .....	3
1.1.	Keep your reader in mind .....	3
1.2.	Keep your authors in mind .....	3
1.3.	Keep your organization's objectives in mind .....	3
1.4.	Avoid duplication .....	4
1.5.	Be prepared to make editorial decisions .....	4
1.6.	Seek professional help .....	4
1.7.	Design classification like it will be 'written in stone' .....	5
1.8.	Design for the total content environment.....	5
1.9.	Design for a number of classification system .....	6
1.10.	Design a practical classification system .....	6
1.11.	Design from the top-level down.....	7
1.12.	Make sure the editorial board is involved .....	7
1.13.	Take your time.....	7
2.	The classification design process .....	8
3.	Classification situation analysis .....	8
3.1.	Organization objectives and strategy analysis .....	9
3.2.	Current content analysis.....	9
3.3.	Common search words analysis.....	10
3.4.	Author and specialist survey .....	10
3.5.	Competitor content analysis .....	10
3.6.	Reader survey .....	10
3.7.	Industry publications analysis.....	10
3.8.	General industry analysis .....	10
3.9.	Initial list verification .....	11
4.	Deciding on classification options .....	11
4.1.	Subject classification .....	11
4.2.	Task-oriented classification .....	11
4.3.	Geographic classification .....	12
4.4.	Date classification .....	12
5.	Classification design techniques.....	12
5.1.	Workshops .....	13
5.2.	Flip board and post-it note approach.....	13
5.3.	Card sorting approach.....	13
5.4.	Prototyping .....	14
5.5.	Individual classification design checklist.....	14
6.	About the Author .....	15

# 1. Principles of classification design

Classification design is one of the most difficult challenges for large websites. The following are key principles you should consider:

1. Keep your reader in mind
2. Keep your authors in mind
3. Keep your organization's objectives in mind
4. Avoid duplication
5. Be prepared to make editorial decisions
6. Design classification like it will be 'written in stone'
7. Design for the total content environment
8. Design a practical classification system
9. Design for a number of classification options
10. Design from the top-level down
11. Make sure the editorial board is involved
12. Take your time

## 1.1. Keep your reader in mind

When approaching classification design, always keep foremost in your mind the reader/visitor. At every step of the classification design process, you must ask yourself the following question: Is this the classification that my readers will understand?

## 1.2. Keep your authors in mind

Keeping the reader in mind is fundamentally important. So, too, is keeping the author in mind. Remember, someone—the author or someone contributing for the author—will have to classify content on a day-to-day basis.

The more complex a classification system you choose, the more skill and time it will require for someone to classify content under that system. If authors or contributors don't understand the classification, then they're not going to classify correctly. If they don't classify correctly, then the whole process falls down.

## 1.3. Keep your organization's objectives in mind

While the reader and author are important to keep mind, so too is the organization. Classification is to strategy, as poetry is to prose. It is the ultimate, concise expression in words of the organization's strategy and focus.

When a reader visits your website, the classification will be among the first things they read. These classifications will create that vital first impression for the reader of what you do.

'Products', 'Services', 'Solutions'. Which would you choose? 'Products for large corporations,' 'Products for small and medium enterprises,' 'Products for home users.' Which would you choose? 'Germany', 'United States', 'Ireland'. Which would you choose?

It's clear from the above examples how the choice of classification can instantly communicate to the reader what the organization is about. Thus, it is essential that the classifications chosen align with the strategic focus and objectives of the organization.

#### **1.4. Avoid duplication**

When people get going with classification design, sometimes it's hard to get them to stop. All sorts of classification terms appear. That's great. Keep them rolling in!

However, the Achilles Heel of classification is two classifications that on the surface seem different, but in reality are the same. This results in:

- Scratching of the head by the reader. They come to the website and wonder if what they're looking for is under 'Services' or 'Solutions?' If they click on the wrong classification, they get frustrated.
- Scratching of the head by the author. When classifying their document, they wonder whether they should classify it under Services or Solutions? If they classify under the wrong classification, the system is undermined.

#### **1.5. Be prepared to make editorial decisions**

At the end of the day, much classification design comes down to opinion. Somebody, somewhere will unquestionably complain. What you're looking for is a happy medium—or the least unhappy medium!

That's life! That's the way it is. You can go round and round the garden with classification. Sooner or later, someone has to make an editorial decision. Based, hopefully, on experience, research, wisdom, skill and good gut instinct, you will have to choose what is the right classification. Editors make these types of decisions all the time. Ultimately, you have to trust your judgement and go with the decision you think is right.

What's vital here is that somebody needs to be in charge. It's great having workshops and lots of feedback. However, at the end of the day, some difficult decisions will have to be made. Somebody needs to have the authority to make these decisions.

Of course, ultimately, the final decision will lie with the editorial board. However, they will be asked to make relatively high-level decisions. Somebody has to have the authority to move the whole process forward before it gets to the editorial board. Ideally, this person should be a member of the editorial board.

#### **1.6. Seek professional help**

If you are classifying a large quantity of content, professional help can make your life a whole lot easier. Librarians and other information professionals have been trained in classification. They have experience and skill.

Too often, those who design for the Web have had this crazy attitude that what they're doing is truly unique. That nobody has faced the problems they face. This is an unprofessional attitude that flies in the face of reality.

Librarians have ate, slept and drank classification issues for most of their careers. Editors face information architecture-type problems every day. They have managed the publication of magazines, books, reports, which involved structuring chapters, creating table of contents, indexes, etc.

Ideally, if you're in charge of a major classification project, you should have a background in publishing or information science. If not, try to get some support from such professionals. It will be of great help to you.

### **1.7. Design classification like it will be 'written in stone'**

Classification design is not a 'suck it and see' design process. Unfortunately, it is often treated like that. Some people think that because a website is so malleable, it can be changed frequently. Yes, the content can and should be changed frequently, but not the classification, not the information architecture.

In information architecture design you are planning an information space that people will want to navigate around. Classification design is a fundamental pillar of this design. It should be solid and long-lasting.

The reason classification should be designed as if it is being written in stone is because people will get used to finding their way around your website using the classification. Authors will get used to contributing content and classifying that content using the classification system. If you keep changing your classification you will keep confusing people. That is not a good thing to do.

Think about how you move about the place you live in. You get used to where the streets, shops and other buildings are. Now, imagine how you would feel if a whole street was removed, or two new streets were added. You wouldn't feel very happy, would you? Most likely, you'd feel disoriented and lost.

Well, that's exactly the feeling your regular readers will get if they find you changing your information architecture. For example, some time ago, Ebay changed the background color of their website. They changed nothing else; just the color. There was uproar. Ebay was forced to change back to the original color.

People do not like information architecture to change. Spend the time to get it as right as possible first time. Think about how it will evolve in the future and plan for that. Otherwise, you are storing up trouble.

### **1.8. Design for the total content environment**

One way of planning for the future is to design for the total content environment, not simply the content you will have at launch, or six months after launch. Spend the time to think through your classification in a thorough manner. Examine how your content will evolve; how new types of content will appear.

The fact that you design a comprehensive classification does not have to mean you will have lots of empty pages on your website. You can specify the content management

software to only publish a classification on the website if there is content published under that classification. In this way, you have a master classification, but you only publish classifications that have content published under them.

Think like an architect. You are planning a skyscraper. You believe that within three years you will be able to rent out 100 stories. However, in the first year you will only rent out 40 stories. Do you build a 40 story skyscraper and then add on the other 60 stories in year 2? Not likely.

So, don't go architecting your classification piecemeal. Don't be adding little bits to it every couple of months that you forgot to plan for. That is simply bad design. It will create spaghetti classification that is a curse to navigate around.

As information scientist, Marcia J. Bates, states: "If you believe your information resource will grow, then design for growth from the beginning. Otherwise, trust me: It will get worse."

### **1.9. Design for a number of classification options**

It's important to have a number of classification options because this will allow for a variety of navigation options for the reader. Numerous studies have shown that readers want a variety of options when navigating through content on your website.

A classification option can be defined as a master grouping for content. Classification options include:

- Subject
- Task
- Geographic
- Date

### **1.10. Design a practical classification system**

Classification is a science and sometimes it is a religion. Those trained in library science often get fraught when they see how the rules are bent and sometimes broken in web-based classification design.

A classification expert should not be a purist but rather a realist. Sometimes a classification that is really a level two will need to be placed at the top-level (level one) because it's a really popular classification for the reader, and/or a strategically important one for the organization. Rules are great, but you need to be able to accommodate exceptions.

Keep the following in mind:

- A three-level classification is what to aim for. A five-level classification is the maximum. Once you go below three levels, people start getting confused and impatient.
- Ideally, you should have no more than 10 classifications at any particular level. 15 should be your maximum. Some people think that you should have less than 10; that people can't cope with a lot of choice. Yahoo has fourteen classifications

- in its core navigation/classification. Yahoo is very popular, so, obviously, people have got used to choosing from a relatively long list of classifications.
- The classic order for a classification is alphabetical. However, practically, it may be that certain classifications are much more important to the organization and/or reader than others. This will necessitate prioritizing this in the list order.

### **1.11. Design from the top-level down**

Don't try and design the entire classification at one go. The top-level—the classifications the reader will see when they arrive at your homepage—should always be designed first. They should receive the major part of your effort.

You should get them right first before paying too much attention to how the lower levels will be organized. This is really important to understand. You can have all sorts of debates about what should be at lower level. However, if your top-level classifications have not been thoroughly thought through, and then need to be changed, that will have all sorts of implications for the lower levels.

Only after you have got sign-off from your editorial board for the top-level classification should you concentrate on the lower level design.

### **1.12. Make sure the editorial board is involved**

Some people think that classification design is a boring, rather inconsequential activity. They couldn't be more wrong. When someone comes to your website one of the first things they will look at is your classification. If they don't see the classifications they want, they will likely leave.

From an internal point of view, managers and other staff can get very tetchy if their favourite product, department, etc., is not represented at the top-level. Therefore, the top-level needs to be thoroughly discussed and signed off by the editorial board. (This board should have representatives from all the key sections within your organization.) Until conclusive agreement has been reached, it is rather pointless taking the design process further.

### **1.13. Take your time**

You should never rush classification design. It's hard work. It's complex work. It's vital work. It takes time to tease out all the implications of choosing a particular classification. If you get it right, you have established a strong pillar in your information architecture. If you get it wrong, it will cause you untold problems down the line.

## 2. The classification design process

The following process is proposed for classification design:

1. Carry out the classification situation analysis. (4. Classification situation analysis.)
2. Develop an initial classification list and verify this for obvious errors and duplications. It can be useful at this stage to place this list in a public area within the organization so as to get more feedback.
3. Initiate classification workshops where you will:
  - a. Decide on style and tone issues. Will the style be concise, using classifications like 'Activities.' Or, will it be more warm, using classifications like, 'Things to Do and See.' Also, how will the capitalization work? Will you capitalize all words in the classification, just the first word, etc.? Will you use ampersand (&) instead of "and" to save space?
  - b. Agree the classification options. (See 5. Deciding on classification options.)
  - c. Agree the number of levels for the classification. Ideally, you should have no more than three levels, and a maximum of five.
  - d. Agree the range of classifications at any particular level. You shouldn't have more than 15 classifications at any one level.
  - e. Initiate a flip-board and/or card sorting approach to choosing potential classifications. (See 3. Classification design techniques.)
  - f. Test each potential classification using the individual classification design checklist. (See 3. Classification design techniques.)
  - g. Agree the first draft of the top-level classification
  - h. Agree the order of the classifications for this top-level
4. Develop a paper-based prototype of this classification.
5. Test this prototype with readers, authors and other interested parties.
6. Revise this prototype. Test again, if required.
7. Prepare a final prototype.
8. Present to the editorial board.
9. Achieve sign-off of top-level classification
10. Design lower-level classifications. It may well be that at this stage you break up into specialized teams. For example, if Finance was chosen as a top-level for your intranet, a team from Finance might work on the lower level classifications for the Finance classifications. If this happens, make sure they are given guidance on best practice for classification design.

Let's now examine in greater detail:

- How to go about the classification situation analysis
- How to decide on the classification options
- Techniques that will help you carry out the classification design process more efficiently

## 3. Classification situation analysis

Classification design should follow the 'geniuses steal, beggars borrow' rule. Your job is not to come up with some innovative way to classify your content. It is to find a classification that works. You will know that a classification is working if you're readers

are happy using it. So, if you find elements of your classification—either internally or externally—that you think will work for you, adapt them to your needs.

The classification situation analysis should cover the following areas:

1. Organization objectives and strategy analysis
2. Current content analysis
3. Common search words analysis
4. Author and specialist survey
5. Competitor content analysis
6. Reader survey
7. Industry publications analysis
8. General industry analysis

As you go about your situation analysis:

- Underline or otherwise mark potential classification terms
- Type these terms into a spreadsheet document
- Beside each classification list the source of where you got it from
- Number the amount of times you come across a particular classification term. This will be a good gauge for its popularity.
- Don't worry about duplication at this stage. List everything that you think makes sense.
- Don't be too analytical at this stage. Type down whatever you think might be relevant classifications.
- Remember, to focus on the top-level classifications. You can collect possible lower level classifications in a separate document, but don't spend too much time on it.

### **3.1. Organization objectives and strategy analysis**

Classification design should always keep the reader foremost in mind. However, if you don't have a strong grasp of what your objectives are, then you can't properly identify your reader. Nor can you clearly identify what is the benefit to the organization of classifying all this content.

Examine your corporate objectives, mission statement and other high-level documentation and see what words or terms stand out. What area of business do you want to be leader in? Do you want to be seen as the best in the industry at customer support? Your objectives should throw up key classification terms.

Perhaps your corporate objectives are too vague from a content classification perspective. If so, work with the editorial board to develop clear content-focused objectives. From these, draw key classification terms.

### **3.2. Current content analysis**

- What classification systems are currently in use within the organization?
- What classification system is being used for the current website, if any?
- Are there any classification systems being used for other non-web databases?
- Are there any paper-based classification systems?

- Analyze a sample of content, isolating potential classification terms.

### **3.3. Common search words analysis**

- If you have a current website, analyse the most common words or terms that people use when searching the website. These words or terms can often give you hints in relation to what the classifications should be.
- Collect these most popular search words or terms.

### **3.4. Author and specialist survey**

- Talk to the people who create the content. How would they like to see their content classified?
- Sometimes, the authors of the content may be too busy to contribute their content. Talk to their assistants.
- Do you have librarians or other information specialists on staff? How would they like to see the top-level classification appear?

### **3.5. Competitor content analysis**

- How are competitors classifying their content?
- Really focus on competitors who have popular websites.
- Are there any common trends emerging towards classification among competitors?

### **3.6. Reader survey**

- Survey your readers. Ask them to list their most important top-level classifications.
- Ask the readers what they want to do when they come to your website and try and catch that as classification terms.
- Make sure that you get representative samples from each reader type.

### **3.7. Industry publications analysis**

- Examine popular industry publications.
- How do they go about structuring their publications?
- What sort of terms are coming up again and again?

### **3.8. General industry analysis**

- Are there any industry groups involved in developing a common metadata/classification approach?
- Are there accepted classification systems for the industry?

### **3.9. Initial list verification**

At the end of the situation analysis, you will have developed an extensive list of potential classifications. You now need to go through an exercise which eliminates obvious duplication and corrects spelling. When you are eliminating duplication, note beside the remaining classification the number of times it has occurred. This is an important indicator of how popular it is.

At the end of this process, you should have a long list of potential classifications stored in a spreadsheet format. These classifications can now be translated into post-its for the flip-board design approach. Alternatively, they can be written out on cards for the card sorting design approach.

## **4. Deciding on classification options**

The following are potential classification options:

- Subject
- Task
- Geographic
- Date

Remember, the beauty of web-based classification is that you can multiple-classify. This allows the reader a variety of navigation options so that they can find the content they need just the way they want to find it.

However, keep in mind here that the more classification options you choose for your content, the more effort will be required for the person who is supposed to classify a particular piece of content.

### **4.1. Subject classification**

This is a classic approach to classification. Here, you organize your content based on defined subjects. Perhaps the best known subject classification is the Yahoo classification. If you go into a library, you will also find subject classification at work. Over here you will find the History section. Over there you will find the Music section.

In most situations, subject classification should become your core classification; the essential classification architecture around which you organize your content.

### **4.2. Task-oriented classification**

While subject and functional classification is an efficient way for an organization to organize itself, it is not always what the reader wants. Very often, the reader comes to your website because they want to do something.

Governments tend to organize themselves by department: The Department of Foreign Affairs, The Department of Tourism, The Department of Industry and Trade. However, from a reader's perspective, two potential problems arise with this sort of classification:

1. People are often not aware of what department deals with their particular problem.
2. Sometimes the task the reader wishes to complete requires content from more than one department.

To solve the above problems and create a more reader-focused approach, consider adopting a task-oriented classification approach. If you were a government you might come up with the following task-oriented classifications:

- Consumer affairs
- Employment
- Going on holiday
- Death
- Birth
- Retirement
- Education
- Housing

### **4.3. Geographic classification**

Consider employing geographic classification where you have content that is geographic specific. For example, Staples.com sells stationary and office supplies across the United States.

Its products are warehoused at various geographic locations. There's no point in a customer in San Francisco finding out it has XYZ photocopy paper, if that paper is currently only available in the New York area.

Staples.com thus classifies geographically. When you come to its website, one of the first questions you are asked is to say where you are geographically based.

### **4.4. Date classification**

Date classification is very useful where you are publishing a specific type of content on a regular basis. For example, the company newsletter may have been published every week over a ten-year period. Classifying it by date will allow you to better organize what will be a substantial quantity of content.

Date classification is often not ideal from the reader's perspective. They would prefer to see subject classification as well.

## **5. Classification design techniques**

Classification design is an iterative process that requires substantial feedback and interaction if it is to be a success. It is not advisable to design classification on your own. You really need to get a group of relevant people together to hammer out the details.

The quicker you can visualize your classification, the better. When you present classification as navigation on a website, then the choices and needs become much clearer.

## 5.1. Workshops

Properly planned workshops are an ideal way to design the classification, once the situation analysis has been completed. However, keep in mind the following when running classification workshops:

- Make sure that the right people are involved.
- The managing editor should run the workshops.
- Avoid religious debates about classification terms. Sometimes, people can have their own pet terms and get into heated debate defending them.
- Don't get carried away. Always be rigorous and realistic about the classifications you choose
- Champion the reader, the author and the organization's objectives.
- Visualize the proposed classification as quickly as possible.

## 5.2. Flip board and post-it note approach

Flip boards and post-it notes are a good way to visualize your classification. Here is a proposed approach:

- Write out all the classification terms gathered during the situation analysis on post-its.
- Stick up flip board pages around the room.
- Place the post-its on these pages in any order.
- Do a very rough mock-up of a webpage and as a result of discussion, begin to place post-its with chosen classifications on that page.

## 5.3. Card sorting approach

This is a supplementary or perhaps alternative approach using flip boards. Here is a proposed approach:

1. Write out all the classification terms gathered during the situation analysis on cards.
2. Ask readers, authors and representatives of management, to go through these cards, choosing what they feel are the relevant classifications.
3. Generally, it better for people to go through these cards quickly, rather than thinking over them for a long time.
4. Try to get 10-30 people to do this.
5. Ask them to choose the top 10-15 classifications that they would want to see at a top level.
6. Ask them to prioritize this list based on the most important getting the highest mark and so on. Let's say that someone chooses 10 classifications. They would then give a mark of 10 to their first choice and so on.
7. Create a spreadsheet with all the classifications listed in alphabetical order.
8. Create a print-out of this spreadsheet, with a box beside each classification. Give a copy to everyone carrying out the card sorting. Ask people to place their marks in the relevant boxes.
9. Create another spreadsheet with all the classifications listed in alphabetical order. Create a column for everyone who is carrying out the card sorting.
10. Place the relevant score in the appropriate column for each participant.

11. Create a total column at the end, which gives a total mark for each classification.
12. When all the respective marks of the participants have been entered, sort your spreadsheet based on the total column, highest mark first.
13. What you will get is a list showing the classifications that have received the highest total marks from all the participants.
14. This classification list and how it is prioritized becomes a very valuable input into what your top-level classifications should be and how they should be ordered.

## **5.4. Prototyping**

It's vital to prototype the classification as early as possible. Classification design can seem very theoretical until you see it presented in some sort of visual form. Creating simple webpage mock-ups on paper is more than sufficient in the early stages. Only when you're looking for sign-off from the editorial board should you consider more graphic-intensive prototypes.

One thing to watch out for here is the graphical design debate. Make sure that people focus on the classifications themselves, not the layout or colors.

## **5.5. Individual classification design checklist**

When examining a particular classification to see whether it is right or not, ask the following questions:

1. Is it necessary? Is this classification in any way a duplicate of another classification already chosen?
2. Is it clear? While it's good to be brief, it's better to be clear. Does this classification have multiple meanings?
3. Is there a simpler word or phrase to use?
4. Is there a shorter word or phrase to use? "Never use a long word where a short one will do," is one of the six classic rules of writing developed by British novelist George Orwell. It particularly applies to classification design.
5. Does this classification follow web convention? Classifications such as Home, Contact, Help, Search, About, What's New, are what people are used to. Coming up with new names for the above will only confuse people.
6. Is this classification roughly the same length as the other classifications at the same level? If 9 out of the 10 classifications at a particular level use one word, and one uses four words, it won't scan well.

## 6. About the Author

Gerry McGovern is an internationally acclaimed content management consultant and author. He has spoken in over 20 countries on content management, and has published three books on the subject.

His two latest books are Content Critical and The Web Content Style Guide (Financial Times Prentice Hall). Content Critical has been described by Knowledge Management Review as a "bible" and by Content Management Focus as the "standard text."

Gerry is editor of the e-business section of BUSINESS: The Ultimate Resource, the largest business reference encyclopedia ever published (Bloomsbury/Perseus). He is a member of Financial Times Knowledge Dialogue, a select group of 'thought leaders' who advise senior executives. He is a regular contributor to variety of magazines, including ClickZ and Information World Review.

Previously, Gerry was founder and CEO of Nua, a developer of content management software and solutions. In 1996, Nua received the Best Overall World Wide Web Business Achievement award from the European Union. Gerry McGovern holds a degree in management from Trinity College, Dublin.

Gerry publishes a weekly email newsletter on web content issues, entitled New Thinking.

To subscribe to New Thinking either:

Send a blank email to:

<mailto:new-thinking-join-request@list.adventive.com>

Or go to:

<http://list.adventive.com/SCRIPTS/WA.EXE?SUBED1=new-thinking&A=1>

[www.gerrymcgovern.com](http://www.gerrymcgovern.com)

[gerry@gerrymcgovern.com](mailto:gerry@gerrymcgovern.com)