Emblems in the Digital Age

by

Mohamed Kotb

Department of German Studies

McGill University, Montreal

July 2002

A thesis submitted to the Faculty of Graduate Studies and Research in partial fulfillment of the requirements of the degree of Master of Arts

Your file Votre référence
ISBN: 0-612-85864-2
Our file Notre référence
ISBN: 0-612-85864-2

Abstract

This thesis deals with the representation of emblem literature in digital media in the modern age. A discussion of issues related to new media such as the advantages and disadvantages of digital media as well as copyright issues is presented. There follows a discussion of different technologies related to modern means of publishing, notably Acrobat technology, HTML, XHTML, and XML, and how they could be best used to serve the goal of dealing with emblems by means of digital media. A discussion of digitizing and indexing emblems as well as CD-ROM technology is also presented. This leads to an evaluation of some Internet web sites and a CD-ROM edition. The thesis concludes with a summary evaluating the success of modern attempts of presenting emblem literature in modern digital media.

Résumé

Cette thèse traitera de la présentation de la littérature emblématique dans les médias numériques d'aujourd'hui. Nous aborderons plusieurs aspects relatifs aux médias numériques, tels leurs avantages et leurs désavantages par rapport aux moyens plus traditionnels de publication. Nous présenterons également le débat entourant la question des droits d'auteur ainsi que les défis que posent l'Internet et les nouvelles technologies pour le système juridique et vice versa. Par la suite, nous examinerons différentes technologies spécifiques à la publication numérique, notamment Acrobat, HTML, XHTML et XML, en mentionnant leurs avantages, leurs désavantages et la meilleure façon possible de les utiliser pour servir notre but, c'est-à-dire le traitement des emblèmes par les médias numériques. Nous discuterons aussi du mécanisme de numérisation des emblèmes en montrant les défis auxquels nous devons faire face. Cela nous conduira à une évaluation de quelques sites web ainsi que d'une édition sur CD-ROM. Nous conclurons cette thèse par un résumé qui évaluera le succès de la tendance actuelle de diffuser la littérature emblématique par l'entremise des médias numériques.

Acknowledgments

This work is the fruit of lengthy discussions and correspondences with my professor Dr. Daly. I met Dr. Daly as a student in the Department of German Studies. It did not take long to realize that I was fortunate to share the same seminar table with a professor of exceptional abilities not only in the academic field but as well in his humane way of treating his students, regardless of where they come from, with a fairness that sets an example, respecting their interests and encouraging them to develop their own opinions. Working with Dr. Daly, has been both challenging and enriching at the same time. Words cannot express my gratitude for a great professor whose existence made all the difference. Special thanks as well are due to my co-supervisor Dr. Richter whose advice and moral support has been of great value to me. Special thanks as well to Dr. Reber who best represents the German word: Menschlichkeit, and whose moral support has been crucial to me in moments of doubt. I would also like to thank my dear friends Karim and his wife Chandai for their tremendous support and my dear colleague Geneviève Dubé for her help with French texts and for her friendship.

Table of Contents

Introduction

The main objective of this work is to bring together what might seem to some as two distinct worlds: the world of emblems having its roots in the sixteenth century and the modern age marked by the growing importance of digital media and by the ever-growing importance of the Internet and its influence on our world. A closer look, however, will reveal that this distinction is a false one. I believe that the two worlds are not as separated as it might seem at first glance. There are two reasons for that. Firstly, while some might think of emblems as a part of the European cultural heritage, I believe that the world of emblems is still present in our world as I intend to show in the first chapter of this work. Understanding emblems will, therefore, help us understand many modern forms of communication, which have their roots in the emblems' heritage. Secondly, the unprecedented means of communication offered by modern technology remain without soul if their usage is limited to commercial aspects. These means should be used to preserve and present culture in its various forms and from various centuries. Consequently, the marriage between emblems and the digital world is not only desirable but also necessary. While good intentions are essential for the success of any marriage, they are not enough. Many issues and details should be discussed, even though they might not seem relevant at first glance.

Although some tend to debate the relationship between form and content in literature, this relationship is an integral part of emblems. It is demonstrable that earlier emblem writers and artists spent much time with their publishers to make sure that their works were presented in an appropriate manner that reflected the message or the content of emblems. David Graham notes:

> "It is well known, for example, that writers of Aesopic fables employ a limited number of well-worn structural, stylistic and typographical techniques calculated to ensure that the point of the story is prominently available not only to the mind but to the eye."[1] (3)

Just as painters need to know about the materials they use in their paintings and surgeons need to be familiar with their tools, emblem scholars in modern ages need to understand techniques related to the digital world and their underlying concepts to ensure the integrity of their work. The development of digital media offers us many advantages, but it also brings some risks. Chapter 2 discusses several of the advantages and highlights many of the dangers related to digital media, some of which are legal issues, specifically issues of copyright which I will discuss in some detail in chapter 3. Realizing the nature of this work and its limitations, I find myself obliged to commit what Peter Daly would call the *error of omission* rather than the *error of commission*, and limit my discussions to the main concepts of such issues related to emblems and to Internet

technologies, namely HTML, XHTML and the emerging XML mark-up languages. Also I will discuss a technology that tries to fill the gap between the world of print and digital presentation, specifically Acrobat. These topics are covered in some detail in chapter 4. Chapter 5 could be considered as an extension of chapter 4 discussing issues related to the digitization process and one of the main publishing media, namely CD-ROM, and the challenges we have to face due to the unique nature of emblems. In chapter 6 I will discuss modern attempts of digitizing and presenting emblems starting with an early attempt by David Graham, "The Macintosh Emblem Project", and evaluating most recent attempts made by several universities and some commercial entities and discuss several aspects related to the content, technical process, and presentation of emblems in these projects. I will also evaluate an emblem encyclopedia published on CD-ROM. While I take into consideration the fact that we are dealing with a relatively new media, I believe any objective criticism should evaluate the positive as well as the negative parts. Although I intend to make every possible effort to make my criticism as objective as possible, I am aware that some might disagree with the matter presented. The sole purpose for this work is to suggest improvement in the hope of making the marriage between emblems and the digital world a successful one. I will present a final conclusion assessing the success of these attempts in achieving what I believe should be our goals in digitizing and presenting emblems. I believe these goals should include the following: preserving, presenting and helping understand emblems and the emblematic world.

Chapter 1

Emblems and Modern Times

1.1 Introduction

What are emblems? The question is by no means a new one, and the answer is likely to differ according to whom you ask. The first emblem book was published in Latin in 1531 under the title "Emblematum Liber". The author was the Italian humanist Andrea Alciato described by Encyclopaedia Britannica as "*The father of emblem literature*" [2] and widely perceived among emblem scholars as the most influential emblem writer. The influence of Alciato was reflected in the various translations and editions of his books, which exceed 170 editions. Modern scholars were deeply influenced by his works in their attempt to define emblem. The influence of Alciato on modern scholars as well as differences related to the definition of emblems are reflected in Daly's writing:

> "As far as generic assumptions are concerned, opinions still differ on what constitutes an emblem, and this has shaped bibliography. Mario Praz had a broad understanding of the emblem, and his bibliography is still the most valuable and informative. Rosemary Freeman, on the other hand, had a narrower conception, restricting the use of the term to the three-part combination of motto, picture, and epigram associated with Alciato."[3] (63)

It is not my objective here to re-define the term or build a theory for its genre. However, this discussion is important for two reasons. Firstly, differences about what constitutes an emblem and the value of the different elements will influence the approach one might take in subjecting emblems to modern technologies. Secondly, this will likely influence the decision of what is to be digitized. In other words the way we define and understand emblems will influence both *what* and *how* emblems will be presented in digital media. Perhaps the most common approach to defining emblems is the descriptive approach, which outlines the three elements comprising an emblem as three parts: *motto, picture* and *text*. The relationship between these three elements as well as the evaluation of each element and its importance for emblem is not agreed upon. Albrecht Schöne uses the word "*Spannung*", which translates as *tension*, to describe the relationship between these elements, a description, which I find vague. I prefer in this context the description used by Daly as he describes this relationship as: "*functional interrelationships between text and image*" [4](67). Schöne believes that the three parts have a double function, which he formulates as "*Darstellen und Deuten*" [5] (22) [representing and interpreting] (my transl.) Alain Michel believes that the pictorial part serves as a memory aid that helps the user to keep the idea presented, which he or she might otherwise forget without the image: "... *la mémoire retiendra plus facilement la*

structure ainsi visualisée" [6] (24) [... memory retains more easily the visualized structure] (my transl.) He attributes the idea of visualization as memory aid to Cicero and ancient Greek rhetoricians. Such interpretations, while they reveal important parts about emblems and the functions of their parts, are by no means conclusive and they reflect the approach of the respective scholar(s) and his (their) understanding of emblems.

The symbolic nature of emblem is another aspect widely agreed upon. Elements in emblems are symbolic representations of human characteristics or other abstract notions in our life. Emblems make the abstract concrete to us, or as John Aiken puts it in a simple definition of emblems for children more than two hundred years ago as: "*a visible image of an invisible thing*"[7] (145). What the message is and how to interpret the different elements is a matter of disagreement among emblem scholars. Schöne presumes certain "*potentielle Faktizität*" [8] (75) [potential facticity] of the emblem without which the emblem loses its validity. Jöns, on the other hand, pays more attention to the artistic nature of emblems and the symbolic mode of thought they represent, which inspires him to dedicate more attention to medieval symbolism. Daly summarizes as follows:

> "In his treatment of the emblem as a 'mode of thought', Jöns lays particular stress on the role of medieval symbolism ('Ding-und Bild-Symbolik') and the allegorical interpretation of the bible." [9] (41)

The symbolism used is culturally dependant. Knowledge of the culture where emblems were produced as well as the cultural background of the emblem writer is necessary for an accurate interpretation of emblems. Indeed, emblems show an interdisciplinary nature, which has contributed to the difficulty of defining emblems. In short, the difficulty of defining emblems and evaluating the different parts could be summarized as follows:

- Several disciplines are coordinated in creating emblems. Understanding emblems requires knowledge in different disciplines.
- The process of emblem books production often involved different countries; the author could have been Italian, the language could have been Latin or another national language, the publisher could have been in another country. The Netherlands enjoyed a great reputation in publishing emblem books.
- The roots of emblems date back to older times. Daly goes as far as to date this to ancient Egypt and shows the influence of Egyptian hieroglyphics on emblem writers[10]. There are differences among emblem scholars, however, on evaluating these resources and their contribution to emblem literature.

- Emblems were not limited to one European culture and national differences between emblem writers of different nations are common.
- Emblem books enjoyed great popularity for over 200 years in Europe and certain differences have taken place with the time.

1.2 Other Forms of Emblematic Literature

Beside emblem books, emblem literature includes other forms closely related to emblems, which I will briefly review here.

1.2.1 Impresa

Schöne dates the beginning of impresa to the end of the fourteenth century where it took its birth in France and was developed later in Italy during the French occupation [11] (42-43). Schöne refers to the close relationship between emblems and impresa when he writes: "*Emblematik und Impresekunst treten so in ein enges Verwandschaft- und Wechselverhältnis.*" (43) [Emblematics and the art of impresa show a close and interchangeable relationship] (my transl.). This close relationship has motivated some to consider emblems as form of impresa. This is reflected in Sulzer's opinion presented by Daly as follows: "*On the relationship of impresa to emblem, Sulzer goes so far as to assert that the emblem is a 'variety of the impresa*" [12] (29). While we might disagree with this opinion, it reflects the close relationship between emblems and imprese. Yet, the impresa, while similar to the emblem, shows two major differences. The first being the *formal construction* of impresa that constitutes only two parts: image and a textual part, the *motto*, composed of only a few words. The other major and perhaps more important difference is the public to whom the impresa addresses its message. Schöne explains this when he writes: "*bedeutsamer als der formale Aspekt sind für die Unterscheidung von Imprese und Emblem die intentionalen Abweichungen*" (45) [more important than the formal aspect are the intended discrepancies] (my transl.). Furthermore, he sees the origin of the impresa in the "*individueller Selbstbestimmung*" [the self-determination of the individual] (my transl.), whereas emblems deal with the "*Allgemeingültig*" [universally accepted] and "*überpersönlich verbindlich*" (45) [binding for all] (my transl.) to deliver a moral message. The personal nature of impresa is reflected in Caldwell's opinion when she states: "*In the sixteenth-century imprese were regarded as signs of identity and expression of human thought.*" [13] (232). The text or inscriptio in an impresa expresses usually a noble thought addressed to highly educated person(s). Emblems on the contrary address far broader public. They address well-educated and less educated person(s) alike. Imprese were often used as personal badges or symbolic representations of an abstract idea or a noble

characteristic. As with emblems opinions differ on what constitutes a perfect impresa and how to interpret the different parts and the relationship between the motto and the picture.

1.2.2 Coat of Arms

Also known as shield of arms or armorial bearings, a coat of arms constitutes graphic element(s) that could be accompanied by a word or a few words reflecting a certain value or attribute. The textual part is often placed above the coat of arms. Sometimes a ribbon is also placed at the bottom of the graphic with the surname on it. The class or gender of the bearer of coat of arms could be distinguished from some certain components in the coat of arms. Coats of arms date back to the twelfth century in Europe, where it was originally used to establish identity in battlefields. Away from the battle, bearing coats of arms used to denote loyalty to the clan and its chief and its traditions. Eventually, coats of arms were used to denote a family, a descent, a clan, a person, a profession or an institution. Figure 1-1 is a presentation of Jeanne d'ARC bearing her coat of arms at the coronation of Charles VII. Figure 1-2 is a presentation of Jeanne d'ARC's coat of arms.

1.2.3 Heraldry

The relationship between heraldries and coats of arms is very close and sometimes both terms are used interchangeably. Heraldries, however, include more complicated symbolic graphic elements and their use follows certain conventions. There are certain elements that are strictly used in heraldries only, most notably the shield. Heraldries are related to different disciplines, especially arts and the science of armorial bearings. Although they date back to the twelfth century, as do coats of arms, their usage was wide spread in the thirteenth century in Europe. As coats of arms they include a graphic composed of several elements of which the main component is the *shield*. In the fourteenth century the *crest* was introduced that was modeled onto the helmet. Other elements in heraldries are the *torse* which is a rope that sits on the top of the helmet, the *mantling* believed to be used to protect warriors from the sun and rain as they traveled, and it was usually one color and presented in heraldries as a mass flowing around the coat of arms, the *supporter*, which is usually a human being or an animal flanking both sides of the shield, the *motto*, which is often placed on a ribbon above the entire heraldry and in less frequent cases it is placed at the bottom of the shield with the surname inscribed on it, the *compartment*, which is used to show the coat of arms being supported from the bottom, the *shield*, which comes in

various forms, and the shape of the shield could indicate where the family came from. As in coat of arms, the gender and the rank of the bearer could be indicated through subtle details in the graphics. Figure 1-3 is a graphical presentation of the different parts of heraldry.

1.3 Coats of Arms and Heraldries in Modern Times

In modern times coats of arms and heraldries are used to denote belonging to a certain group or organization and they represent its values or reflect its history. Many institutions, schools, universities, associations, and even churches use coats of arms and heraldries to reflect their mottoes or histories. In many cases coats of arms are granted as an encouragement or reward for services rendered by a person to the group or to society, or to denote a certain position and privileges within a society. Individuals could also design their own coat of arms. The use of heraldries and coats of arms is strictly limited to the person to whom the coat of arms or heraldry was designed in the first place and to his own legitimate and rightful line of descent. It is therefore considered highly inappropriate to purchase and bear someone's coat of arms as it denies someone the legitimate and exclusive right of the usage of the coat of arms. Apart from distinguishing the bearer from others, they may include a certain message. In some countries the use of coats of arms and heraldries is controlled by the law. Many firms use coats of arms and heraldries as well as emblematic motifs as their logo or trademark, which is protected by copyright laws. To mention but a few, the car manufacturer Mercedes-Benz uses the combination of a circle and a star as a trademark (figure 1-4). The star is traditionally a symbol of nobility and the circle indicates infinity. Tac-Software, a company developing educational software programs uses the owl, known for its sharp sight and often associated with wisdom, as a trademark symbol (figure 1-5). The lion, a motif often encountered in heraldries and often associated with power, is used as the trademark of Metro-Goldwyn Mayer (figure 1-6). In recent times there seems to be a growing interest in using coats of arms and heraldries. The medieval tradition seems to be reviving again. In the course of my research I came across several Internet sites offering their help in designing and selling coats of arms for persons or companies or verifying newly designed ones, which indicates to me that it is a growing business as well.

1.4 Emblematic Manifestations in Modern Times

While emblem books belong to the past, manifestations of emblematic modes in modern times are present in many aspects of our modern life. Not only coats of arms and badges are widely used, but as well different strategies related to emblem literature and in some cases old emblems are being introduced in modern contexts to serve certain purposes. It does not take much

effort to realize that something like an emblematic mode informs the naming choice of modern weaponry, whether the name is *phantom* or *stealth* for jet fighters or *leopard* for tanks thus hinting at certain capabilities of the weapon by using animals or mystical creatures and figures. Military operations are given carefully chosen code-names in a process that is, in its essence and result, an emblematic one. The United States and its allies gave the name *Desert Storm* to their operation in the Arabian Desert during the Gulf War to emphasize the devastation the enemy could expect. There was a not-so-hidden message for both enemy and allies alike, who happened to be historically and geographically tied to the desert.

Countries choose their flags and national symbols with care. A careful analysis of such symbols could reveal certain messages or emphasize certain political changes. Egypt abandoned the traditional eagle in its flag as it created the United Arab Republic with other two Arab countries to assure them that Egypt did not seek any domination or try to extend its protection over them. The traditional eagle in the Egyptian flag was replaced by three stars of equal size. Eventually, the eagle was reinstated in the Egyptian flag as the political project failed and Egypt regained its traditional name, while confirming its Arab and republican nature in the name: Arab Republic of Egypt.

The emblematic mode is not limited to military and political domains. Many high-technology firms use emblematic modes in naming their products in the process of development or as they are introduced to the market. While the emblematic mode is encountered in most, if not all our aspects of life, the most important manifestation is perhaps modern advertising, which I present here in some detail.

1.5 Emblem Literature and Modern Advertising

One of the most important links between emblem literature and our modern world is modern advertising and its strategies of persuasion. Daly refers to the similarity between emblem, which constitutes picture, motto and text and modern advertising, which can also contain the three elements. He designates the strategy followed by both emblem writers and modern advertising when he writes:

> "Just as the emblem uses a symbolic picture to make visual and concrete the abstract idea contained in a motto, so too the advertisement can render an abstract statement visual through an illustration." [14] (353)

He is cautious, however, not to assume that the similarity of both worlds is a deliberate one:

> "I do not assume that it is necessarily emblematic. There is no reason to suppose that copy writers and commercial artists have stumbled upon

emblem books." (349).

While in principle I agree with him, the similarity is sometimes so strong that it is hard to accept coincidence as an explanation.

The watchmaker Breitling uses a motif as its trademark, (figure 1-7), strikingly similar to the picture presented in one of Alciato's emblems, (figure1-8), where we see a dolphin enclosing an anchor. Both symbols are important for sailors. The dolphin represents speed and the anchor represents stability and security. People believed that dolphins offer help for sailors in distress and the anchor is perhaps the most important element offering security for ships. The anchor is used by the watchmaker and the dolphin figure is modified to produce the letter B, the first letter of the brand name. There are two wings surrounding the anchor and the letter B. The wings could be interpreted as eagle's wings presented in a very similar way to its representation in emblem books. The notion of flight is confirmed in the written text, which refers to pilots.

The similarity of strategies of persuasion is encountered in many advertisements. Schöne refers to a technique used by emblem writers that was often used in literary works where symbols refer to more than what they present. This requires what Schöne describes as: "*die stillschweigend-selbstverständliche Voraussetzung der Beweisführung*"[15](68) [the tacit, self-evident assumption of the presented proof] (my transl.). The same technique was used later in literary works, which employed emblems as argument in what Schöne calls "*Argumentum emblematicum*" (69) [emblematic argument] (my transl.) to prove something, defend a position or to prove the contrary of someone's argument. He gives an example of this technique using an emblem of Joachim Camerarius, *Oculis vita* (67), (figure 1-9), where we see a couple of ostrich birds looking at their eggs and breathing on them to bring them to life. The content of the emblem is presented in the drama *Epicharis*, which was presented in the theatre in 1666. Voulsius Proculus utters the words: "*Des Straußes Auge kan die Jungen lebend machen*" (68) [The Ostrich's eyes can bring its chicks to life] (my transl.) The same emblem was used in different contexts by literary figures in other works to prove their point of view. Schöne comments as follows: "*...auch hier soll der Angesprochene überzeugt werden von einem Sachverhalt, den er nicht glauben mag*" (69) [...here too, the addressed person should be convinced of something he would not believe] (my transl.). Several examples are cited in his work and such strategy of persuasion, namely using what people believe to be true to convince them of what they might put into question, is often encountered in literary works as well as in modern advertising. The German carmaker Opel published an advertisement in Der Spiegel Extra[16], (figure 1-10), for the *Corsa*, which is classified as a small car. The advertisement shows Albert Einstein standing next to the car and pointing to it with his pipe. In the background we see mathematical equations on a

blackboard. The title or the motto reads: "*GRÖßE IST RELATIV*" (24-25) [size is a relative matter] (my transl.). The motto hints at Einstein's theory of relativity, one of the greatest scientific achievements of the past century; most people believe of its accuracy. The text reads: "*Wissenschaftlich betrachtet ist Größe eine relative Größe. Bezogen auf den Corsa, sind zu unterscheiden...*" [From a scientific point of view, size is a relative matter. Applying this to the Corsa, one can differentiate...] (my transl.). Different aspects of size related to the interior room of the car as well as to other size-related aspects are represented to indicate that the small car offers large space and many advantages. The meaning of relativity of what is big or great is transferred from the theory of relativity to the small car. The choice of the word *Größe* is deliberate. It can mean physical size or greatness. The text uses the words "*absolute Größe*" [absolute greatness] (my transl.) to refer to the market value of the car. The text refers as well to the greatness of its look: "*Das großartige Aussehen noch gar nicht eingerechnet*" [We did not count the great look of the car] (my transl.). We are led to believe that the car, while small, is a great one.

A similar technique is used by the Swiss financial institution Credit Suisse in an advertisement in the German magazine *Der Spiegel* [17](10-11), (figure 1-11). The title of the advertisement reads: "*Etwas mehr Schweiz könnte Ihrem Leben nicht schaden*" [a bit more of Switzerland could not harm your life] (my. transl.). Next to the title of the ad is a red sports car. The color was deliberately chosen. On the rear of the car and very close to the motto is a white highlighted cross occupying a portion of the picture of the car. The only colored objects in the ad are the red car and the white cross (the surroundings of the car are presented in black and white), a combination, which unmistakably refers to the Swiss flag. The luxury car and its luxurious surroundings clarify what is meant by more of Switzerland. The combination of the flag and the deliberate use of the country's name hints at the common idea many people have about Switzerland as the land of banks and financial institutions. The text of the advertisement refers directly to that: "*Dass die Schweiz hervorragende Vermögensberater hat, ist weltbekannt*" (10) [It is known worldwide that Switzerland has excellent financial advisors] (my transl.) The trust people have in the country is being channeled to the financial institution in the advertisement.

Another striking similarity to emblems is encountered in an advertisement for a brandy published in *Der Spiegel*, (figure 1-12). The main motif is that of a young lady sitting, or rather posing, on the back of a wild bull. The title of the picture is the brandy name being advertised. The text, which is placed under a bottle of the brandy reads: "*Un poco macho un poco angel*"[18] (9) [a little bit macho, a little bit angel] (my transl.). The picture is produced in black and white and the color of the slender young lady harmonizes with the black body of the bull. The race of the young lady is hard to determine due to the alteration of her body color and her hair is cut short

as if to exclude any notion of ethnicity. She could be black or white. The young lady seems to be in perfect harmony with the bull that remains peaceful as if it was serving the woman sitting on its back. There is a striking similarity between this motif and the one in the unicorn emblem presented by Schöne (110)[19], (figure 1-13), where the unicorn is lying down in front of a young lady, resting its head on her lap while the young lady is touching its head. In the background we can see some snakes slithering from the peaceful scene. The wild bull in the advertisement presented above is a perfect modern replacement for the unicorn. Beside its strength, the bull is known to be an untamed animal. People had similar ideas in the past about the unicorn. The text is written in Spanish although the magazine is a German one. The choice of the language is deliberate. Spain is famous for bullfights. For most Germans, Spain is a favorite tourist destination. Of course what is being encouraged in the advertisement is not the chastity of the old unicorn motif but the consumption of the advertised product.

Any comparative study of emblems and modern advertising is likely to reveal more similarities than we might at first think. Indeed, emblems are very much alive, not only as mode of thought but as well in their strategies and their themes, which are directly or indirectly related to emblem literature.

Chapter 2

Information between Print and Digital Media

2.1 The WYSIWYG Illusion

The invention of the printing press revolutionized the way we produce, preserve and distribute information. It moved this process from individual effort that was often susceptible to error to a mass communication media. It became an industry that kept on growing and has reached a high level of precision today. However, one basic feature remains unchanged: what could be printed depends on the technical capability of the producer according to the principle: *what you print is what you get*. The story of digital or electronic publishing is a different one. In its debut, technical developments were based upon the needs of computer scientists and engineers. Monitors were mere output devices to help engineers interact with computers. Computers and monitors grew in capabilities and now they are part of our homes and offices. However, one thing remains the same: computers still function on the level of bits and bytes. To accommodate computers to our needs, many systems have been developed. Unfortunately, they were not always compatible with each other. What was displayed on the screen looked different when printed. A user interface document preparation system was developed to overcome this problem and was known as WYSIWYG (What You See Is What You Get). The WYSIWYG created an illusion and an ideal that was never attained and it was often described as *what you see is more or less what you get* or *what you see is all what you get*. What is presented on the computer screen depends upon two major factors: hardware and software.

2.2 Important Factors

2.2.1 The Hardware Factor

The on-screen displayed document depends largely upon the monitor's capabilities and that of the graphic card supplied with the computer. The quality and fine details of the presented font(s), picture(s) or any element(s) presented on the screen depends upon the quality of those two components. A good quality graphic card allows for a good display especially when the computer is supplied with a good quality monitor. Most computers now use color CRT (Cathode-Ray Tube) display monitors, which use a combination of three color phosphors (red, blue, and green), arranged in adjacent trios of dots, called pixels. The more pixels on a given screen size, the higher

the image quality. This quantity of pixels is called resolution. Support for different resolutions depends upon monitor's specifications. Most CRT-monitors can be set to different resolutions. A monitor of high quality usually allows for high screen resolution and for high refresh rate of the displayed content. Notebook and laptop computers use other kind of displays, such as LCD (Liquid Crystal Display) or TFT (Thin Film Transistor) display screens. This technology is now advancing to the desktop monitors as prices become more affordable. They offer many advantages over the older CRT monitors. They consume less energy, have less weight and put less strain on the eyes, but they do not offer the same capability of changing resolutions as CRT monitors. Setting the monitor to a higher resolution increases the number of screen pixels within the same monitor area. This results in smaller screen pixels and smaller displayed page, since the number of pixels in the page itself stays constant. The on-screen size of an image depends upon three factors: the *size* of the monitor, the *resolution* of the monitor, and the *resolution* of the scanned image. Larger monitors use larger pixels than smaller ones resulting in a bigger, but less sharp on-screen image. When developing content for the Internet or CD-ROM, careful thought should be given to the resolution choice of images included and they should be tested with different monitors of different sizes as well as different resolutions. In other words, what is developed for digital media depends not only on technical capabilities of the producer, but must take into consideration those of the intended users as well.

2.2.2 The Software Factor

Printing books depends on the format supported by the publisher. Displaying content developed for digital use depends on the capability of the system used by the user. If the font used by the developer is not installed on the users system, the system may replace it by a default font or another font of the same font-family. Unpredictable results may occur as well. If the format used for digitized images is not supported by the user's system, the content will simply not be displayed. In other words, the software installed on the user's computer will be crucial in displaying the content being developed. Due to the many differences of systems it is virtually impossible to have an ideal solution. Choosing the lowest common dominator will ignore the more advanced capabilities of most users. Developers should develop their content for 85% of users' machines. It should be noticed, however, that what was in common use a few years ago is no longer valid for most users today. In the past, developers used to develop content for 14-inch monitors with 640 by 480 screen resolution. Most developers now develop content for 800 by 600 resolution.

2.2.3 The Intended Public

Apart from the major issues of hardware and software capabilities, one should always consider the public to whom the content is addressed. Users who are willing to play games on their computers are more likely to invest in more advanced monitors and graphic cards and will be happy to manipulate their screen resolution or other individual settings to adapt to a certain game. People studying literature are less likely to have the same behavior or capabilities.

From this general presentation, and taking into consideration how users utilize the Internet, we should notice the basic differences between printed and digitized media as follows:

- Designing for printed material is a two dimensions effort, where the dimensions of the page are clearly fixed. These dimensions vary widely in digital media.
- Print reflects a much higher quality of what is presented, whether it is images or text. Content displayed on the screen puts considerably more strain on the eyes of the reader. For this reason, digital content should be divided in a manner that does not force the user to read for a lengthy period on the screen.
- In printed books, the publisher must make a decision prior to publishing regarding layout and design issues. In electronic publishing it is a good idea to leave as much of these up to the user and not hardwire the preferences of the developer or assume that all users have the same settings or the same software.
- The method of communication in books is usually printed text and perhaps some images or illustrations. On the Internet more visual elements are necessary to attract and keep Web surfers. In printed media using colors and images is a costly process; digital media, on the other hand, allows the incorporation and production of such elements for negligible costs.
- The Internet not only provides access to information but it is also a visual experience. One must notice, however, that too many visual elements might distract from the message intended.
- The user's experience in printed media is limited to the book being read. The user of Internet or CD-ROM can move directly to another area of interest.
- Internet users are not very patient. Studies show that a typical Internet user will be willing to wait 10 to12 seconds on line for a page to download.
- Buying a book or choosing one in the library is often a decision made after some thought. Visiting a Web site could happen accidentally or through casual navigation. For this

reason it is important that the site states clearly what its mission is and not assume that the visitor knows this in advance.

- Contrary to books, Web pages are materials that are being delivered through many different networks. For this reason, Web documents should be as light as possible. HTML files are generally very light on their own. Most Web pages include images and in some cases other files as well. These files should be kept as light as possible. Both the resolution of scanned images as well as image format influence considerably the size of images. Different image formats use different compression scheme. Other formats do not use compression; these are more appropriate for printed forms but not for the Web.
- Design issues influence as well the speed of transferring the documents over the Web. For instance, some content might be light itself, but if included in a complex table it will download very slowly, since the browser will show nothing before loading the entire table.
- Internet users do not always go first to the home page. In many cases they go directly to certain content in a Web site. For this reason one should always provide contact information as well as a link to the home page on all pages.
- Most Web pages are created using HTML, which has gone through several revisions. Browser makers created their own tags as well. One should use as many cross-platform tags as possible. It is also a good idea to use an older version of HTML rather than the latest one, which might not be supported by older browsers.
- The software that has been used to print a book is irrelevant to the user. On the other hand, software used to create Web pages could influence the experience of the user. Some software might add components that do not work with all browsers or require specific extensions on the server.
- Contrary to books, Web pages include a hidden but very useful part on the Web page that is usually included in the head section of the Web page. The value of a Web page could be significantly enhanced by careful attention to the different elements in this section. The title of the page for instance is generally displayed on the title bar of the browser and it should be chosen with care as it offers users as well as some search engines information about the content of the page. The browser also uses it to bookmark the page. For this reason it should be as descriptive and as short as possible.
- Many users save the content of the visited page for further study. When these pages are loaded later on, the browser shows its location on the hard drive or storage media as the URL (Uniform Resource Locator) address. To make it possible for the user to come back

and find the site, which could be of crucial importance if the material is used for scholarly purposes, the address of the Web site should be included in the heading section on all the Web pages of the site. The user can find this by viewing the source code of the page.

2.3 Dangers of Digital Media

The advantages of digital media and the Internet are sometimes too obvious to mention and in many cases they are taken for granted, as computers and the Internet become an integral part of modern society. To mention but a few, digital media is much cheaper to produce, transport and distribute. A few CD-ROMs can hold the contents of a small to mid-sized library for negligible costs. Searching and accessing a certain content can be performed within seconds and the content can be further processed using other applications. The same content can be shared by many users through networking. CD-ROMs do not need a small army of employees to maintain them as libraries do, they have no vacations, and never ask for salary increases. Moving or traveling no longer means leaving precious works behind. The Internet on the other hand offers the possibility of reaching a much larger audience than any one could have dreamed of prior to its development. The wealth of information that can be accessed through the Internet is simply overwhelming. Users' experience is no longer limited to the content presented. Hypertext makes it possible to extend this experience immediately. Digital media and the Internet have changed fundamentally the way we think of scholars today who bear little resemblance with scholars only a few decades ago. A visit to any modern library will show a majority of students spending most of their time in front of computer screens interacting with or accessing knowledge in digital form.

Yet, these advantages do not come without a price. With the birth of the digital age came much enthusiasm, but also many concerns. While some of these concerns reflect a typical fear of change, nostalgia for older times, or misunderstanding of the nature of the digital media or the Internet, it would be irresponsible to dismiss all criticism. I will present a brief history of the Internet and then I will discuss some of the most repeated criticisms of the Internet, its resources, and of digital media. I will also address some issues that are not commonly addressed.

2.4 The Internet: Definition and Historical Development

Most books date the beginning of the Internet to the late sixties as the United States government's Department of Defense, during the cold war, tried to create a reliable and secure communications network that would survive an all-out war, in a project that was known as ARPANET (Advanced Research Projects Agency Network). I believe the history of the Internet

began much earlier. The history of the Internet is as old as the history of computers. In their earliest stages computers were not developed for personal use, as we know them today. They were large and very expensive machines that were used in big firms and institutions that could afford them. The need to make these computers communicate with each other became clear. The history of computing is also the history of networking. Indeed, one of the most powerful operating systems in use today, and not surprisingly on the Internet, is the Unix Operating System, which was developed in the sixties. The development of networking went slowly because of different standards used by different companies. As the advantages of networking became obvious, the US government attempted to make use of these capabilities for military and strategic purposes in the project mentioned above. To make computers communicate with each other, a common computer-language, known as a *protocol*, had to be created that would enable computers to communicate regardless of the platform or operating systems used. The standard protocol was invented in 1977, and was called TCP/IP (Transmission Control Protocol / Internet Protocol). By that time, other firms and universities developed their networks as well. Eventually, these networks were connected together. This huge collection of networks was given the name Internet. In 1985, the National Science Foundation (NSF) began a program to establish Internet access across the United States. It opened the door to other educational and governmental institutions and international research organizations as well. The Internet was finally a reality. Although the possibilities offered by the Internet were exciting, its usage remained fairly limited till the Web was developed in the early nineties and a *killer application* was created (the Mosaic browser) that gave the Internet a friendly face. The Internet became universal, and many people began creating their own Web pages. The success of the Web led most people to use the terms World Wide Web and Internet interchangeably. However, we should remember that the Internet is nothing but a designation for networks connected together across the world. We should also remember that the Web and the Internet are not one and the same. In this sense, the Internet could be compared to a huge collection of printing devices connected together in a publishing house that no one owns.

Yet, we must admit that more often than not, printed materials enjoy more respect than Internet resources. I believe there are two reasons for this. Firstly, people trust tangible media more than virtual ones. Secondly, the circumstances under which the printing press was invented and developed created an aura that still influence the way we perceive books and printed materials today. The invention of the printing press took place in a culture where reading and writing were privileges limited to the well-educated in the society. The first book printed was the bible. Eventually, printing became more affordable and became a mass medium for communication, and the material that was printed and distributed became much more varied in subject matter as well.

This development took a few hundreds years. The Internet on the other hand, has gone through this development in less than a decade. The Web that was invented to facilitate communication between scientists in CERN (Centre European de Recherches Nuclaire) became a mass medium for all kinds of content. Reading and even creating Web sites was never limited to any group of scientists or highly educated people. Unfortunately, those who criticize the Internet for frivolous content tend to forget that such material is readily available in printed media. Not everything in books is serious. Some books are evil. The history of the twentieth century is the best proof of this. One of the worst crimes in the history of humanity was preceded by a book that was printed and distributed in the twenties and thirties of the past century. Most of the criticism addressed towards the Web applies as well to traditional print media.

2.5 Criticism of the Internet and Digital Media

The Intention of the Content Provider is put into Question

Wang, Das, Sunio and Schlein cite the famous acronym of Heinlein [science fiction writer] "TANSTAAFL: There Ain't No Such Thing As A Free Lunch." They believe "This thought may well apply to the World Wide Web as well" [20] (149). This criticism is found in most books dealing with the evaluation of Internet resources. It reflects a traditional capitalist approach that implies some kind of relationship between what one pays and what one gets. This principle is sometimes truthful in our daily practices, but not always. The Internet has a different mechanism. One of the driving forces of the Web is the sincere desire of some individuals or institutions to distribute content they want to share with others, regardless of the added costs. One example of this is the W3C (World Wide Web Consortium), the body responsible for the development of the Web. Published recommendations of the W3C are the definitive guide for Web standards. They are offered free of charge. Many governments also support different Web sites for public interests and many universities and respectable institutions make educational content available free of charge on the Web. This does not always guarantee the quality of content, but payment does not guarantee quality either. We should also notice that publishing on the Internet is a much cheaper process compared to older media, where each copy requires additional costs. It is true that more hits to a Web site require more resources; still, costs are minimal compared to traditional media. On the other hand, many commercial Web sites offer a wealth of information hoping to attract as many visitors as possible and thus advertisers. What is being sold in this concept is not the content to the users, but the users to potential advertisers. Many also offer content free of charge, hoping users will be encouraged to buy the same product in a more traditional form or in other media such as CD-ROM. Many successful products were initially offered free of charge to attract

attention. The product or some updates were later offered for a fee. The Web might not offer free lunch but it offers free access to a wealth of information of which some are excellent resources.

The Organization of the Content

Finding certain content or a Web site can be a very challenging venture as Miall notes:

> "The Internet is poorly organized in comparison with the library: "Trying to find information on the Web," says Debra Jones, "is like walking into a library after an earthquake."(1411)[21]

On the one hand, this criticism reflects a conception of the Internet as a huge library. It is not. There is no organization whatsoever of its content. Even if we regard the Internet as a library, something I strongly reject as it contradicts the very technical nature of the Internet presented previously, we must admit that finding books in a huge library is usually a more challenging task than in smaller ones. On the other hand, we must admit that retrieving information using available search engines is a real problem caused by the technology used to create Web pages. HTML, the language used to encode Web pages, concentrates more on the presentation of the text and pays no attention to its meaning or to the context of the text. Even more skilled search techniques help little as Miall asserts:

> "...even full-text searching provides access only to words, not to concepts. We still lack an effective, organized way to research the field of literary scholarship provided in thousands of books and journal articles (and, now, Internet sites), since the appropriate tools have yet to be developed." (1407)

Some search engines offer popular Web pages first, others show pages where the searched words are most mentioned or where the words are mentioned in the *meta* tag, or simply according to the payment made by the Web site to the search engine firm.

We must notice, however, that the Internet and the Web are not full-ended works. One of the developments that is currently taking place is the introduction of a promising coding scheme to Web pages that preserves the meaning of presented text known as XML (eXtensible Markup Language). Creating Web sites using XML and supporting search engines will substantially reduce this problem. We would be able to search for a certain "John Smith" the writer, the publisher or the emblem scholar for instance. I should mention, however, that this concept is not new. It has existed since the late sixties with the invention of SGML, an extremely powerful markup language, where structure and meaning of the text are preserved. XML is a lighter version of SGML. On the other hand, there is already an SGML initiative for encoding texts of literature known as TEI (Text Encoding Initiative), which has been used successfully since 1994. Due to

the fact that XML is a lighter SGML, it is compatible with TEI. XML allows us to create our own markup schemes and thus preserve our own meaning for the content encoded. It is possible for instance, to encode Shakespeare's works according to emblematic content. The question remains, however, who decides what concept is included or what the concept, here emblematic, is to mean. It is virtually impossible to find encoding schemes that would satisfy all humanistic scholars. Most encoding schemes are likely to concentrate on more objective criteria. For instance, TEI offers tags such as: <abbr> to contain an abbreviation of any sort, <author> to designate the author of a text, <sp> to indicate the speaker in a dramatic work or <stage> to indicate the stage direction within a performance text or fragment, and <authority> which supplies the name of a person or institution responsible for making an electronic file available. Thus no specification will tell us where to dedicate more attention to a certain passage, or draw attention to what might be controversial content. The very nature of literature makes it virtually impossible to create a coding scheme that responds to all needs of literary scholars and I fully agree with Fano's argument mentioned by Miall that "*human knowledge cannot be classified with sufficient precision for literature search purposes.*" (1407)

Preservation of Texts

Preserving knowledge has been always a challenging task in the past, and in the digital age it is even more of a challenge than we might at first assume. This problem is expressed by Miall as follows:

> "We face practical problems over the reliability of electronic texts and their long-term preservation... There is no guarantee that the Internet medium will remain much as it is now, able to support the reading much as it is now, able to support the reading of texts in the future in the way that printed texts have remained accessible for hundreds of years." (1411)

It is an unfortunate fact that many electronic documents were lost because of the incompatibility between different processing applications. Writing a document using a certain software makes the document hostage to the software used in creating it. XML effectively solves this problem. XML frees documents from the processing application. One of the basic concepts of XML is that XML-documents are readable to both humans and computers. Even if we do not have the processing application that gives the XML document its final form, we can still recognize the different parts of the document by reading the tags surrounding them. Reading a book marked according to XML scheme would allow us to recognize information related to the document such as: the author, the date of creation, the language used to write the document and other relevant information. It also

preserves the structure of the document and the different parts included within it such as the different chapters, subchapters, quotations and so on. The flexibility of XML makes it possible to create different coding schemes according to the nature of text encoded or even to create a coding scheme for one document only.

Whether we use XML technology or not, this is less of a challenging issue than another one, which is often ignored. We still need to save the content in a medium of some sort. As technology develops, older hardware standards are replaced by more recent ones. If the drive required to read a certain medium is no longer available, there will be no way to access the content. Theoretically, we can preserve many software applications on CD-ROMs or other media to ensure the readability of content created by such software. Hardware, however, is not as easily preserved. The floppy drive available in all PCs today is a living witness to this problem. Many drives with a larger capacity and faster access have been introduced. Unfortunately, they are not compatible with each other and the only drive that is compatible between all PCs is the old, slow and space-limited floppy drive. Many firms discontinued the production of certain drives and some firms totally disappeared. For this reason we must save electronic documents on several media, and a continuous backup on newer ones is a necessary precaution.

The Usability of Electronic Media

Using computers in their earliest stages required engineering skills. Although they have become more user-friendly, computers still have the reputation of being difficult to use. Miall confirms Andrew Odlyzko's complaint that:

> "Internet medium have been designed to serve developers much more effectively than users: "Little attention was paid to human factors. The result is that both networking and computing are frustrating for the end users." (1410)

We should not forget that many other products in our modern society are not very user-friendly either. We spend more than a few years learning to read and write and some never master the latter properly. Yet, no one suggests that we should abolish the comma and punctuation marks or make them more *writer-friendly*. People go to driving schools and usually need some time exercising their skills before they can be considered as experienced drivers. Why should we expect computers and electronic media to be entirely user-friendly products? Despite the fact that computer technology has come a long way, it is still in its infancy. While we might expect it to be more *user-friendly* in the future, we should not expect using it to be an effortless task. Yet, the dangers of underestimating the requirement of using it properly are more far-reaching than we

might assume. This applies equally to developers and users. Indeed, one of the problems of the digital media is its apparent ease that hides more complicated issues often ignored by users and many developers alike. Most Web pages are living examples of badly written HTML and ignorance of basic functionality of the Internet technology. I will give here one example of this: the date of last edition or update of a Web page or a Web site. Many Web sites write the date of the last edition or update to inform users if there is any new content. This could be easily added by including a sentence mentioning this fact encoded with a simple HTML code such as <p>, indicating a new paragraph. Many developers use WYSIWYG software that might add this automatically to Web pages. This approach has two problems. Firstly, the software will consider any change, even if it is a minor correction such as adding or deleting a comma, as latest edition or update. From the user's point of view this is irrelevant and he or she might be frustrated to find out that there is no real update even if the page indicates that. WYSIWYG HTML-editors could be compared to the grammar or proof editing tools of word processing. They make an assumption about what is right and what is wrong and while they can be helpful, they can also cause many problems if the developer is not knowledgeable enough to avoid bad software assumption. In other words, experienced developers are more likely to take advantages of WYSIWYG editors and avoid their pitfalls, while inexperienced developers, for whom such software products seem to have been developed in the first place, are more likely to suffer of their drawbacks. Unfortunately, the latter is more often the case. Secondly, such automatic addition of the latest update might require some component on the Web server, and if not available, the user will miss the update notice. On the user side, browsers could be configured to load a Web page according to different settings. If these settings are deliberately or erroneously configured to load pages from the *cache* files on the user's hard drive, where visited Web pages are usually stored, the user might end up having the same old page from his or her hard drive instead of getting the latest page from the server.

One common myth about the Internet is the phrase: connecting to the Internet. The Internet is not a place where all users can go and see the same product. Actually users *never* connect to the Internet. Users connect to ISPs (Internet Service Providers), and the user's experience depends largely upon the service offered by his or her ISP. Some ISP use Proxy servers, which save frequently visited Web pages and deliver them to the ISP's users upon request in order to speed up the delivery process of Web pages. If the proxy servers of the ISP, or for that matter the company, the university, or the location where the user might be at the moment of connection, is not offering the latest page available of the required source, the user will end up believing that the content he or she is getting is the latest available. While this might not be a

problem for Web sites that are not frequently updated, in some cases this could be a crucial factor if the latest news, for instance, is to be required. The Internet experience is a whole package of factors involving developers, users, software and delivery mechanisms that we need to be aware of. It is not as simple as clicking on a few icons as some IT firms want to convince us.

Reliability and Disappearance of Web Sites

We have to deal with two different kinds of reliability problems: *reliability of delivery* and *reliability of content*. Both problems are widely observed and Miall gathers both as follows:

> "Moreover, literary sites on the Internet, although they may benefit from the creativity of design the Internet offers, are often unreliable, with inadequate bibliographic preparation; each has a different design, following principles that may be obscure; and many sites have been short-lived" (1410)

The problems mentioned above are less common in printed media which enjoy more stability than Web sites, which can be easily created and disappear due to the lack of necessary support. Contrary to books, Web sites are not an ended work once published and they must be maintained. We must differentiate between Web sites created by enthusiastic individuals, and well-funded Web sites of recognized institutions. The latter are less likely to disappear and more likely to be considered by careful scholars. Web sites do disappear, but in some cases it could be a false disappearance due to temporary technical problems or alteration of the interior structure of the contents offered. Web servers are capable of responding to certain hits within a certain amount of time. This capability is relevant to several factors of hardware support, as well as the server-software used to run the Web site. Many Web servers are configured to deny access to additional requests if they get more hits than they can handle.

Web sites are usually accessed through a URL (Uniform Resource Locator). The URL includes several parts: *protocol*, which describes the protocol being used (for the Web it is usually HTTP), *domain name*, also called the host name such as *mcgill.ca* in the following URL http://www.mcgill.ca, the *port*, which is usually 80 (Web browsers automatically use this default setting and users do not have to enter it), the *path*, which is the directory path to the file or resource being requested and the file, which is the actual name of the file or resource being accessed. The following is an example of all these parts: http://www.emblems.arts.gla.ac.uk/SM23B/l028.html. If the user tries to access a certain file on the server directly, and if the structure has been changed, the user will get the dreaded "404" error message indicating that the site is not available. Changing the structure of a Web site is a common occurrence on the Web. Many organizations offer some search mechanism to help users find the

desired content. Unfortunately, many universities' Web sites offer a poor search mechanism, if any.

Another serious issue is the selling of some reputable Web sites which happens when the holder(s) is (are) no longer capable of providing the necessary support or simply want to make a profit by selling the Web site. Buyers usually keep the original URL but may run the site for different purposes. This development is more likely to take place with commercial Web sites and with Web sites with the extension ".com" and less likely to take place with Web sites with the extension ".edu" ,".gov" or ".org" for instance.

The *Reliability of content,* on the other hand, is an important issue. "*Assessing Internet resources..*" writes Gibaldi, "*is a particular challenge.*" (26)[22] Most Internet Web sites pay attention to the presentation of the content but rarely offer adequate information about the people responsible for the site or their credentials. The rigor applied to printed materials should apply as well to Internet resources. Reliable and less reliable resources, are equally available on the Web which gives Wang, Das, Shnio and Schlein a reason to wonder:

> "With the vast amount of data available, how does the naïve web surfer determine the validity of the information that she finds?" (149-150)

This problem, while theoretically easy to solve by setting guidelines regarding the use of Internet resources and limiting these resources to recognized ones, is tightly related to different practices and behavior related to acquiring knowledge. This problem is mentioned by Miall who warns of what he describes as: "*flattening the structures of knowledge*" (1412). He believes knowledge is more than just information as presented on Web pages:

> "Induction into a discipline is a matter not only of acquiring information but of assimilating a set of theoretical frameworks, assumptions and practices that can take several years to master." (1412)

It is true that there is an evident danger resulting from the fact that acquiring information from reliable or less reliable sources is equally easy or difficult on the Web. Students in the Middle Ages often traveled long journeys to visit a well-known scholar or a college. The process itself obliged them to give some thought about the resources they were seeking. This process of examining resources may disappear or be reduced to a minimal effort. In many cases the Web seems like a flea market, where the good, the bad and the useless are virtually next to each other. What might be of interest to someone, is of no value to another. What scholars previously learned was not only the content of a book, but rather a whole set of accompanying elements. Unfortunately, the danger of flattening human knowledge threatens more traditional institutions

such as universities, where basic principles and elements of mind-education are being replaced by the ever-changing needs of the economic and technical demands of society. In such a situation, knowledge is reduced to technical manuals (that is what computer books are after all) that need to be memorized for an exam or employment purposes. Human languages seem to lose ground, at least on the labor market, in favor of programming languages and interaction between civilizations is being replaced by computers networking or trade related issues.

The Influence of the Media on Content

This is an issue to which we need to dedicate more attention. It is noted by Miall who writes:

> "More urgent, we have hardly begun to ask the relevant questions about reading practices and how these are influenced by the nature of the medium (especially when it includes multimedia elements)." (1411)

Many make use of the different multimedia capabilities offered by digital media. If not used with care, the message intended will be overshadowed by other effects. Multimedia elements are far more convincing than written text. Advanced techniques in digital images could very easily be abused to alter or falsify human knowledge, instead of preserving and presenting knowledge. While the content of a text and its integrity can be easily verified, this is a far more difficult task with multimedia elements.

Another serious, less-spoken of, problem is the shift taking place in the way knowledge is being preserved and produced from the hands of traditional institutions to software manufacturers. Evaluating software products usually follows a number of principles that used to be limited to computer software such as price, user interface, stability, speed, integration with other software products and so forth. Applying the same criteria to digitized knowledge bears the risk of underestimating the value and the reliability of content in favor of other factors that might not be as important to scholars. Some software manufacturers reach agreement with computer manufacturers to distribute certain software packages for a reduced price or even free of charge with new computers. Users will be unlikely to resist the temptation of using the free software and will be less inclined to buy other more reliable and more expensive products. Users are also put at risk of assuming that a well-known software manufacturer will be a good source for knowledge as well. While this might be true, the basic assumption of such a relationship between the two is entirely false. The experience of software manufacturers and their ability to produce stable, user-friendly products, as well as their ability to package different software together and facilitate the

integration between different products they offer, is sure to give them advantages over other traditional institutions specialized only in producing one sort of product.

Digital Media and Value Systems

Alonso recognizes the advantages of the new media, yet he seems to be worried about the influence of new media on ideological and political agency and the threat of the latter being compromised by what he calls the *Internet sublime*. He shares the same fears of the threat posed by the Internet to classical ideological and value systems with Sven Birkert whom he quotes as follows:

> "My core fear is that we, as a culture, as a species, are becoming shallower, that we have turned from depth-from the Judeo-Christian premise of unfathomable mystery-and are adapting ourselves to the ersatz security of a vast lateral connectedness. That we are giving up on wisdom, the struggle for which has for millennia been the central to the very idea of culture, and that we are pledging instead to a faith in the web. What is our idea, our ideal, of wisdom these days? Who represents it? Who even invokes it?" (1300)[23]

We should remind ourselves that the Internet is not about replacing any religion or value system. We should also remind ourselves that religions and value systems, in all their varieties, have not only been a source of wisdom, but of conflict as well. This is unlikely to change, primarily due to the fact that many claim that wisdom is confined to their own value system.

Legal and Cultural Challenges

The Internet poses serious challenges to legal systems and to copyrights, issues, which I will address more in detail in the following chapter. In the past, many countries were hesitant to export their books to countries with which they had no copyright-agreements. Published materials on the Internet can be accessed from anywhere. CD-ROMs or other digital formats can be easily transported or even made available through Internet connection.

Enthusiasm about using digital media and the Internet to disseminate literature or cultural content often ignores the fact that some content might be in conflict with other cultures and legal systems. Even within the same society different individuals or groups might advocate values that contradict fundamentals of the modern society, or at the least put these fundamentals at risk. With few requirements, individuals will be able to reach a potentially wide public. We might realize that we have freed a genie from a bottle and while it might serve us, it may also do just the opposite.

2.6 Final Thoughts

Much of the criticism addressed to digital media and the Internet combines content and medium, an approach I disagree with. Some articulate shortcomings that need to be addressed. Others expect perfection of the new media, an ideal that is unlikely to be attained. We must not forget that digital media and the Internet are somewhat recent technologies and they go through a phase of development, which offers advantages and disadvantages. We should remember as well, that these technologies were invented in the first place to overcome inadequacy in traditional media and to respond to some needs related to the unprecedented growth of human knowledge. We should not consider them as a *replacement* for older methods, rather as an *extension* of these methods.

Chapter 3

Copyright and the Internet

Any discussion of the Internet and digital media would perhaps not be complete without discussing questions related to copyright. If the talent of the writer makes it possible for him or her to create, and the technical tools provide the possibility of publishing and distributing the work, aspects related to the law are equally important. A discussion of issues related to copyright is for this reason necessary.

3.1 Definition and Historical Development

The Canadian "Guide to Copyrights" defines copyright as follows: "*it is the right to copy*"(8)[24] a definition, which I find not only too simple, but too narrow as well. The word right implies both an ethical and a legal right. While most of us today think of the legal aspects related to copyright, we should not forget that the moral rights were recognized prior to the legalities, as we know them today. This moral right was perhaps the main factor in an early dispute in the following story cited by Bielefield and Cheeseman:

> "In 567 Saint Columba (521-597), an Irish monk, on a visit to a neighboring monastery, made an unauthorized copy of the abbot's Psalter. When the abbot found out, he demanded that Columba immediately return the copy. When Columba refused, the abbot sought the help of a local king, who ordered the return of Columba's copy to the abbot under the principle "To every cow her calf." (5)[25]

Apart from the conflict and the outcome of this incident, one notices the technique used in copying, which was here merely handwriting. The authors inform us that the world's first copy right law was passed in Britain in 1710 and was known as the Statute of Queen Anne ca. 233 years after the first book was printed in England and ca. 255 after the invention of the printing press; an invention which dramatically changed the way books were produced. The authors believe that there was a strong link between the invention of the printed press and the copyright law. They argue that it took this long:"

- to experince the real impact of the invention of the printing press
- to develop the new legal concepts necessary to deal with that impact of technology

- for the Parliament (the legislature) to become aware of the importance of this technology for the national economy
- to bring all of the parties together - the Crown, the Parliament, the Stationers' Company (the publishing industry), authors and public support." (13)

While I doubt that the aspects related to getting the parties together were behind this delay, the relationship between technological development, commercial aspects and copyright laws is hard to ignore.

3.2 Internationalization of Copyright

While the Statute of Queen Anne was a national law, copyright laws were developed later to assume an international dimension. This is most likely due to the development related to the print technology in various countries, as well as to the development of communication and transportation methods and the growing commercial value of the different copyrighted items. Copyright laws have developed in various directions. The main aspects to be noticed are as follow:

- The scope of what is being protected has been continually extended to include not only books, maps and charts, but several other activities such as musical compositions, photographs, works of arts ...etc.
- The duration of the protection has been constantly extended from 14 years in the Statute of Queen Anne to 50 or 70 years after the death of the author.
- The territorial protection extended beyond national borders, as different nations started to conclude bilateral agreements for mutual copyright protection. The major development in this direction was the creation of the Bern Union 1885, which resulted in the Bern Convention in 1886. The Bern Convention went through several modifications, the most significant of which took place in Paris 1971, and at present it constitutes the major international convention for the protection of literary and artistic works. Signatories of the Convention constitute the Berne Copyright Union.

Technically speaking, there is no such thing as an International Copyright that will automatically protect an author's work throughout the world. Protection against unauthorized use in a particular country basically depends on the national laws of that country. The Bern Convention tries to create a minimum set of rights for authors. These rights can be extended in national laws or through bilateral agreement between the members of the Union. With the internationalization of world trade, countries willing to join the GATT (General Agreement On Trade and Tariffs) agreement are required to sign one of the three international conventions which

are: the Bern Convention, the Universal Copyright Convention (UCC) and the Trade-Related Intellectual Property issues (TRIPS).

3.3 The Internet Challenge

The challenges that technology and new media pose to the current legal system were demonstrated two years ago in a famous case concerning a Canadian Web site. The Web site, icravetv.com, which started 1999, was broadcasting the programming of 17 Canadian and U.S. TV stations online without permission, in what was described by the TV networks as: "*one of the largest and most brazen thefts of intellectual property ever committed in the United States*"[26]. Although the Web site did not charge a fee for its service, it attracted many advertisers from Canada and across the borders. American broadcasters and sports leagues challenged the Canadian Web site in a US district court. The defendants argued that their activities were permissible under Canadian law and submitted a declaration by a Canadian law professor, Michael Geist, who confirmed that (§24)[27]. Nevertheless, the Web site lost the case and a restraining order was issued by U.S. District Judge Donald Ziegler on January 28, 2000. The decision of the court was widely perceived and presented in the press and on the Internet as the order of an American judge to shut down a Canadian Web site[28]. For many it was a case of law and jurisdiction[29]. Some wondered if a US judge has jurisdiction in Canada and whether he has the right to shut down a Canadian Website. However, a careful reading of the "Findings of Facts" as presented by the court reveals the following:

- The case was perfectly within US and the court's jurisdictions (beside many aspects relevant to the nature of activity, the nationality and residence of the defendants, the judge referred to the Bern Convention which states: "*The applicable law is the copyright law of the state in which the infringement occurred. . .*")(§24)
- The defendants, while not charging any fee for accessing the Web site, were actively involved in marketing their Web site in the US and attracting paid publicity on the Web site.
- The content that was broadcasted on the Web site was a product of US firms and protected by US laws.
- The plaintiff was not seeking to shut down the Web site, but rather to stop the Web site from broadcasting the content to US territory.
- The issue was not actually shutting down a Canadian Web site in a US district court. In other words, if the Web site had able to come up with a technology that limited access according to geographic areas and national territories, it could have continued its

activities. Since such technology was and still is not available, the only way to limit access was to shut down the entire Web site.

While the judge's decision seems to have a solid basis according to the American judicial system, and while it conforms with the Bern Convention[30], especially with Article 5, giving the domestic law the relevance in such a case, and as well with Article 16, which allows seizure of the copyrighted work where the work enjoys protection; the decision poses serious questions as to what can be permitted on the Internet. The Internet by its very nature makes political and geographic borders irrelevant. Indeed, one of the main reasons why the Internet has now become what it is, is its universal nature. Is it possible to see in the future endless lawsuits based upon the fact that a certain content of a Web site is copyright protected in another country or against the law in another country? What about cultural differences between nations that make what is morally and legally accepted in one country immoral and perhaps illegal in another? Are we supposed to develop a technology that makes the Internet limited to geographical, political or legal borders? It is clear that the court's decision, while a major victory for US broadcasters and sports leagues, poses serious questions that must be addressed. One thing is clear, however, the entire case and its outcome were closely related to substantial commercial interests. It is very unlikely that we will see similar cases, if the parties involved do not have the same commercial and financial power and interests. This might leave us with an open question about the ethics of copyright laws and their practices, which in the first place are supposed to protect the moral rights of all authors equally but ends up being used solely to protect commercial interests related to powerful institutions capable of supporting expensive legal procedures.

3.4 Fair Use

The term *fair use*, described in the Canadian Copyright Act as *fair dealing*, is a term that has never been clearly defined. The Canadian "Guide to Copyrights" describes the situation as follows:

> "The line between fair dealing and infringement is a thin one. There are no guidelines that define the number of words or passages that can be used without permission from the author. Only the courts can rule whether fair dealing or infringement is involved." (11)

One major aspect related to fair use and how it could be interpreted is the copyright warning included in most copyrighted works and mostly reads as follows: "*No part of the content of this book may be reproduced or transmitted in any form or by any means without the written permission of the publisher*". Cornish notes that:

> "This statement has never been tested in law. It is generally thought unlikely that it would stand up in court as it tries to prohibit what the law allows." [31] (45)

He rejects its interpretation as a contract between the user and the copyright holder. He argues that "*the general opinion is that it is there to frighten rather than be enforced!*" (45), an interpretation which I disagree with. I believe that the warning text constitutes a contract between the user and the publisher. It is a precondition set unilaterally by the copyright holder. While I agree with him that "*actual enforcement of it would be a very difficult thing to do, and costly in legal fees to establish as binding*" (45), I find it difficult to ignore the legal nature because of difficulty to enforce. Were it so, we would have to abandon many laws and legal rights for the same reason. Yet, the contractual nature of the warning does not necessarily mean that it is legally valid. Basically, a contract is any kind of agreement between two parties. Some contracts take the form of mutually agreeing upon certain conditions and obligations; others take effect by setting conditions and obligations unilaterally by one party and the explicit or tacit recognition of these by the other party(ies). Most of our daily activities are considered to be a form of the latter. Thus purchasing some goods or services usually takes effect by setting the terms and the price of the offered services or goods by the provider and the customer(s) accept(s) or reject(s) such conditions. Some contracts require written forms and mutually written recognition from both parties to be legally valid such as marriage and insurance policies. In most cases, however, these written formalities are not required. People do not need to sign a contract or any written form when they buy some goods or take a meal in a restaurant for instance, all which constitute forms of contracts and legal acts. Article 1387 of the Civil Code in Quebec describes what constitutes a contract as follows:

> "Art. 1387. A contract is formed when and where acceptance is received by the offerer, regardless of the method of communication used, and even though the parties have agreed to reserve agreement as to secondary terms."[32]

For a contract to be legally valid, - regardless of its form, written or not, tacitly or explicitly accepted, and regardless of the way the conditions were set, unilaterally or not- certain conditions must be met, however. One of the most important conditions is that, the contract cannot prohibit something that the law allows. While the copyright warning agrees in form with the previously mentioned definition of the contract in Article 1387, it does not conform to its content - given the fact that it prohibits explicitly the reproduction of any part in any form without previous permission. Article 1411 of the Code Civil reads as follows:

> "Art. 1411 A contract whose cause is prohibited by law or contrary to public order is null."

To avoid this conflict, some publishers use a slightly different form of the copyright warning citing in general the cases where copyright is permitted. Apart from any conservative interpretation of the law, whether for or against the copyright warning and issues related to it, the Canadian Copyright Act (Article 29)[33], as well as the Bern Convention (Article 10iii), make provision for fair use. The main guidelines are the following:

- The purpose and motive of the reproduction: commercial use is more restricted than educational ones or use for criticism or purpose of news reporting.
- The nature of the copyrighted work.
- The amount of copied or cited passages and how they relate to the whole work. Cornish notes that not only quantity counts but quality as well. Thus there might be an infringement of copyright if the copied part is a crucially important part of the work, even if it is only one page.
- The potential effect of the use on the value of works commercially available.
- Educational and non-profit and charity organizations enjoy special rights.
- It is up to the legal authority in each country to set the guidelines and judge what is fair use, and what constitutes an infringement.

3.5 Final Thoughts

There are two aspects, which seem to be missing in the discussion of fair use. Firstly, fair use is not only good for the users, but also for the copyright holder as it frees up resources that otherwise would have been wasted if the copyright holder has to respond to each request to copy a part of the published work. Secondly, there is the notion of creativity, which is a pre-condition for a work to be copyrighted. Can any work be considered as 100% creative? The cultural influence of society and its contribution in different creative works is a factor that is often neglected. Giving absolute exclusive rights to the author or the publisher threatens the very same source of creativity that led to the creation of the work. The notion of fair use deserves to be treated with more respect, not simply as an activity that we tolerate.

Chapter 4

Digital Publishing Technologies

4.1 Acrobat Technology

4.1.1 Introduction

The development of Acrobat technology started long before the Internet became what we know today. As computers were heavily used in the business environment and in the publishing industry, people realized quickly the difficulties of transferring and sharing documents due to different file formats used by different applications, different operating systems and font requirements. One could have exchanged files in ASCII text format, but text alone is not sufficient in most cases. In many cases graphics had to be included as well. The business community was looking also for a solution to archive its documents avoiding the endless piles of papers in their archives and thereby reducing paper consumption. What was needed was a technology that would enable the user to create, view and print documents while maintaining the documents' integrity without the need to install any of the fonts or applications that were used to create the documents. Since one of the main uses for such a technology was the office - at least initially - some security features were needed as well (office documents have legal value and should not be tampered with). Adobe Systems Inc. responded to this challenge by creating the PDF (Portable Document Format). The dream of a paperless office remains unattained as Frank Romano notes with humor:

> "With the increasing attention being given to the Internet, the idea of a paperless world seemed tangible. However, now that we know the truth of the statement there'll be a paperless office when there's a paperless bathroom." (33)[34]

Acrobat technology was very successful in achieving the objectives it was created for. With the development of the Internet and the need to share and transfer documents across different platforms and systems, Acrobat technology was available as an ideal solution. Although in its earlier versions Acrobat suffered from some limitations such as the lack of hypertext features, fade presentation on the screen and the lack of color support, it has overcome many of theses limitations. The interest in Acrobat technology as an ideal format for sharing and transferring documents assumed a global dimension. Romano rightfully notes: "*Acrobat is emerging as the software capable of creating the near ideal digital document*" (35). This success allowed some to compare Acrobat - particularly after inserting hypertext feature, and the capability of embedding

sounds and multimedia files - with HTML and some even wondered if HTML was going to be replaced by Acrobat. I believe this discussion is unnecessary. A comparison between the two technologies is like a comparison between a Ferrari and a truck. While a Ferrari is an elegant and fast car, it is ill suited for moving furniture.

4.1.2 How to Create Acrobat Documents

To read PDF files, one needs to have Acrobat Reader installed. To encourage the use of its product Acrobat reader is offered free of charge on the Internet. To create or modify PDF files one needs to purchase the Adobe Acrobat suite of applications. The suite includes three different applications, *Acrobat reader*, to read the files, *PDF writer*, to create PDF files from word processing applications, mostly for simple documents, and *Acrobat Distiller* for more sophisticated and challenging documents.

4.1.3 Advantages of Acrobat

- Acrobat offers an almost perfect solution for bridging the gap between DTP (Desktop Publishing) and digital publishing distributed on CD-ROMs or on the Internet. Regardless of the applications or operating system used to create the document, the document will be presented and, if needed, printed exactly as created and as shown on the screen. It offers an ideal solution for transmitting and distributing documents between users with different operating systems.
- It offers many useful features such as: search capability and the capability of including images, multimedia files and other formats.
- Adobe Acrobat allows the creation of language rich documents. This allows the presentation and printing of different languages such as ancient Greek, Arabic and Hebrew without any need to install fonts or to use an operating system that supports any of these languages.
- The security feature gives the creator of the document full control over how the document should be used. The creator of the document could allow or forbid editing, copying or even printing the document. The latest version (as of this writing) is version 5.0, which supports 128 bit encryption, which is the most secure encryption scheme available at this moment (it is used by the US government for classified documents).

4.1.4 Disadvantages of Acrobat

- The user may not have the Acrobat Reader installed.
- The size of PDF files is usually larger than simple HTML files, and they require more computer resources.
- Acrobat is a proprietary technology, which means any further development remains in the hands of Adobe System Inc. While the Acrobat reader is currently offered free on the Web, Adobe has no obligation to continue to do so.
- To create Acrobat files one needs the Adobe Acrobat Suite, which is a relatively expensive product.

4.1.5 How to Use Adobe Acrobat

- Adobe Acrobat offers an ideal solution for the reproduction of emblem books. Emblem books could be scanned, in total or in part, and presented on the Internet or via CD-ROM. Such digitized forms in PDF format could be easily archived for preservation, reprinting or sharing with other institutions and individuals.
- Adobe Acrobat offers the possibility of scanning or *capturing* documents directly and saving them in PDF format.
- Using the security feature available, one could distribute emblem books or related materials without having to worry about issues related to copyright, printing or changing the nature of the document. Only the creator of the document has full control over what is permitted or not.
- Adobe now supports progressive display, which makes the transferring speed of PDF documents faster on the Internet (previous PDF documents needed to be downloaded in total before they were displayed, which was a frustrating experience specially with long PDF documents). However, I believe CD-ROMs are more suitable for distributing PDF documents, especially lengthy ones.

4.2 HTML (HyperText Markup Language)

4.2.1 Introduction

The World Wide Web Consortium describes the HTML as the "*lingua franca for publishing hypertext on the World Wide Web*"[35]. While many in the computer industry tend to exaggerate the value of the technology presented, this description seems to be an understatement.

Without HTML most likely the Web would not have become what it is now. HTML pages are simple text pages that have been marked up using tags, which could be described as small notes surrounding the text of the document such as <H1>. To denote the end of the element marked up, most tags require closing tags which are preceded by backslash like this </H1>. Some tags have optional end tags and others have no end tags. The beginning of HTML is the <HTML> tag, which declares the type of the document for the browser. Following the <HTML> tag are two main sections of HTML document: The *head* element, which contains data about the document. It usually includes tags such as the <TITLE> tag, which gives description for the document and is displayed on the title bar of the browser and the *meta* element. The *meta* element can contain information, also called *metadata* because it is data about data, about the author of the document, the organization that created it, keywords for search engines, language attributes and any other type of information the developer might wish to include about the document. Some elements of scripting languages, which are mainly used to create some kind of interactivity, are also included in the head element. The *body* element is the main page that the user sees when he or she opens the page. Typically the body element includes the document presented and the tags denoting its structure such as <H1> for heading level one, the beginning of paragraph <P> and formatting tags such as <B> for bold or <I> for italic. Many formatting tags take a *semantic* approach leaving formatting of the different elements to the browser. An example of this is the <STRONG> or <EM> tags which instruct the browser to emphasize the element or make it stronger than the surrounding text. Some browsers might emphasize the element by displaying it as bold or as italics, or even in a different color. Other tags reflect the physical approach forcing the browser to present them in a certain manner. The <B> tag is an example for this.

4.2.2 HTML between the Past and the Future

The Web and HTML were initially designed for use by academics in a fairly limited way. The original design of HTML reflected those needs. Two main goals were to be attained: platform-independence and simplicity. In the beginning, HTML was largely an informal effort. The first formal version of HTML is version 2. In 1993, the National Centre for Supercomputing Applications (NCSA) at the University of Illinois released Mosaic, a Web browser, which became very popular in the Internet community. This presented a great opportunity for HTML and the Web to grow and many users created Web pages using HTML. In its origin HTML was more concerned about the logical structure of the HTML document than the presentation of text, which was supposed to be the function of the browser.

From its debut till the first formal version, HTML did not include a large number of explicit formatting tags. There was no means to center text and graphics, for instance, and only limited ways to format graphics on the page. There was a small set of tags that provided formatting based on the needs of the average academic paper. The Web experienced a rapid growth and along with it changing demands for HTML. Professional designers, writers, layout artists, and many from the DTP industry needed more control over the presentation of the Web pages. Changes in the following HTML versions reflected these demands. HTML 3.0 incorporated special graphics features, background colors and images, tables for displaying data, and other features that allowed for more control of the presentation of the Web page. Browser makers developed their own tags as well, which offered some extended functionality of the HTML but were not supported by other browsers. HTML tags offered in its latest version (version 4.01), were mere recognition of what browser makers developed rather than setting guidelines of how HTML should evolve. What started as a simple, structural and platform-independent markup language, became a complex set of tags, attributes and browsers' extensions. HTML was losing its simplicity and its platform-independence.

4.2.3 How to Create HTML Documents

HTML files are basically plain ASCII text files that can be created with any text editor such as BBEdit for Macintosh, or Notepad for Windows. The HTML filename ends in ".htm" or ".html". Word processing programs could be used as well, but they are not recommended as they add their own formatting tags and file extensions automatically. The user should be careful in this case to save the HTML files as text files only with ".htm" or ".html" extensions. On a more professional level there are other editors which help speed up the process of creating HTML documents, which can be generally divided in two categories: *text, or HTML Editors* that help the developer by listing the different tags and their attributes and insert them where the developer wishes. Some of these programs assign shortcuts for tags as well. The other category is the *WYSIWYG* (What You See Is What You Get) editors, which operate much like a word processing package and write HTML code and help put together not only a Web page but a Web site as well. While this category is very helpful, one should note, however, that in some cases these programs make decisions about the logical structure of the document and add tags, which do not necessarily reflect the intention of the developer. Careful attention must be given to the tags added and their conformity with the HTML standard since some software packages might add proprietary tags, which do not function with all browsers.

4.2.4 Advantages of HTML

HTML is the language of the Web. Apart from that, HTML offers many advantages and features that encourage its use in other electronic publishing forms.

- HTML files are extremely light. Different materials can be written in HTML format and be saved in portable form such as disks, CD-ROMs or any other medium. In fact, many computer companies and publishing houses have produced publications using HTML format.
- The hyperlink feature of HTML is one of its strongest arguments making searching and navigating the documents as easy as a click on a link.
- HTML makes provision for including images or other multimedia features. One can embed some scripting language allowing some interactivity with the user. Some search mechanisms can be embedded that allows searching the Web site for certain content.
- Virtually every computer in use today is provided with a browser of some kind. Reading and accessing HTML files, as well as creating them does not require any additional investment or installation.
- Learning HTML for basic Web pages or Web documents is not a challenging task. However, learning more sophisticated features is more demanding.
- HTML, if used according to the W3C recommendations, is platform-independent and HTML pages can be viewed and created on any OS (Operating System).
- HTML is an open standard. The W3C is the main body responsible for HTML specifications.

4.2.5 Disadvantages of HTML

Most of the criticism addressed to HTML ignores some basic design issues related to HTML. Many criticize HTML for offering poor formatting capability. Jim Buyens writes:

> "It bears repeating that, despite its popularity, HTML is one of the worst page description languages in use. Its inventors, in the interest of device dependence, stripped HTML of virtually all capabilities to control page appearance. And page authors, in the interest of visual communication, have tried to regain control ever since." [36] (333)

Such criticism ignores the fact that HTML was not meant to be a formatting or DTP Program like Adobe Page Maker or QuarkXpress. The problem with HTML, from my point of view, is simply its success. One might argue that the success of HTML has revealed its points of weakness, while others might argue that this success required more of HTML than its original design intended to

achieve. The answer to this is more philosophical than anything else. For the most part, the exact syntax of an HTML document is not rigidly enforced by the browser. This often led to abuse of HTML. Charles Goldfarb notes:

> "HTML serves as a useful cautionary tale. It actually has a fairly rigorous structure, defined in SGML, and available from the World Wide Web Consortium. Everybody tends to treat the rules as if they were from the World Wrestling Federation - they ignore them." (42)[37]

To cope with malformed and badly written HTML, the browser makers make provision for badly written HTML documents. I give here one example of this. While HTML has a clear rule for nesting elements, which one could describe as *first in, last out*, many ignore this rule and write a bad code as follows: <B><I> Hello world! </B></I> to display this: ***Hello world!*** The previous code is wrong nesting and it should be written as follows <B><I> Hello world!!</I></B> to avoid overlapping of tags. To avoid crashes of HTML pages or Web browsers, browser makers make provision for such badly written code. This is a very lengthy and expensive process. Browsers become bulky and require more computer resources (hard disk space, processing power and memory). This makes it difficult to display Web pages on smaller devices that do not have the computing resources of desktop computers. With more tags offered in the latest versions of HTML that hardwire content and presentation, changing content or formatting became more difficult especially for Web sites with a huge content that often requires modifications. HTML became difficult to learn as well. A complete HTML reference book would probably require hundreds of pages.

4.2.6 How to Use HTML

- Writing good HTML is perhaps the best way to present emblems on the Internet. Emblem books could be scanned; images could be embedded in HTML files and distributed on CD-ROM.
- Since HTML is the language of the Web, the form could serve as an excellent method for cooperation between institutions in possession of different emblem books. Such cooperation is a precondition due to issues related to copyright. Distributing books in HTML takes advantage of user's skills using his or her browser. No additional learning is required.

4.3 XHTML (eXtensible HyperText Markup Language)

4.3.1 Introduction

As mentioned previously, HTML was at a crossroads. It was clear that the solution was not to add more tags. XHTML was presented as a solution for these problems. XHTML is almost the same as HTML, but designed to work with a new technology, many believe it holds the future of the Web: XML (eXtensible Markup Language). XHTML is a combination of HTML and XML. XHTML consists of all the elements in HTML 4.01 combined with the syntax of XML. Berners-Lee, W3C Director and the creator of HTML describes XHTML and its mission as follows:

> "XHTML 1.0 connects the present Web to the future Web, It provides the bridge to page and site authors for entering the structured data, XML world, while still being able to maintain operability with user agents that support HTML 4."[38]

XHTML aims at realizing two main goals:

Portability: means giving devices with small computing power the ability to access Web pages. This could be realized through enforcing a mechanism of good HTML writing that adheres strictly to the rules.

Extensibility: allows for more creation of tags according to the needs of different institutions. Furthermore, separating content and structure allows frequent modification of vast documents without going through the laborious work of changing the whole content. To do this XHTML recommends using CSS (Cascading Style Sheets), which takes charge of the formatting of different elements on a Web page. CSS follows the same concept used by some word processing applications known as style. The developer of Web pages can concentrate on the content of the Web page and CSS instructs the browser as how to format the different elements of a Web page. This ensures consistency through the Web site. In other words, any HTML document can be converted to XHTML following simple rules of good coding of HTML, strictly following certain conventions related to syntax and case-sensitivity of XHTML, adding the DTD (Document Type Definition) as the first line of each XHTML file and paying attention to some different rules of using tags' attributes.

To facilitate the migration from HTML to XHTML W3C created three different types of DTD for XHTML:

Strict: is the most rigorous standard of XHTML. It normally uses CSS to display and format the different elements of a Web page but does not support frame-related elements.
Transitional: is a more tolerant DTD and takes advantage of HTML presentational tools. It is best used with browsers that do not support CSS. It is also the most used standard of XHTML.
Frameset: is used for Web pages that use frames, which divide the browser's window into two or more windows.

4.3.2 How to Create XHTML Files

XHTML could be considered as a clean version of HTML. Thus any tool used to create HTML could be used to create XHTML as well. When using WYSIWYG editors, some additional cleaning handwork is required, however. Some software products are available and serve as converting tools to XHTML. They ensure also the cleanness of the document. The W3C offers validation service to ensure the conformity of the document with the DTD type declared at the beginning of the file.

4.3.3 Advantages of XHTML

- XHTML can be considered a better HTML thus offering all the advantages of HTML and ensuring portability of Web pages.
- XHTML is a bridge between the language of the Web today, HTML, and the Web language of the future, XML. Web pages written in XHTML will remain compatible with XML browsers in the future.
- Separating content and presentation allows for consistency through the Web site and facilitates changing the content of a Web site.
- XHTML as an application of XML has a native support for Unicode, a universal character set that supports all human languages. This allows for a language rich content and for displaying content in different languages including Hebrew and ancient Greek for instance.

4.3.4 Disadvantages of XHTML

- XHTML is not as forgiving as HTML and strict adherence to XHTML rules is required.
- XHTML recommends using CSS for representation, while helpful in the long run, it requires an additional learning curve.

4.4 XML (eXtensible Markup Language)

4.4.1 Introduction

XML is perhaps one of the most discussed technologies recently. As in such cases, it is sometimes hard to differentiate between myth and reality. I should perhaps start by saying what XML is not. It is not a programming language and it is not compiled. XML is not a software program and it does not do anything on its own. One of the main reasons for confusion, at least for the novice, is the three letter acronym describing it. The word *markup* reminds us of HTML and the word *language* is more misleading than descriptive. One of the basic elements in any language is its inclusion of a set of words and grammar rules specifying how to use these words. This concept applies not only to human languages but also to programming languages. Applying this concept to HTML, one could describe the tags as the words and the rules defining their usage as the grammar. XML virtually does not define any particular set of tags. It provides a standardized framework, or grammar rules according to which different markup languages could be created. This is the reason some describe it as Meta Markup Language, a language that is used to create other markup languages. The markup approach of XML is fundamentally different from HTML. While HTML marks up elements mostly to denote their presentation using formatting tags, XML marks up different elements as to what they *mean*. Thus instead of the following markup: <B>John Smith</B>, XML-based languages would markup the text as to what this "John Smith" means in the context of the document. If he is the author of a book, one could use the following markup <author>John Smith</author>, if in another context he is the owner of a corporation, we could use the following scheme <owner>John Smith</owner>...etc. This way the content and its meaning are preserved. The tag <author>, or <owner> is a descriptive tag that gives us information about the content itself. This is why XML is often described as a *metadata language*. Metadata can be defined as data about data. A third and important feature of XML is the way it handles a document. While HTML offers a limited set of tags to describe the structure of a document, XML documents form a tree that must include a root from which other branches come out in a very strict manner. XML documents usually start with a declaration that would look like this <?XML version="1.0"?> and they have the standard extension ".xml". The browser does not have to recognize the different tags in an XML document, because XML is *not* about *presenting* or *formatting* documents. XML describes *structure* and *semantics*, not formatting. Because XML does not use predefined tags, the browser does not know how to format or display them. The tag <table> could mean an HTML table or maybe a real table. In order to display XML

documents, XML uses a language called XSL (eXtensible Stylesheet Language), which functions much like CSS with XHTML documents. Indeed, CSS could be used to format XML documents as well but it is recommended to use XSL, which is a far more powerful formatting language than CSS. While CSS is merely about presenting or formatting XHTML documents, XSL does more than that. If XML documents are to be displayed in a Web browser, XSL can transform the XML document to HTML document. One might ask: "Do we have to go through all this just to display or to transform XML documents to HTML?" The answer is simple: while HTML documents are and will always be HTML documents, XML documents preserve the logical structure and the meaning of the different elements that could be transferred to other devices or transformed to other formats. This is perhaps the strongest argument for XML, which could be described as the *reusability* of XML documents. With simple modification of XSL, XML documents could be easily used or manipulated as the need arises. To do this XSL consists of different parts to define XML parts and patterns, transforms and formats them.

So far it seems that XML documents need another language, XSL, to be displayed and transformed. But there is one crucial element as well that is necessary for XML: DTD (Document Type Definition). DTD defines the document structure with a list of legal elements that could differ from one XML document to the other and sets the rules for what is allowed in the XML document. While XML sets the general syntax rules of the document, DTD sets the rules for the different elements in the document. The DTD is the first thing the browser reads after the document declaration. It could be included in the document or referred to as an external reference. XML does not force developers to write DTD for each XML document. One could write a well-formed XML document following the strict rules of XML, but it is strongly recommended to write DTD for any serious work with XML. Any XML compatible application usually processes the three different parts mentioned above: XML, XSL and DTD. There are other technologies involved as well that could help XML do more. The Hypertext feature, for instance, in XML is the task of another language known as XLL (eXtensible Link Language), which offers more powerful features than its counterpart in HTML. XLL can be used to link to pictures or other documents.

4.4.2 How To Create XML Documents

XML files can be created using any simple text editor and saving it with the standard file extension ".xml". To check the conformity of XML document with XML rules, parsers are used (parser are software that verifies the structure of the document). There are two kinds of parsers: the *validating* parsers and *non-validating* parsers. Validating parsers check the XML-document

markup against a DTD. Non-validating parsers check to make sure that the document is well formed (a well-formed XML-document is a document that follows strictly the rules of XML without having a DTD).

4.4.3 Advantages of XML

- To start with, XML is not limited to the Internet. XML has a lot to offer to the publishing industry in its various forms as well. The incompatibility between different systems and applications dealing with text and documents made documents prisoner to the program that created them and much was lost due to the lack of software that could open and edit documents. To overcome these difficulties, SGML was invented back in the sixties and it became an international standard since 1986 (SGML -Standard Generalized Markup Language -ISO 8879). Trying to cover as wide an area as possible, SGML ended up being too complicated to be implemented in most publishing houses and its usage was limited to institutions with a large volume of documents. XML could be considered as a small or stripped down version of SGML and it still offers most of the capabilities of SGML. This gives many advantages to XML since SGML is well-established and well-documented standard that has been used with success in the past decades.
- As a subset of SGML, XML documents are valid SGML documents as well. The main approach XML takes towards documents is the separation between three elements of the document: *content*, *presentation,* and *structure*. This offers many advantages. The same content can be manipulated and be displayed in different formats to different devices and platforms. The same content can be formatted for on-screen display or optimized for printing.
- XML documents keep their structure when transferred, and this can be very important for documents where relationships between elements are relevant.
- XML data is text data; that makes it entirely platform-independent. XML is a very convenient format to move data between systems. The same content can be used many years after its creation.
- XML can be used to store data. Some believe that XML will replace databases one day. We must notice, however, that XML is not a database, neither does it offer a replacement to the database. XML can be used to store data in files or in databases. There are products, which facilitate XML integration with databases. Database makers provide more and more support for XML.

- XML is a non-proprietary, public format. It is not a proprietary development of any company that could make any changes to the standard without notice. No specific commercial software is required to open or edit XML documents. The body controlling the XML standard is the W3C (World Wide Web Consortium).
- Compared to SGML, XML is relatively easy to learn.
- XML enjoys great support among the major software firms who are including more and more support for XML in their applications.
- XML supports the Unicode Character Set (ISO/IEC 10646), which makes it ideal for language rich content.

4.4.4 Disadvantages of XML

- XML offers virtually unlimited capabilities but we should not forget that XML does nothing on its own. The advantages of XML depend on the variety of applications supporting it. The number of browsers and applications supporting XML is still somewhat limited.
- While the standard of XML has been published and agreed upon, there are many other standards related to it that are still in the developmental process. Some technologies are likely to replace others as well.
- While Web pages created using HTML could be done by an individual effort, XML and related standards require teamwork to be implemented successfully.
- XML has a very rigid and unforgiving structure and its rules must be followed rigorously. More effort is needed to markup a document using XML than HTML.

Chapter 5

Digitizing Emblems

5.1 Introduction

Digitizing emblem literature requires a combination of hardware and software. Besides the basic component of a modern computer system, digitizing emblems mainly requires a scanner and some added software. Scanners are optical devices that capture printed matter, text or images, and transform them into a digital signal that can be manipulated by computers using different software such as OCR (Optical Character Recognition) software or graphics software. Like all input devices, a scanner needs a driver, which is a device-specific software that allows a computer to recognize it and work with it. Since scanners were designed to work with different applications, a common interface was designed to allow software applications and image-capturing devices, scanners or digital cameras, known as TWAIN (Technology Without An Interesting Name). Virtually all scanners come with a TWAIN driver, which is recognized by most software applications. Besides those two basic software components, device driver and twain drivers, most, if not all, scanners come with software that is capable of handling basic OCR and image editing. However, the capabilities of such products remain limited. For more serious work, it is recommended to use more advanced software. This software is usually sold separately. Due to the fact that emblem books are old and some emblems may be in poor condition, some digital enhancement of the content might be necessary. The decision regarding the software to be used depends upon price, capabilities and licensing issues.

In digitizing emblems one should differentiate between two main approaches:
Digitizing the entire content as a picture: this offers the advantage of reproducing an exact picture of the entire content. This approach, however, has its disadvantages as well. It requires huge computer resources and the textual content will not be searchable. Some fine details, such as text in small size, might not be properly presented. This problem can be resolved by scanning the image in high resolution, which will increase the file size even more.
Digitizing the textual part(s) and the picture(s) separately: Theoretically, this allows text to be searched and substantially reduces the file size. One could take advantages of modern OCR software capabilities in recognizing texts as well. Modern OCR software can be trained to recognize new forms of text and fonts, and they achieve a high accuracy rate. Some basic problems related to the language such as the problem of inflected and conjugated word forms remain, however. I will discuss linguistic problems later in this chapter. Old texts, in particular texts in ancient Greek, might pose a serious challenge because they were not written by modern

computer software using standard fonts and because of inconsistencies in writing. We should also notice that there was no real standard in writing emblem texts. Too many irregularities in writing forms, using fonts, irregular spacing between words and often misspelled words as well as the poor condition of emblem books available today make recognizing emblem texts a very difficult task for modern OCR. We must be cautious, however, not to separate all the textual part(s) from the pictorial part(s). Texts written on imprese, heraldries and coats of arms, for instance, are an integral part of the image and should not be thought of as detachable text, even if we use text-tools in editing software to enhance them if necessary. Emblems on the other hand, offer a combination of text and image. The separation in the process of *digitizing* does not necessarily mean separation in the process of *presentation*. As I mentioned previously, the different parts of an emblem constitute a unit and, in my view, they should not be separated.

Another important decision that must be made during the scanning process, is the image format. This decision will depend upon the purpose of digitization. If the purpose is to reprint emblem books, one should choose a format that reflects high quality in print and does not compress the image such as TIFF (Tagged Image File Format) or other software specific formats. The downside of such a format is the huge file size, which makes them entirely inadequate for the Internet. In order to make image files small enough to fit the Internet, files must be compressed. Different compression schemes offer a variety of benefits and drawbacks. A graphic compression scheme that does not change the image or cause it to lose any quality is called lossless, while a scheme that results in a change in the image data or a loss of quality is called lossy. Lossy compression is mostly noticed when the image is printed since printers use a much higher resolution than computer screen. On the screen, however, such loss in quality may not be perceived by the human eye. Thus a TIFF or JPEG format, for instance, might look the same on the monitor, but not in print. I present here in brief the most common formats available for the Internet.

5.2 Graphic Formats

GIF (Graphics Interchange Format) is a proprietary graphic format developed by CompuServe. Although it is common, the use of this format could pose problems related to copyright. Theoretically, CompuServe could ask developers of software or Web sites using GIF format to pay a royalty to CompuServe. GIF uses a lossless compression scheme, supports up to 256 (8 bit) colors, allows for simple animations and supports both interlacing and transparency.

JPEG (Joint Photographic Experts Group) is a lossy compression scheme that supports 16.7 million (24 bit) colors. Standard JPEGs do not support transparency or interlacing. However,

the progressive JPEG format does allow interlacing of JPEGs. JPEGs are best used with color-rich images.

PNG (Portable Network Graphics) was developed in response to legal problems over the GIF format. The PNG format is a lossless compression scheme that allows either an indexed 256 color image or a 16.7 million color image. PNG also supports interlacing, 256 levels of transparency, and a variety of other options offering thus many advantages of both GIF and JPEG. PNG can be used with any graphic. Unfortunately, it is not supported by all browsers. This is changing, however, and I expect this format to be dominant in the near future.

The fact that print-oriented formats are different from Internet-oriented formats does not necessarily mean that we must digitize twice. In fact, most if not all, graphics software can transform print format into JPEG, GIF or PNG format. We must notice, however, that this process is a one-way-street. Print formats can be compressed, but the contrary is not possible. It would be wise in such cases to transform a copy of the digitized format, and keep the original for possible printing later on. I also recommend using different applications in the scanning processing. For instance, we could use the capture capability of Adobe Acrobat to create a PDF directly from the scanner, and use PhotoShop to create a print version. While this might prolong the scanning process, we should recall that it is always easier to make two or several copies of the printed matter available, than making one copy, only to discover later that we have to copy or digitize again. The fragility of emblem books available today should be a concern as well. Making several digitized forms of the book while it is available on the scanner is always better than having to go through the entire process again, which could possibly cause damage to the book.

5.3 Indexing Emblems

5.3.1 Introduction

It is evident that the mere scanning of emblems is not enough to make them accessible to scholars and students. Gaining access to a collection of emblems is problematic when they are not indexed, or translated into a textual form. To make emblems accessible to scholarly research we must find a way to index and catalogue them. In fact many use the term *digitization* for the entire process of scanning and indexing and cataloguing the content. Indexing and cataloguing visual material is one of the major topics that occupy scholars of art, literature and several fields of computer science, as Michael Lew notes:

> "Visual information retrieval lies at the crossroads of multiple disciplines such as databases, artificial intelligence, image processing, statistics, computer vision, high performance computing, and human-computer intelligent interaction."(3)[39]

I will deal here briefly with the problem of cataloguing, indexing and retrieving visual media in general and with emblems in particular. I will also discuss problems related to the indexing and retrieving of text. At first sight this might not seem relevant to retrieving visual matters, but indeed it is, since the main tool we use in indexing and cataloguing images is usually text. We should also remember that emblems include a textual component, which needs to be considered in any indexing and cataloguing system. Information about emblems themselves or metadata related to emblems is also written in text format.

5.3.2 The Unique Problem of Visual Information Retrieval

In dealing with visual media in general we use text-based communication method or words. Indexing a collection of digitized emblems or visual media is made through describing the content using words. Indexing and cataloguing scanned images will have two main purposes: to maintain control over a collection, and to facilitate research. Several indexing and cataloguing systems are available today. Earlier systems in the field used words as the main indexing and retrieval mechanism. More recent systems use complex mathematical coding of different parts of the image, especially coding of colors in their different variations and contrasts and different textures in the image to index and retrieve images. Due to the absence of these elements in most printed emblems, I will restrict my discussion here to the somewhat older method of indexing images through text description or textual data related to the image. In indexing visual materials

using words there are basically two main approaches: using a set of *controlled vocabulary*, and using *free text*. I will briefly outline both approaches and their advantages and disadvantages.

Using a set of *controlled vocabulary* limits the description to a set of words previously provided for the indexer. This is known as *keywording*. Keywording is one of the most important aspects in the process of indexing visual materials and constitutes the traditional approach to indexing images. The main advantage of using keywords is that the controlled vocabulary makes indexing more precise. Further, it is easier to search on a keyworded database, because the searcher can choose from a finite list of keywords to frame his or her search. On the other hand, it takes longer to add terms using this system, because the indexer has to look at a long list of keywords and thesaurus entries for each picture. This can be a lengthy process especially with systems offering a huge list of controlled words. On the other hand, the indexing itself is likely to be less descriptive because the indexer only uses certain words to describe the pictures(s) to be indexed. Also, it takes longer to search, because searchers need to consult a thesaurus and a keyword list. We must notice also that some of these systems, such as Iconclass, use a set of controlled words including a combination of computer-based digits and human words. This procedure aims at enlarging the list of the controlled words and at the same time ensuring the uniqueness of each word provided in the indexing list. However, this renders a complex process even more complex.

Using *free text* indexing, on the other hand, does not rely on a controlled vocabulary. Instead, the indexer uses his or her imagination, knowledge, and linguistic skills to describe the content. Using *free text* is a much easier and quicker process to index images and visual material(s). The indexer has the freedom to associate all themes or items in a picture and to discover the meaning of the content. Free text indexing is easy to index and allows a great deal of flexibility when assigning terms. As a result, the indexing is likely to be more accurate, because terms are unique to each image. However, because of the lack of vocabulary control, indexers could use synonyms to describe similar images. There is a greater chance to miss images if descriptions are not applied consistently. However, we should differentiate here between two different kinds of description: describing the visual item(s) in a picture, and describing the meaning of the picture as a whole. Both methods, *controlled vocabulary* and *free text*, reveal basic problems in dealing with the meaning of the visual element(s) presented. Thus, indexing the subject of an image becomes a combination of a linguistic descriptive effort as well as an analytic, intellectual process. The extent to which the indexer analyses the concept of a work, or what it means, and his or her translation of the meaning into descriptive terms, will contribute to searcher retrieval. For instance, one could describe a picture of people gathering outdoors as follows:

people, table, chairs, food, drinks, bottles, garden and so on. One could also use words describing the meaning of the picture such as festivity, celebration, party, summer and so on. The description of the item(s) included in the visual work is likely to be more accurate than the description of what the picture is about or what it means. The latter is likely to reflect the subjective perception of the indexer. There is, for instance, a great difference between describing an activity such as running or jogging. In addition to this basic problem of describing the meaning of some picture, there are some specific problems related to the nature of emblems, which I will outline here.

5.3.3 Problems in Indexing Emblem Pictures

Should everything in an emblem picture be described and indexed? Can a useful distinction be made between primary and secondary motifs in pictures? Is it sufficient to describe only symbolic motif? The answer to these questions should take the following into consideration:

- Describing the items individually does not necessarily describe the relationship between the items, which might be of a crucial value in elucidating the meaning of the image. This is particularly important with emblems, where the entire picture is more than the addition of the individual items. For instance, a snake embracing a woman or a man peacefully has a different meaning from a snake biting a man or a woman. A picture description, which only lists a string of motifs, is insufficient.
- Unlike many pictures, emblems do not always represent objects as they are in reality, neither do they offer abstract notions as many works of arts do. Items in emblems are often symbolic representation of something else. A woman may be a Greek Goddess, a virgin, a princess or something else, although the textual components will often indicate what is meant. Each of these terms has its implications and differs greatly from the others, although the visual presentation may be identical. Using any of these terms is likely to depend upon the interpretation of the indexer of the emblem and will influence both the search process as well as the perception of the reader.
- We should also remember that emblem writers and artists had limited technical capabilities and in some cases lacked knowledge of some creatures in their emblems. Is the animal pictured a leopard or a tiger, for instance? The text might offer some clarification, but this is not always the case. Using general terms such as "animal" may appear to offer a simple solution, but will cause later problems in searching, since too many results are likely to occur.
- Emblems were produced in the pre-Linnear period. Daniel S. Russel comments on the difficulty in naming such creatures as halcyon and pelican[40]. We should not index each

and every detail in the image. The likelihood of a higher recall is increased with the increased amount of text describing a picture, however, precision is likely to decrease. While this applies generally to all pictures, we should be careful that some minor details in emblems might have a great value to the meaning of emblems. How does one judge the importance of such elements? The relative size of any specific item does not necessarily relate to its importance for understanding emblems.

- While it might be necessary to index some elements in some emblems, the same elements might be of no particular value in other emblems. For instance, in many emblems the presentation of clouds is a mere background. In other emblems, however, one could argue that the presentation of such an element is not casual but contributes to the meaning of the emblem as a whole, and should be part of the indexing process.
- Pictures cannot be "read" in isolation, i.e. ignoring the inscriptio and subscriptio. It would make no sense to pretend the texts do not exist. A Greek hero or a Roman God may well be identified in the text and this knowledge should inform the attempt to describe pictures.

Once emblem pictures are indexed, we need to deal with problems related to indexing text itself to make it easier to retrieve.

5.3.4 The Problem of Text Indexing

Text retrieval may appear deceptively easy compared with searching images. Scholars are becoming more aware of the difficulty in recent years due to the unprecedented growth of the Web and the increased amount of electronic information available. We are surrounded with large quantities of information, but unable to use that information effectively, because of its tremendous size. Web search engines are the most common solution to this problem. Here text search is done using different search schemes, some of which are as simple searching for a word or a combination of words using Boolean operators; others offer more advanced search mechanism(s). Often this is not as effective as users wish. I have discussed this problem partially in chapter 2 and attributed this problem largely to using HTML as a markup language for Web documents and indicated that HTML pays no attention to the meaning of the text or the context in which similar or identical terms are mentioned. I have also suggested that using XML as a possible future markup language for the Web will substantially help solve this problem. Yet, there are certain inherent problems in indexing and searching text. One of the main problems is the presentation of different characters available in different languages.

5.3.5 From ASCII to Unicode

Fundamentally, computers understand only two digits: 0 and 1. In order to present different characters and letters, we need to present them as series of 0s and 1s. Different companies developed different models, which were not compatible. To solve this problem the ASCII (American Standard Code for Information Interchange) code was developed. ASCII is the basis of character sets used in almost all computers today. ASCII is capable of presenting 128 different characters including white space, numbers, most basic punctuation, and unaccented letters a-z and A-Z. This leaves us with the problem of presenting and retrieving characters not available in the ASCII model. To overcome this problem many ASCII extension encoding schemes were developed, but they were not compatible with each other and data was susceptible to corruption as it traveled among different systems. As a solution to this problem, the Unicode Standard was developed by the Unicode Consortium, a non-profit organization. The Unicode Consortium presents the advantage of the Unicode Standard as follows:

> "Unicode enables a single software product or a single Website to be targeted across multiple platforms, languages and countries without re-engineering. It allows data to be transported through many different systems without corruption."[41]

We must note, however, that Unicode went through different revisions including different levels of support for different characters. Not all software products, which claim to support the Unicode Standard, support the latest version. The original goal was to use a single 16-bit encoding, which allows the presentation of more than 65,000 characters. The latest version (as of this writing) is Unicode Version 3.2.0, which provides codes for 95,221 characters from the world's alphabets, ideograph sets, and symbol collections and has about 6,700 unused code points for future expansion in the BMP (Basic Multilingual Plane). Also, there are over 870,000 unused supplementary code points on the other planes. More characters are under consideration for addition to future versions of the standard. Major software manufacturers have adopted the Unicode Standard and are implementing it in their applications. For accurate text retrieval on the Internet, ideally we need to have the different applications implied in the delivering mechanism supporting the Unicode Standard, notably the server, the indexing software, the Database and the user's browser. Major browsers as well as Web servers now offer support for the Unicode Standard. However, not all database systems or indexing systems fully support Unicode. In the near future this should be less of a problem as support for Unicode becomes more available. I fully agree with Markus Kuhn's assessment that:

> "Unicode is well on the way to replace ASCII and Latin-1 in a few years at all levels. It allows you not only to handle text in practically any script and language used on this planet, it also provides you with a comprehensive set of mathematical and technical symbols that will simplify scientific information exchange." [42]

To retrieve text from a CD-ROM, ideally the database and the indexing mechanism of the database included on the CD-ROM, as well as the OS should support Unicode Standard. Currently not all desktop OS fully support Unicode. The most used Desktop OSs supporting Unicode in use today are the Apple Mac OS 9.2, Mac OS X 10.1, Microsoft Windows NT, Windows 2000 and Windows XP as well as the latest versions of Linux.

5.3.6 Indexing Text

Resolving problems related to the representation of different character sets relevant to European languages is essential in dealing with textual data related to emblems, but it is just the first step in indexing text for later retrieval. To index words an IRS (Information Retrieval System) needs to recognize what a word is. Recognizing word boundaries is easier in English than in other languages. In English words are separated by white spaces or a punctuation mark. In other languages such as German, we have the notion of compounds, where compounds are not separated by spaces. The second question we should ask is what should be indexed? Words or phrases? While it is possible to search for a certain phrase, we should differentiate here between phrase indexing and retrieving a phrase as a string of letters. The latter does not constitute phrase indexing, since it is only capable of recognizing the textual string, not the context. For instance we can retrieve *president of the U.S.*, but we might fail in retrieving the same phrase after a slight change such as *presidents of the U.S.* Recognizing phrases is not trivial and requires some intelligent analysis of the content and some awareness of possible variations of the phrase. Using words for indexing is the most common approach.

5.3.7 Noise Words

The question about what to index is relevant to the question, what not to index. Some words are hardly content-bearing such as articles, pronouns, prepositions and conjunctions. These words are called *stop words* or *noise words* and they should not be indexed. We should also notice that *noise words* are context sensitive. For instance a word like computer has a meaning and it is usually indexed, but in a computer book it is so often repeated in too many different contexts that it becomes a noise word and should not be indexed. This will save space and make the indexing mechanism faster and more effective.

5.3.8 Stemming and Lemmatization

To make the search mechanism even more effective two main indexing concepts were developed: *stemming* and *lemmatization*. The two concepts are closely related and the two terms are often used interchangeably. This is not accurate, however. I will discuss both concepts here and outline their advantages and disadvantages.

Stemming

In its simplest definition, *stemming* provides us with the stem of a word. A stemmer has a set of rules, which it uses to remove suffixes, and other endings of words such as "tion", "nal", "ly" and so on. For instance, a simple stemmer does a simple plural removal as in plays (play), studies (stud) and endings such as "ing" as in playing (play), "ful" as in handful (hand), and "ness" as in seriousness (serious). Further it can make more intelligent decisions related to some forms of plural that search for *ies*, but not *eies* or *aies*:, or *es*, but not *aes*, *ees*, *oes*; and s, but not *us*, *ss*. Stemming can also reduce a verb by eliminating added letters in conjugated forms such as played (play). The stem is then supplied with the wild card * that would allow retrieval of all words including the stem, for instance, we would be able to retrieve "braves", "bravery", "braveness" just by using the stem brav. Stemming is an effective tool in indexing and retrieving information. We must notice, however, that it has its limitations and it can cause many erroneous results.

Problems with Stemming

- Stemming can result in non-words such as lovi (stem lov).
- Stemming is problematic with irregular forms such as "feet", "women", "teeth", etc.
- Some unrelated words share the same stem such as "police" and "policy" (stem polic).
- Stemming shows different results in different languages. This is due to the nature of each language. English is perhaps the easiest language, where stemming is most effective.
- Stemming is problematic with compound tenses and various forms of conjugated verbs.
- Most stemmers do not consult a lexicon or check the dictionary for the correct word.
- Stemming is problematic with compounds as in the case of many German words such as "Liebesgedicht" (love poem). This is less of a problem in English since almost all compounds such as Computer Science, Trade Agreement, are separated by a white space. There are exceptions for this such as "database" or "bookshelf", but in German and other European languages compounds are very common.

- Stemming is problematic with idiomatic expressions, where words could have different meanings according to the context.

Lemmatization

M. Boot defines lemmatization as follows:

"[Lemmatization is] *generally defined as the transformation of all inflected word forms contained in a text to their dictionary look-up form*"[43]. According to this definition, to lemmatize a text is to reduce each word in that text from its inflectional and variant forms to its base form. Lemmatization thus covers a wider area than stemming and offers several advantages over stemming.

Advantages of Lemmatization

- Lemmatization always checks a lexicon or a dictionary to prevent problems that occur as a result of different words with different semantic meanings sharing the same stem such as "universe" and "university", stem (univers). This also solves the problem of non-words, which occur with stemming.
- Lemmatization deals better with compound forms often available in German and other languages. Lemmatization is often combined with part-of-speech tagging to avoid those mistakes. In compounds such as "Regierungsvertrag" (government contract), for instance, the main word is "Vertrag" (contract) and the lemma is provided for this word.
- Lemmatization deals with conjugated forms of nouns in languages where this occurs.
- Lemmatization deals effectively with words, which can be used as verbs or nouns by tagging the word(s) as a verb or as a noun according to its specific usage in the sentence.
- Lemmatization is capable of dealing with different inflectional forms of verbs including compound tenses and with irregular verbs as well.

Lemmatization is a lengthy and repetitive task in which computers could effectively help. There are several software products known as lemmatizer(s) available for several languages. These products are usually language specific. Programs such as *Eyeball* and the *Micro English Parser* for English, *LexiTex* and *Lemmatiseur* for French, and Lemma2 for German, are a few examples of programs in this category. Due to the sensitivity of the task, manual editing is strongly recommended to avoid bad software assumptions. Manual lemmatization is also possible and is considered to be more accurate, but it is usually a very lengthy and costly process.

Problems with Lemmatization

- Idiomatic expressions are problematic with lemmatization. The same word could have two different meaning depending upon the context. For instance the word "play" could have two distinct meaning such as "play" (as a theatrical performance of a drama), and "play", (as a game).
- Lemmatization is computationally more expensive since it requires a lexical look-up.
- Tagging of parts of speech could prove to be erroneous, especially with words where the compound has a distinctive meaning from the words it includes. For instance, in German the word "Bahnhof" is a compound of "Bahn" and "Hof", two words, which are semantically not related.
- Texts in emblems do not correspond to the standard of language today, which makes any lemmatization by means of software very difficult.
- Emblem texts, and the metadata related to them are very likely to include several languages, which could prove to be a very challenging task for any lemmatization products, which are language specific.

5.3.9 Conclusion

The interdisciplinary nature of emblems becomes more evident as we digitize them. Emblems share common aspects with both textual data and visual media. Although there are several indexing and cataloguing systems for visual media available today, none of them was specifically developed to deal with emblems and their unique nature. Emblem texts present a challenge also. They were written in several languages and periods, prior to the development of these languages, as we know them today. Offering a normalized version of emblems will greatly help in solving the problem of indexing and cataloguing emblems.

5.4 CD-ROM

Once emblems are available in digital form, we can publish them on the Internet or on CD-ROMs. I discussed previously the Internet in some detail. I will elaborate here on CD-ROM technology.

5.4.1 Introduction

One of the most important developments in digital publishing is the invention of CD-ROM (Compact Disc Read-Only Memory). The term "disc" is used rather than "disk" to differentiate between magnetic disks and optical discs. CD-ROM is a non-volatile optical data storage medium using the same physical format as audio compact discs, readable by a computer with a CD-ROM drive. As the name implies, CD-ROMs use compact discs. In fact, they have the same physical disc format as the ones used for music. The story of developing CD-ROM began in the music industry as Sony and Philips introduced it in 1982 as a format of recording music known as Compact Disk or simply CD. CDs offered several advantages for the music industry over the classical tape used for distributing music recordings. CD's features such as durability, random access features and audio quality made it incredibly successful, capturing the majority of the market within a few years. Since it was not possible for the general public to duplicate CDs, it offered the best protection for copyright. Soon after the compact disc became popular for storing audio signals on optical material, the benefits of storing computer information in this manner became apparent. CD-ROM followed in 1984, but it took a few years longer to gain the widespread acceptance enjoyed by the audio CD. This was due to several factors related to conflicting standards, difficulties of hardware configuration, the lack of support by OSs and the high price of the technology. Also the lack of suitable media that could take advantages of such a medium for the majority of computer users was an obstacle for its widespread use. In the eighties and the early nineties most people used computers for text-based applications that did not need the capabilities of CD-ROM. As computer capabilities evolved, computers became more than just word-processing machines. Rich multimedia content that took advantages of the huge capacity of CD-ROMs was developed. As computer applications, data and database grew in size, the advantages of CD-ROMs became obvious for both users and software manufacturers alike. It was cheaper and easier for software manufacturers to distribute their applications on CD-ROMs. CD-ROM offered an ideal solution at this time for copyright protection against illegal copying of software since it was not possible to duplicate CDs. For users, it offered them a convenient way for installing huge software applications and a reliable medium for keeping software without being worried about the degradation often occurs to floppy disks. It gave users as well the

possibility of making use of multimedia capabilities and allowed users to use computers for more than word-processing. In fact, the "multimedia revolution" was largely a result of the availability of cheap CD-ROM drives. While many use the term multimedia in relationship with music or computer games, we should not forget that the multimedia revolution has allowed as well for a revolution in the way we learn. As the public interest in CD-ROM grew, prices kept on falling, and the problem of conflicting standards virtually no longer exists. CD-ROM quickly became a standard part of modern computers and an ideal publication medium for encyclopedia, software applications and databases. With the huge success of the new medium, hardware manufacturers developed CD-Writers, which allowed users to duplicate their CD-ROMs and use CD-ROMs as backup media. At the beginning this was not much of a concern since the prices of the raw CD-ROMs and that of CD-Writers were relatively high. As prices of media and CD-Writers kept on falling, the once secured medium against duplication became the ideal medium for duplication. This has a strong effect on the software industry as well as the music industry, which had developed this medium in the first place. The software industry claims to lose 11 billion US$ through illegal copying of software[44] and the music industry claims a loss of 50% of its revenue or as much as 4.3 billion US$[45]. I must add, however, that such estimations are based on the assumption that users who take advantages of pirated software and music would have bought the real product if they did not have the pirated copy. Such an assumption is entirely false. It is very unlikely that a student or a citizen in a developing country who buys a pirated copy of an expensive software product would be willing or capable of buying the much-inflated prices of most software products. However, to protect their intellectual property, music and software manufacturers are developing different methods of copyright protections.

CD-ROMs are a huge topic in the computer world today. Without getting too technical about the interior mechanism and the functionality of CD-ROM and CD-Drives, I will present the basics of this medium. CD-Drives have evolved and beside the classical CD-Reader we have two different kinds of CD-Writers that are capable of reading and writing CDs: the first is classified as Write Once, Read Many (WORM) drive. Often the acronym CD-R (for CD-Writer) is used to describe such drives. With WORM drives, after information is stored on a disk, the data cannot be changed or deleted. The other is the CD-ReWriter, which allows for writing and modifying the content of the CD. The latter is a recent development and we must notice that CD-RW media cannot be read by CD-ROM drives built prior to 1997 due to the reduced reflectivity (15% compared to 70%) of CD-RW media. Most CD-Writers in use today support writing and re-writing of the CD-RW. Blank media can be divided accordingly into two categories: WORM blank media known as CD-R (CD-Recordable) and Re-Writable CDs known as CD-RW.

Standard pre-recorded CDs have their information permanently stamped into an aluminum reflecting layer. CD-RW discs have a phase-change recording layer and an additional silver (aluminum) reflecting layer, which allows several writing and modification of the content.

As far as digital publishing is concerned, CD-ROM will likely remain as one of the most important media in publishing in the near future. This is due to the several advantages it offers which can be summarized as follows:

5.4.2 Advantages of CD-ROM

- CD-ROM is an ideal storage medium that offers the best price-ratio for any storage medium. A single CD-ROM can store up to 650 or 700 MB, enough to store more than 300,000 text pages. This can be of great advantage in digitizing emblems, which need to store and distribute high volumes of data including pictures and text.
- CD-ROMs are almost readable on any computer since virtually all computers today are supplied with CD-ROM drives.
- CD-ROMs offer high access speed and high transfer data. Recent advances in technology have also improved their performance to levels approaching those of hard disks in many respects.
- CD-R is considered a helpful technology for archival purposes as the data cannot be accidentally modified or tampered with.
- CD-ROMs are portable; there is no need for an Internet connection to access the content.
- CD-ROMs can be easily duplicated (after paying attention to the copyright) and shared among institutions or individuals working on digitizing emblems.
- The storage capabilities of CD-ROM and the high transfer rate make it possible to store images in high resolutions that is impossible or at least not advisable to transfer over the Internet.

There are, however, some drawbacks with CD-ROM.

5.4.3 Disadvantages of CD-ROM

- CD-ROMs are not easily updated. Data presented on CDs can only be updated through an upgrade of the product. Some manufacturers present constant updates through the Internet. This does not update the content saved on the CD itself, however.
- CD-ROM content is, in most cases, platform-dependant. CD-ROMs usually store more than information. In most cases information is presented within an indexing and retrieval

system that provides search mechanism for the content presented. Developing content for a specific platform or OS is not an evil per se. It offers users and developers a common ground and the developer can count on certain skills acquired by the user in his or her dealing with the OS. On the other hand, it has the disadvantage of depriving users of other OSs or platforms from gaining access to such a program. Developers must re-write the program if they wish to make it available for another OS or platform and in many cases this could prove to be very difficult if the software was written in a programming language that is specific for a certain OS. Although Microsoft Corporation enjoys a dominant position in the OS market with approximately 90-95% of desktop users using its Windows OSs, we must notice that there are certain incompatibilities between the different versions of Windows. This puts users in a difficult position requiring them to constantly update their software each time they update their OS or remain stuck with their old OS to keep their product in use. If the user chooses the latter, he or she is likely to miss new products developed for the new OS. Although one can install and use different OS on the same computer, this option requires advanced computing skills and poses serious technical challenges as well. Several issues of compatibility between different OSs and hardware arise. Hardware requirements differ substantially among different OSs as well.

- The storage capacity, while more than enough for most users, is still too limited to include certain applications and databases.

In fact, this problem was the main reason behind developing the emerging technology known as DVD (Digital Versatile Disc or Digital Video Disc). DVDs can be considered as an advanced form of CD-ROMs. DVD drives are backward-compatible with CD-ROMs. This means that DVD players can play old CD-ROMs, and video CDs, as well as new DVD-ROMs. Newer DVD players can also read CD-R disks. The main advantage of using DVD over CD-ROM is the huge storage capacity offered by DVD, which can hold up to 17GB of data. This capacity is expected to extend to 29 GB in the near future and perhaps beyond that. The DVD specification supports access rates of 600KBps to 1.3 MBps. While DVD was originally developed as a medium to distribute DVD-movies, more applications are now distributed in DVD format. DVD drives have become more affordable and it is possible now to duplicate DVD using DVD-Writers. This poses serious questions concerning copyright and possible duplication of DVDs. Powerful interests of Hollywood are starting to show their muscles to develop legal restrictions on the technical development of this medium. Indeed, the story of DVD is almost a repetition of the story of CD-ROM in terms of price, conflicting standards, higher storage capacity and copyright protection. It

will be interesting to see how different commercial interests, copyright and legal issues and technical development come into conflict and how society will react to this.

Chapter 6

Emblem literature On the Internet and CD-ROM

6.1 The Macintosh Emblem Project

In her essay about Spanish emblems Sagrario López Poza writes:

> "The peculiar structure and contents of Spanish emblem books require that those who approach their study have a knowledge of art history, Spanish philology, Latin philology, history, rhetoric and a certain command of the erudition of the Spanish Golden Age." (347)[46]

Although López Poza restricts her comments to Spanish emblems, these remarks reveal the many-sided nature of emblems in general, which require deep knowledge in various disciplines to be fully understood. This many-sided nature of emblems had the unfortunate effect that in dealing with emblems, literary scholars have dedicated more attention to the textual parts of the emblem and less attention to the image. In his essay on French emblems, David Graham criticizes the approach of Alison Saunders and her questioning of the value of the image in emblems and its necessity for understanding emblems. He states:

> "... the simultaneous presence on the page of figure and text is a fact, and it poses problems of analysis which must be met head on and which have nothing to do with the esthetics of a given emblem." (1)[47]

Indeed, the importance of the image should never be underestimated. This importance was underlined by Couton whom David Graham quotes:

> "nous avons beaucoup trop oublié que ce sont d'abord des livres d'images" [we have been altogether too forgetful that these are picture books first of all]" (4-5).

Yet, in his attempt in the early nineties to digitize emblems in the Macintosh Emblem Project[48] we notice more interest in the text than the image. David Graham, in order to facilitate the search and classification of emblems, separated the texts and pictorial parts of the digitized emblems. This was meant to achieve two main goals. Firstly, to make emblems searchable through text. Secondly, to save computer resources due to the large file size pictures require. This was a necessity in the light of the limited capabilities of desktop computers in the early nineties. This approach, while helpful in achieving the goals desired, is not perfect and we should pay attention to its drawbacks. Searching emblems through text search could cause one of two situations, which

we would describe as a *false hit* or a *false miss*. A false hit would happen if the text string or word is encountered, but not in the context searched for. For instance, we could encounter the word "love" in several emblems, but the context of love might differ greatly. If the researcher is interested in finding emblems related to the topic "love", in the sense of love between men and women, he or she is likely to find the word "love" in other context such as love of power, love of wisdom and so on. A false miss would occur if the topic or the context being searched for does not include the word "love". For instance, some emblems might have love as the main topic without referring to it using the word *love* in the text. Furthermore, this method creates a false separation between the text and the picture. Emblems have their own method of communication, which includes both text and pictures. Both entities are closely integrated together and their combination constitutes a unit that cannot be easily separated without disturbing the way the message is being perceived. The reader's experience of emblems is a visual as well as a textual one. The experience resulting from this separation could end up being a false one. We should also remember that the same symbol could have different meanings according to the context. For instance, snakes were often used as a symbol of wisdom in some emblems, but in other emblems they were used as a sign of danger. In some cases, it is easy to make the distinction; in others this is less obvious. In my own interpretation of the unicorn emblem in chapter 1 I assumed that the snakes in the emblem were going away from the peaceful scene. This reflects perhaps my own fear and perception of snakes as life threatening creatures, an estimation many modern readers are likely to share. However, I would not exclude the possibility of including the snakes in the emblem as being a part of the peaceful scene to demonstrate that they no longer present any danger for the young maid. Not only symbols are susceptible to interpretations, but the text as well. As Daly notes:

> "It goes without saying that later readers and writers, and indeed whole communities, can receive and change texts in ways that no longer correspond to the original intentions of the text or its maker." (394)[49]

It is obvious that searching emblems using text alone, while effective, does not address such questions.

6.2 Glasgow University

Glasgow University offers an emblem Website dedicated to French emblems. The Website is accessed through the URL: http://www.lib.gla.ac.uk/emblems/index.html. The editors are Alison Adams and David Graham. E-mail contact information to the editors is provided. Very little information is offered about the editors and their credentials. Interestingly, there is a link

leading to the home page of the technical assistant developing the site but no adequate information about the editors.

The site was updated several times in the course of my research but no substantial changes to the content in terms of addition or deletion were made. The site was last updated (as of this writing) on 09-05-2002. The source of the information presented on the site is adequately documented.

6.2.1 Content

The following editions are available:

Alciato/Lefevre *Livret des emblemes*, Paris: Chrétien Wechel, 1536, offering 113 emblems. The table of contents shows two main references schemes in naming emblems: the L scheme for the scanned *Alciato/Lefevre* edition in the site and the D reference referring to Daly's Index "Emblematicus: Andreas Alciatus", which follows the Padua 1621edition of Tozzi. Due to this fact, there is an inevitable discrepancy between the two numbering schemes of the two references. Emblems were sorted according to the L scheme, and all links work properly with the exception of L001 D001, which actually links to the L003 D011.

Alciato/Aneau *Les emblemes*, Lyons: Macé Bonhomme, 1549, offering 201 emblems. There are two references in numbering emblems: the A scheme for the scanned Alciato/Aneau edition and the D reference. In this table of contents there are some unusual discrepancies between the two numbering schemes. It is not clear if all the emblems from Daly's index were included. If that was the intention, then six emblems seem to be missing. These emblems are D026, D037, D047, D097, D147, D170 and D170. If the intention was to include all emblems in the Aneau edition, then there are four emblems, which seem to be missing. These are the emblems that correspond to Daly's D061, D062, D066 and D101. The titles of these emblems are mentioned in Latin and French. However, no link is provided and no A reference for these emblems, which suggests that it is not just a missing link in the Web page.

The text is kept in the original language, early modern French. No translation or normalization to the recent language standard is provided. Editor's intervention in dealing with the text is kept to a minimum and only a few corrections were made. However, I have noticed some inconsistencies in writing some words. I am not sure if these differences were in the original text, or just simple editing mistakes. The word *ainsi*, for instance, was written *ainsy* as in L040 and in other cases *ains* as in L027 and L023. The word *Dieu* is sometimes written in small caps *dieu* as in L018 and L053, and in one case written as *Deu* as in L010. The word *eaue*, sometimes written *eau* as in L038. The word *Lhas* in L039 does not seem to mean anything.

Very few links to other Web sites are presented. This is somewhat strange since more links are actually provided by other Web sites or pages from the same university offering important information about current digitization projects by the Glasgow University. It would have been a good idea at least to link to those pages.

No copyright notice is included. We should note, however, that not including the copyright notice does not render the effort of digitizing and presenting emblems unprotected. A copyright notice might help if legal issues arise.

Since the site is dedicated to French emblems, it is presumed that users will be knowledgeable in French. Not providing some normalized text, next to the original, might pose some challenge for scholars or students less knowledgeable in sixteenth-century French.

6.2.2 Design and Presentation

Presentation of Text

The motto is presented in French and Latin. The Alciato/Aneau 1549 version is presented with the dedication but no dedication is available in the Alciato/Lefevre version. The overall design is simple, easy to navigate and consistent. There are some shortcomings, however. The display of textual content is not optimal. The motto is disproportional with the rest of the text. This is due to the fact that the developer chose a heading level (<h2>) to display the motto. This could be solved by choosing a font attribute one size larger than the rest of the text, or using formatting tags such as or <b>. In general, the problem of displaying text is largely due to the fact that the developer mixes the semantic approach (where the browser decides how to display text) with a the physical approach (where the developer forces the browser to use special fonts and font attributes). Each approach has its advantages and disadvantages. One must make a decision on which design approach to take and not to mix them. The problem is further aggravated by using the <NOBR> tag in the head and foot parts of virtually all Web pages. The <NOBR> element stands for NO Break. This means that all the text between the start and end of the <NOBR> elements cannot have line breaks inserted. While <NOBR> may be essential for those character sequences that one intends to keep together, it should be used carefully. The <NOBR> forces the browser to display the content of the <NOBR>element on the same line, which results in an odd appearance when the user alters the display windows properties or font attributes in his or her browser. In such cases, the user is forced to scroll horizontally to see the content of the line.

Presentation of Pictures

Only the pictorial part of emblems is scanned. Pictures are offered in GIF format. The pictures are presented in low resolution to keep the file size to a minimum, which reduces download time. Yet, I get the impression that this was overdone. Images are not as clear and as sharp as one would expect them to be. The <IMG> tag inserts pictures in a Web page. One can, and should, specify the dimensions of the image as they are presented on the screen, which might not necessarily reflect the physical size of the picture in the original book. Enlarging the pictures or simply changing the *width* and *height* attributes of the <IMG> tag does not resolve the problem due to the low resolution of the scanned pictures. The main factor of the quality presented is the resolution of the scanned image. The problem could have been solved by scanning images in two different resolutions, one for displaying in all pages, and the other could be linked to the picture presented for users who wish to have a deeper look into emblems.

The entire content of available emblems is included in two main tables of contents provided in text format, and providing the title of emblems in both Latin and French. Unfortunately, this is presented in a complex table used to control the position of the text. Using complex tables and cells causes two main problems. Firstly, the download time is prolonged due to the fact that the browser needs to draw the entire table before the text is downloaded. Secondly, the text has an unpleasant appearance. The browser formats the cells according to the largest content of any given cell in the same column. Some table cells include a few words, while others are crowded with text. The developer could have adopted a simpler approach providing the content of the table of contents as text separating the line with <br> and <p> tags instead of the complicated table. Text in Latin and French could be displayed in different colors to make a clear distinction between them. The entire title or motto of the emblem is provided as a link, which clutters the page appearance even further. This could be resolved by making the link attached to the emblem reference number.

The developer makes poor use of the *meta* tag. The *meta* element is used within the *head* element to embed any useful information not defined by other HTML elements. Such information can be extracted by search engines in identifying, indexing and cataloguing documents. In fact, the only attributes used in the tag are the ones added automatically by the software used to create the site. In general, it seems that the developer relied entirely on the software used to create the site without examining the code added by the software or testing the Web site with different browsers and settings.

The site offers no search mechanism. This is not a major problem since the site offers two main tables of contents, which include all emblems available on the site.

6.3 Universidade Da Coruña

Universidade Da Coruña, Spain, offers a Website dedicated to emblematic studies in general and Hispanic emblematic literature in particular. The site is accessible through the URL http://rosalia.dc.fi.udc.es/emblematica/inindex.html. The site is offered in both English and Spanish. The Web site offers an example for the way Web sites for scholarly purposes should be documented. The research team, which includes several professors and graduate students in several disciplines, is properly introduced. Contact information includes, together with the traditional e-mail, telephone and fax numbers. The site seems to be frequently updated. The last update (as of this writing) is May 13, 2002. No copyright notice is included in the site.

6.3.1 Content

The site offers an extensive bibliography of almost 900 entries dedicated to emblem studies and more precisely to Hispanic emblematic literature. The site offers detailed information about Spanish emblem books from the sixteenth and seventeenth centuries that have been digitized and analyzed by members of the *Investigation Group on Hispanic Emblematic Literature at the University of Coruña, Spain*, giving details of some 27 works. The site offers many links to other Web sites dedicated to scholarly emblematic studies. The site offers access to a virtual library containing presumably as many as 27 works with some 1,816 emblems.

6.3.2 Design and Presentation

The site offers a clean and appealing design. The developer has adopted and retained a semantic approach to mark-up the Web pages. This allows a neat display of the content in different browsers and in different display settings. The design suffers from a few shortcomings, however. The content is presented in a few extremely long pages. This causes long download time and affects the site navigation. Links are made for parts of the documents that follow. In any case, the user will have to wait till the entire page is downloaded. The content could have been divided into smaller pages to reduce download time and the link mechanism changed accordingly. The developer uses a picture as texture for the page. Apart from the extra download time for the picture, the color of the text and that of the background do not contrast well, which makes reading more difficult than necessary. Instead, the developer could have used a color as background, which does not require any additional download time. The color of the text should always be chosen to contrast well with the background color. Also, I find it unnecessary to use small pictures for fax, telephone and contact information. Using a counter is not necessary either as it

adds no value to the page. For obvious reasons, I suggest changing word *Inglés* in the Spanish text to *English* and the word *Spanish* to *Español* in the English text.

The site links to the main page of *Proyecto de Investigación sobre Literatura Emblemática Hispánica*, which shows a different design raising the question if the two parts belong to the same site. It is always a good idea to keep a consistent design through the entire Web site.

The main page of *Proyecto de Investigación sobre Literatura Emblemática Hispánica* offers access to a *database and virtual library of Spanish emblem books.* There ends the good experience with the site. The virtual library is optimized for a high resolution (1200 x 760 or higher). This forces users, especially with small monitors, to scroll horizontally and vertically. The virtual library is optimized as well for specific browsers (Internet Explorer 5.x or Netscape Navigator 4.5 or later versions). Designing for specific browsers is not a good idea unless there is a special reason to do so. To enter the library a subscription is required. The subscription process, regardless of its technical deficiencies, asks the user to enter information, which might be considered confidential. Such information is accessible later for any user or subscriber through the *informacion de socios* link. There is no obvious way for a subscriber to unsubscribe or delete information he or she had previously entered. The wisdom of the entire process remains a mystery to me.

Assuming we are able to access the library, we are confronted with poor implementation of DHTML (Dynamic HTML) features. (Technically speaking there is no such a thing as a Dynamic HTML. The term is widely used to denote a combination of HTML features and other technologies such as CSS, scripting language [JavaScript or VBScript], Java Applet or other technologies, to create some kind of interactivity or dynamic change of the content or decoration of the Web page). Because of the poor implementation of DHTML, standard browser functions are no longer available. For instance the back button of the browser does not lead us to the previous page but to the main page of the *Research Project on Hispanic Emblematic Literature*, where one could access the virtual library again, after signing on one more time.

The main page of the Database related to the virtual library is another poor implementation of DHTML. The mouse pointer keeps changing forms and messages show up and disappear. The database itself offers two fields to be searched: choosing a work and choosing the author. Apart from the unusual reversed order of the fields (one would expect to choose an author then choose one of his works), the search result shows the query of the latest choice, not the combination of the two. The only help available is the *Manual de usario* [user's manual], which

appears to be empty. While the search query shows a specific number of emblems presumably available, no emblems seem to be accessible.

The virtual library offers also a chat room. Apart from being unnecessary it does not function. The entire concept of the virtual library and the related database should be reconsidered.

6.4 The University of Illinois at Urbana-Champaign

The University of Illinois at Urbana-Champaign offers a Web site dedicated to German emblem books as a part of its digital imaging project. The site is accessible through the URL http://images.library.uiuc.edu/projects/emblems/. The Web site merits a special attention. Among all American universities, the University of the Illinois has one of the most extensive collections of emblem books available. As the site informs us, the collection holds nearly 650 original editions and 400 emblem books in reprint or microform format. It is thus one of the largest collections in the United States. Due to this fact, digitization as well as accessibility of the collection through the Internet holds a special significance for the university and emblem scholars and students. For users, such an endeavor would allow them to access a wealth of emblem literature. For the university, it helps in preserving its valuable and rare collection by limiting the damage that is likely to occur through physical access.

Principal researchers involved in the project are Mara Wade, Associate Professor and Patricia Hardin, a graduate research assistant. E-mail contact information is provided. The Digital Imaging Project aims at presenting an on-line catalogue of 67 German emblem books (with ca. 2500 emblems). The research project only focuses on German emblem books, whereas the definition of this is both language and country specific. It is not clear, however, how many of these books have been actually digitized or are available on line. The source of emblems presented on the site is well documented. The site was last updated (as of this writing) 02.04.2002. The site sets a set of *Conditions of Use* that reflects dramatic restrictions on the use of the site.

6.4.1 Content

The site offers a searchable collection in a well-designed and functional database. The user can search the collection according to several fields, which are: *emblem id, short title, motto (Latin), motto German, descriptors, topos and author*. The user can also combine several search fields for a more accurate and limited result. The search result is then presented in a frame beneath the main page. A high-resolution version of the

searched emblem is easily accessed from the picture thumbnail provided in the search result. The picture is then presented with two main sets of information: the picture itself offered in a very clear, high resolution version, and descriptors accessible through a tab attached to the picture itself, which is scanned in its entirety giving detailed information about the picture and the technical process used in scanning the picture.

6.4.2 Design and Presentation

There are basically two types of data presented: the emblem, which is scanned and presented in its entirety, and the textual data related to the emblem. This includes metadata about the emblem and the motto presented mostly in Latin. No translation is provided. The German text is not presented in a normalized version according to the standards of the language today. The separation between motto and text is not clear. The motto in Latin is presented in a separate field, whereas the motto in German actually includes the text of the emblem itself. This might cause confusion for some users, especially for users who have difficulty with early modern German.

The metadata include information about the emblem such as the author, publisher, artist etc. It also includes a description of the pictorial content of emblems. The description method uses a controlled vocabulary that is more human-language oriented but also could be searched using the Iconclass classification system. Evidently, this approach tries to find a balance between *controlled vocabulary*, and *free text* indexing. The descriptions are written in English, but are made up of strings words, rather than sentences. The different fields related to metadata are stored in database for search purposes.

The picture is offered in two resolutions, one low resolution for fast download, and another version of the same emblem available in a higher resolution offering a near-print quality. I find this approach optimal. Some digital enhancement using a photo editing professional software program was done as well.

The site offers a clean and appealing design. The user is guided by clear instructions on how to navigate and is given appropriate information about pictures prior to download such as resolution and file size. The search mechanism is effective and reliable. However, one must question the wisdom of the way the search result is displayed. The search result displays thumbnails of the emblems corresponding to the search query in a frame beneath the main page. Apart from the inherent problems caused by frames (they are not supported by all browsers, the user is confronted with several windows and different scroll bars at the same time), the frame size

is offered in a very small size making it impossible to properly recognize the different pictures. This forces the user to enlarge several thumbnails to find the emblem he or she is looking for. This might turn to be a lengthy exercise since these emblems are provided in a relatively high resolution. If the user does not pay close attention to the existing frame on the lower part, especially when using smaller monitors, he or she is likely to overlook the entire frame and search result. It would have been better to display the search result in separate window including the title(s) of the emblem(s) found. Text files download much faster than thumbnails and are capable of providing more information and more results than the four thumbnails displayed in the lower frame.

The database uses Latin and English in naming the different fields of the database. Mixing language, while minimal, is not a good idea. One last remark is related to the poor use of the <META> tag, which was observed in the Glasgow project, and which also seems to be ignored by most emblem Web sites.

6.4.3 Digitization Process

The site offers detailed information about the digitization process, some of which seem to be contradictory. The *Technical Report* accessible through the URL http://images.library.uiuc.edu/projects/emblems/technical.html, mentions that "*Images were saved in Tiff format, zero compression*", whereas the technical description of the process in the emblems presented on the site clearly shows that there no print format was chosen in the process. I believe that was not a very wise decision. Pictures in JPEG format, as I mentioned previously, are more Internet-oriented, but not the best format for print. It is obvious that print-oriented formats such as TIFF or other proprietary formats require much more resources then JPEG or GIFF. Yet, this should not be an obstacle. The market offers today many possibilities for affordable storage media. Using CD-ROM as a storage medium to save such formats would probably be the most affordable way to do this. CD-Writers are not as expensive as they used to be and many computers are equipped with these CD-Writers today. Such print-oriented formats could be easily manipulated in a later stage to produce JPEG or GIFF formats. It offers as well the possibility of reprinting the emblem books in a better quality than compressed format. Apart from the superior quality of print, print-oriented format is not compressed. This allows us to keep the best possible version for further development in digital imaging. Compressed formats such as GIFF or JPEG cannot be uncompressed. The contrary holds true for print-oriented formats, which can be compressed or manipulated later thus offering a better quality in any new image format that might develop in the future.

6.5 Mnemosyne Partners

Mnemosyne Partners offer a Website dedicated to issues related to digitization and dealing with pictures by means of modern technologies. The company chose the somewhat difficult to remember name of the Greek Goddess of memory Mnemosyne as the name for its Web site. The site is accessible through the URL http://www.mnemosyne.org. The provider is located in The Netherlands. The site claims to have been developed by a small and dedicated group of people with firm roots in several disciplines related to art history, humanities and computing, which turned later to be a company offering its services in digitization projects, electronic publishing and other related matters. The site is not dedicated solely to emblems, although it includes some emblem books. The approach considered by the site is worth discussing due to the unique approach it takes in dealing with emblems and pictures in general. The site claims to use SGML/XML markup and the Iconclass application in their digitization. They deal with emblems or pictures in two main tracks: the picture itself, which is described using the Iconclass system, and the data related to the picture as well as the text, which are stored in computer files and dealt with by means of three standards (if we ignore the earlier attempts they had abandoned using the inconvenient ASCII text files): SGML, TEI and more recently XML. I will discuss both tracks separately. As I explained in pervious chapters, XML guarantees many advantages in dealing with data in general and text in particular. TEI is an application of SGML and XML is a small SGML version, thus compatibilities between these three standards is guaranteed and the migration from one standard to the other is made with relative ease. Prior to the development of XML the site used SGML to deal with text and metadata related to emblems. Now they seem to be migrating from SGML, which is a more powerful markup scheme but also more complicated, to the XML standard. To do this, they transformed SGML files to XML using some utilities. This approach allows for easy transformation of XML files to other formats using XSL. One of the possible transformations, for instance, would be converting XML files to HTML for presenting on the Internet as they did with some emblem books. So it appears that the company is moving in the right direction allowing for great flexibility and reusability of the data stored. This is so far one of the best approaches in dealing with emblems. The site does not mention if DTD was created specifically to deal with emblems. Using SGML, as they have done in the past, would indicate that such a DTD is available, since SGML requires a DTD. Yet, it is not clear if such a DTD was designed to deal with the textual part of the emblem or solely with textual information or metadata related to emblem files such as vendor, publisher and so on.

Pictures, on the other hand, are dealt with by the indexing system known as Iconclass, which I will discuss here briefly. Due to the complexity of the program, I will limit my discussion

to the basic concept of the program. Iconclass is a subject specific international classification system for iconographic research and the documentation of images developed by Henri van de Waal (1910-1972), professor of Art History. The concept of identifying, describing and classifying pictures, visual symbols and works of arts is not new. The earliest attempts in this field were made in the sixteenth century. The most important work in the field is perhaps Cesare Ripa's work *Iconologia* (1593). In this work, Ripa presented different personifications and motifs and enriched them with textual description and a simple explanation of the motif and its meaning. Iconclass seems to depart, however, from the *free indexing* used by Ripa, a concept, which is still widely used today. The program employs a set of *controlled vocabulary* to describe the different elements of the picture. This vocabulary contains approximately 28,000 definitions. Due to this fact, usage of the program is likely to be very complex for developers and users alike. The search result of the query will substantially depend upon the entries given by the developer or indexer of the picture. In other words, the more words we use in describing and indexing pictures, the more search results we get, putting in peril thus the main aim of the program of facilitating access to pictures. Limiting entries to a minimum, on the other hand, will result in less search results but might exclude the desired picture.

There are also basic problems related to the usage of language. Language is alive and evolving and we cannot restrict its usage to any number of words, not even 28,000 words. We need to be aware of the fact that many terms are identical or have subtle differences. If several developers are working on the same project, the likelihood that they will employ different terms for the same item is high. In fact, this is one of the major problems related to developing software by companies employing several programmers in the same project, even though the developers deal with a limited vocabulary related to computer concepts and programs. In many cases, one observes some inconsistency of terminology in the same program, even if there is one developer.

Describing what seem to be identical objects is likely to depend upon the cultural context in which it is used. Thus describing a picture of an ancient Greek temple in terms of a medieval church or an Asian temple would be inaccurate. Language is context sensitive, both culturally and historically. Mere translation, which seems to be available with Iconclass in several languages, is not likely to solve the problem.

6.5.1 Content

The site offers varieties of works and promises to include more in the future. I will deal here solely with emblems as they are presented on the site. There are only two emblem books available on the site:

Alciato's, *Emblematum Libellus* (Paris: Christian Wechel, 1534), offering 113 emblems. Emblems are presented in their entirety. The quality of the scanned picture makes me wonder if any digital image enhancement by means of image editing professional software was made. The texts of the emblems, including the motto, are reproduced in the original language, Latin. No translation is available. The emblem is accompanied by a framework including bibliographical information as well as metadata related to the matter presented, references and the entries related to Iconclass.

Hadrianus Junius' *Emblemata,* Latin edition, (Antwerp: Christopher Plantin, 1565) offering 58 emblems and 6 more pages for front matters. The pictures are presented in JPEG format. It appears that there are two versions of the same picture: a light one for presentation as the user accesses the main entry, and a high-resolution version of the same emblem (the source code indicates the presence of larger pictures). Unfortunately I was not able to access the high-resolution version. The reason for this could be broken links or the unavailability of the content on the server.

The source of emblems presented on line is accurately documented.

6.5.2 Design and presentation

Emblems are presented in separate frame sets, where the content is available on the right frame, and the titles of emblems are available in the navigation frame on the left. Frames are usually troublesome as a navigation mechanism. The user is not only confronted with two different pages at the same time, but with two different scrollbars and two navigation schemes also. To further complicate matters, there are links on both the main page and the main links frame. It is difficult in such mechanism to predict what will happen when the user clicks on a link. If users create a bookmark in their browser they may not get the page again when they follow the bookmark at a later date, since the bookmark doesn't include a representation of the state of the frames on the page. If the user attempts to print the page, he or she prints actually the page of the active frame, which may not be the one he or she wishes to print. Search engines have trouble with frames since they don't know what composites of frames to include as navigation units in their index.

6.6 Studiolum CD-ROM Evaluation

Studiolum is a CD-ROM edition touted as an emblem encyclopedia. To the best of my knowledge, it is the only CD-ROM edition available in this category. The program was developed by Tamás Sajó. The CD-ROM offers an extensive collection of works related to emblem literature

and some works of the antiquity likely to have been a source of knowledge for emblem writers. The collection of texts and images is bundled with an indexing and search software known as Folio View 4.20 to facilitate finding and retrieving information. The collection is accompanied by a lexicon offering bibliographical entries about several writers, some of whom are emblem writers, others are writers of the antiquity.

6.6.1 Content

The program offers access to a wealth of precious works, a total of 40 works by 15 authors, and a lexicon providing bibliographical entries about 18 authors. Some of the authors belong to the fifteenth century and the period after that such as Alciato (1492-1550), others are authors of antiquity such as the ancient Greek poet Anacreon (582- 485 B.C.). The works are available in their original languages. No translation is provided. The texts are enriched with links leading to annotations, comments and, if available, to other editions of the same work. This is particularly helpful with works offered in several editions and several languages. However, such editions cannot be considered as mere translation due to the important differences among them. For instance Alciato's *Emblemata*, Augsburg edition (1531), in Latin, shows 104 emblems, while the Padua edition (1626), available in Italian and Latin, offers more than 200 emblems.

The content in Studiolum is organized in two principal ways:

a) Hierarchically: authors, books, chapters and text.

b) For search purposes the content is further divided into fields, marked in the texts on the basis of language, type of the texts and so on. According to the developer, there are 31 fields, yet the *hit list property* offers actually 49 fields to be searched, if we add the *relevance rank* and *partition available,* the number raises to 51.

The works presented in Studiolum are not critical editions, rather a reproduction of the available works described by the developer as: "*annotated electronic facsimiles*" [50]. The developer claims that the original form of the texts was kept as far and as often as possible, and even their errors. Bibliographical entries related to the source of the text presented are not available. The only information available is the city of publishing and the publishing year. Perhaps that is all that is available in the printed version. However, the author could have mentioned the library(ies) and location of his resources for the edition; a practice that is necessary with old books and materials, because no two copies of the same printing are necessarily identical. It is not clear if the works are presented in their entirety or not.

From emblematic point of view, the works included in the CD-ROM could be divided into three main categories:

Emblem Books

Such as the different editions of Alciato's *Emblemata*, offering 104 Emblems in the Augsburg edition (1531) and 212 emblems in the Padua edition 1626.

Books Directly Related to Emblem Literature

Such as the different editions of Giovio's *Dialogo dell'imprese*. Of special importance is the Lyons edition (1574), offering more than 130 imprese accompanied by the motto written in Latin and description and explanation mostly written in Italian. The bearer of impresa and his position in the society are in most cases identified. One must also appreciate, among others, the rare collection of about 250 old medals, mostly from Roman times, as presented in Erizzo's work with detailed description and comments in Italian. Paradin's work *Devises Heroïques*, Lyons edition (1557), which includes 326 emblem pictures accompanied by description and explanation is another example of books with great value for emblem scholars. Including Ripa's exceptional work *Iconologia*, offered in three different editions, adds a great value to the collection; especially the London edition (1709) offering 326 pictures with detailed description and explanation in English.

Non-Emblematic Books

There are works that cannot be considered as emblematic and could be easily classified in a different category such as Baldini's *Mascherata*, Florence edition (1565). This does not mean that such works are of no interest to emblem scholars. Indeed, many of these works are of great value to emblem scholars as they reflect the literary spirit of the times. Due to my unfortunate lack of knowledge in Latin and ancient Greek, I am in no position to judge the content of works presented solely in ancient Greek or Latin.

Studiolum is of great value to emblem scholars, yet I have difficulty to consider it as an emblem encyclopedia, not only due to the fact that it includes non-emblematic works, but also because it does not offer enough information specific to the subject. I am not assuming that an emblem encyclopedia should include all emblematic works, although that would be a great achievement. The purpose of Studiolum was initially to reproduce the sixteenth-century and the seventeenth-century humanist library. The only coherent relationship seems to be the period of publication, which mostly concentrates on works published between 1517 and 1709. The only exception is the Budapest edition (1997) of Ripa's *Iconologia*. Some of the authors belong to the fifteenth century and the periods that followed it, whereas others are authors of the antiquity such as the great ancient Greek poet Anacreon (582-485 BC).

6.6.2 Links to the Internet

World Wide Web links generally allow users to get more detailed, specific and current information on thousands of topics. Unfortunately, Studiolum does not offer any links to Web sites.

6.6.3 The Intended Public

It is clear that the product is designed for emblem scholars as well as scholars in other disciplines related to ancient art and literature. Good knowledge of several languages, and most of all of Latin, is presumed. Yet, one must question the wisdom of not providing any translations or online dictionary to help the user get over the language barrier. The likelihood that all scholars and students who might be interested in the product will be aware of all the languages available is very slim. Unfortunately, there is not enough bibliographical information or any recommendations for further reading. The information offered about the authors whose works are included in the CD-ROM is somewhat limited. There are also some entries about authors whose works are not mentioned in the CD-ROM such as Amaseo and Decius Baronius on the one hand, but no entries about some authors whose works are included such as Baldini, Doni, Giovio, Paradin and Vico.

6.6.4 Authority and Documentation

Scholars need to identify and specify their resources properly. Unfortunately, such information is absent in Studiolum. During the installation process the name *Tamás Sajó*, presumably the developer, is given. Clicking the *about* option, in the help menu in Studiolum, does not provide any information about the product or about the company or the individual who created the program. This is particularly alarming, since the product seems to address scholars, not the general public. The only information available is related to Folio View 4.20, the search and indexing engine used by Studiolum, but not Studiolum itself.

6.6.5 Usability of the Program

Readability

The software allows the displaying of different fonts and offers the possibility of changing the zoom, which is a very helpful feature.

Print Preview

Print as well as print preview are available. The user has several options to set-up the page in print mode.

User Comments and Annotations

To add comments the user is required to create *shadow files* using the option *New* in the FILE menu. One must create the shadow files in another storage medium such as HDD. The user may enter his or her comments and can send them to the developer as suggestions to improve the product. Yet, copying from the database itself is prohibited. This is somewhat strange. In my view this is some kind of one-way sharing. The user might also confuse the *shadow* file with the main *library* file.

Help

Most users hate to use help files and would prefer to use a program that is designed to be as intuitive as possible. Studiolum is far from being intuitive and help files are also virtually not available. Trying to access help files using any available option leads to the main empty window of Folio View 4.20, (figure 6-1). The help offered by Folio View 4.20 seems to be the standard help on how to use the search engine, not how to use Studiolum.

Search

Searching the database is basically a word-search. The simple query search does not support Boolean operators, thus searching for *salamander and fire* would generate hits of all the words mentioned in the Boolean phrase. This applies as well to the Boolean operator itself such as "*and*" or "*or*". Some inconsistency has been noticed as well. For example search for *paradin + giovio* would generate hits related only to the last word, in this case, *giovio;* searching for *salamander + fire* would generate hits that correspond to *salamander* or *fire*.

Folio View documentation instructs us to use F3 for simple search. Doing so gives us a large window offering us as many as 51 fields to choose from, (figure 6-2). Specifying fields to be searched does not function as one would expect. For instance querying a word like *amore* gives results in languages that were not specified in the field options. The advanced search query offers the possibility of searching by scope using the brackets "[...]". Using this possibility shows nine choices, which need explanation such as *field, partition* and *popup.* Using the bracket does not seem to resolve any problem and it generates no hits. For instance, using *[Alciato]* or *[Alciato,* does not generate any hits. Grouping simple words using the Boolean operator *and* is well supported in the advanced search mode. Other Boolean operators are not equally supported and

show a great deal of inconsistency. Grouping conditions using quotation marks " " is not supported. Phrase search is supported. Yet, the phrase must be *exactly* as mentioned on the Studiolum CD, otherwise the search engine would give no result, thus searching for "*the salamander lying in the fire*" would generate one hit, searching for "*the salamander lying in fire*" would generate no hits. The search feature does not support truncation, thus searching for *Alciat* for instance, would give only results where the word *Alciat* is available. The search does not support different derivative forms of words. Words must be written *exactly* as they are indexed, thus searching for *salamander* will not produce other derivative forms of the word such as *salamander's, salamanders* and so on. Searching for the derivative forms on the other hand will generate no hits. The number of hits shown in the advanced query dialog box does not always correspond to the result shown in the hit-list frame. Some parts of the Studiolum CD were not indexed at all, for instance entries in the introduction are not indexed. Searching for some of the writers mentioned in the encyclopedia does not always generate hits as the case with the search query *Cartari*. In general the search mechanism is unnecessarily complicated, inconsistent and not very effective.

6.6.6 Interface

In its simplest definition, the interface is the way the user interacts with the software. There are two basic kinds of interface: command-line interface and graphical user interface known as GUI. For many users, the interface *is* the software. While this is fundamentally false, it reflects the importance of the interface. The interface should be as intuitive, comprehensible and as easy, *for users*, as possible. There are guidelines for developing an interface, but there are no absolute standards to which all software developers must adhere.

The Studiolum main interface contains several elements: the *title bar*, *menu bar*, *tool bar*, a bar including 7 tabs, different sets of resizable frames and status bar at the bottom of the screen, choose from, (figure 6-3). The number of items displayed on the menu bar and their content as well as the frame sets available in the main window varies according to the chosen tab. The idea of resizable frames might be a matter of taste, yet the complicated frame sets give the program an unnecessarily complicated appearance. In some cases the pictures in the *Document* frame did not display at all and one would have to keep on resizing the frame to display the picture. The *Object* tab has no obvious function but displaying the main icon of Studiolum. Using Studiolum for the first time, the user is confronted with a wizard similar to the Microsoft database Access program. Throughout the program many varieties of dialog boxes similar to database and word processing

programs are available. In fact, for successful work with the program some knowledge of both kinds of programs is required.

Interface Terminology

The terminology used in different dialogue boxes is mixture of English and Latin. Mixing different languages in the same interface is not a good design idea, even if the presumed user is knowledgeable in several languages. In some cases different terms are given to the same feature in dialog boxes. For example, the term *branch* refers sometimes to the same items named *liber* in other dialog boxes.

Choice of Icons

Developers use well-designed graphic icons to enhance the usability of applications. The developer of Studiolum uses icons very similar to Microsoft applications, which is a good idea, yet the purpose of using these items is entirely different from that of Microsoft applications. Some icons are identical in appearance but perform different actions. For example, the *expand to level* icon and the *expand entire table of contents* look identical with light color differences, that might not be perceived by users with eyesight difficulties. The simple search icon is identical in appearance with the *advanced search query* tab. The main idea of using icons is to perform actions, yet in some cases clicking an icon would force the user to make more choices through more dialog boxes as is the case with the *zoom* icon.

The content of the CD-ROM is presented as one huge file in the *Document* frame; this makes scrolling a hazardous action that might lead the user too far from the intended location.

The Options Dialogue Box

The *Option* dialog box, accessible through *Tools*, offers 8 tabs, each of which displays many different options, (figure 6-4). One would wonder, what all these options are and how could they influence the functionality of the program. Apart from requiring too much time to try them out, many choices do not seem to influence the usage of the program and different options show a great deal of inconsistency. HELP, as mentioned previously, does not offer any help at all. Many unnecessary options are provided as well, such as the possibility of replacing the + sign in the *Content* frame by different book icons in different variety of colors. On the other hand, one misses such an important feature that would allow the user to collapse all branches and sub-branches in the *Content* frame, especially if the user clicked the *expand entire table of contents* button, which

causes all branches and sub branches to be expanded. It was always easier for me to quit the program and start again instead of collapsing them individually.

6.6.7 Platform Conformity

The program is developed for the Windows platform. One would expect the program to reflect standard Windows operations and behavior. Windows users know that grayed-out options are options that are not available. The developer does not seem to respect this general rule in the program's design. In many cases options that are not available are dark but still perform no actions. For example, in the menu bar, *Windows* we have several options such as *cascades*, *title horizontally*, *title vertically* which perform no action (they are intended to be used when the user opens several windows). Right-clicking an item usually displays context-sensitive menu showing functions related to the item(s) selected such as copying. In most cases context-sensitive menus offer a long list of non-available options. In some cases options, presumably available, lead us to another submenu with all options grayed-out. In some frames right-clicking is not supported at all. Clicking different items, such as text or graphic shows standard Windows behavior like highlight or the square surrounding the images, yet some essential Windows functions are not allowed such as copying for instance. Instead of informing the user that he or she is not allowed to *copy*, the interface displays a message warning the user that he or she is not allowed to *edit* the database. Generally, error messages do not offer any help options nor do they guide the users to any possible solutions.

6.6.8 Technical Issues

Stability

The program is relatively stable; yet working with other graphic intensive applications at the same time such as other encyclopedia or multimedia applications seems to cause some conflict. Either Studiolum or the other program crashed and in some cases a reboot was necessary.

Speed

The program is relatively slow. It takes too long to display a dialog box or react to user's action.

Screen resolution

The program works successfully in several resolutions.

Upgrade

The program does not offer any information about possible upgrade in the future or expenses related to it.

Customer Support

No customer support is provided, no supporting Web site or even an e-mail link to the developer. Online support is provided only to the search engine used to index the program (Folio View 4.20).

Installation

The program installed and worked successfully on several versions of Windows. The installation program creates a desktop shortcut for user's convenience. Software packages usually have a text file, mostly named *readme,* which provides information related to the installation process, technical requirements, compatibility issues, trouble-shooting advice and the latest information about the software. The *readme* file in Studiolum is used merely as a copyright reminder. No contact or technical support information is provided. The program requires a hard disk space of 9-10 MB. Although not explicitly mentioned, one could copy the *library.nfo* file to the hard drive and work without the CD-ROM, which is a great advantage for mobile computing. The program runs significantly faster as well. No information about memory requirement is available. Using Norton Utility to control the usage of memory, the program requires about 11-12 MB. The program requires the installation of some fonts provided on the CD to display Greek characters. In some cases manual installation of the fonts was necessary. The process of installing these fonts is simple for experienced users but might pose a challenge for less experienced users.

Uninstall

The software does not provide any uninstalling mechanism. To uninstall the program one must rely on the uninstall feature of the OS or a third-party software.

Test Procedure

The program installed and worked successfully in the following versions of Windows: WIN 95, WIN 98 (French and English versions), as well as on Windows 2000 professional edition (English version) other operating systems were not tested. For the sake of accuracy, most of my evaluation work was conducted using WIN98 SE English version. The program has been tested on several machines and several hardware configurations.

6.6.9 Final Thoughts

Studiolum is of great value to emblem scholars. Unfortunately, the program suffers from serious shortcomings in classification and search mechanism and from a poorly designed interface. Most of these shortcomings are due to the search and indexing engine used in Studiolum (Folio View 4.2). Indexing and search engines are as necessary for encyclopedias as pillars are necessary for a building. Yet, no one would like to live in a house where pillars hinder the placing of furniture and the freedom of moving within the house. Most of the time one gets the impression that one is working with Folio View not with Studiolum. In order to include as many features as possible, and to offer a professional word processing program, the size of the program is large and its performance and consistency suffer serious shortcomings.

It is regrettable that the developer of Studiolum does not allow the user to copy texts or content of the material presented. The principle seems to be: you may see, but you are not allowed to touch, thus eliminating one of the greatest advantages of digital media. All this makes for poor usability of the program. Hopefully, these defects will be addressed in the further developments of the product.

Conclusion

Digital media offer great opportunities to preserve and present emblem literature, and make it more accessible to scholars, students and the general public. However, digital media have limitations and shortcomings that need to be addressed. The multidisciplinary nature of emblems creates additional challenges that need to be met. We should not expect a perfect solution, but it is our mission to make the best possible use of the technology available today. As might be expected, modern attempts at digitizing emblems show several shortcomings. These shortcomings are due in part to the limitations of technology, but also in part to the improper use of it.

In my introduction I have mentioned that we should attempt to achieve three main objectives: preserving, presenting and helping understand emblems and the emblematic world. The objectives of preserving and presenting emblems is well underway, although not yet far enough to achieve what would be the dream of any emblem scholar to have the entire corpus of emblems available in digital form. Unfortunately, one gets the impression that different institutions around the world are working in isolation. Often I had the impression that the attempt to digitize emblems is a scattered and fragmented one. Different institutions with one common goal resemble a group of islands in an ocean with no real connections between them. A few Web links among several Web sites constitute the only bridge among them. No real exchange of data or cooperation seems to exist among them. In presenting the different technologies I stressed the value of the possibilities offered by these technologies, which facilitate cooperation, transfer and share of data among different institutions. To use these possibilities effectively we need to address several issues, some of which are legal related to copyright. Some Web sites impose dramatic restrictions on the use of the materials presented and in some cases we have the one-way-sharing attitude.

The third objective depends greatly on efforts made by emblem scholars to bring emblem literature closer to students and the general public. In fact, I had often the impression that scholars are writing and presenting their materials to scholars only. We should not expect the general public or interested students to be knowledgeable of several languages, especially Latin and ancient Greek. Modern languages in their early modern form present a challenge for most students and non-native speakers. Offering normalized versions of emblem texts would be one step in the right direction. This will make emblems more accessible to the general public, and at the same time make them easier for computer technology to index and classify. We need to provide readers with articles and related bibliography. Also, I believe that some effort should be made to generate more interest in the general public and help it understand emblematic literature.

The international Society for Emblem Studies will hold a conference from the 10th to the 14th of September 2002 at the University Da Coruña, Spain, where researchers from all over the world will gather together. Hopefully, these issues will be addressed.

End Notes

[1] *See* Graham, David. " 'Voiez icy en ceste histoire.. ': Cross-Reference, Self-Reference and Frame-Breaking in Some French Emblems." Emblematica. AMS Press, Inc., New York, Volume 7, Number 1 (1993): 1-23.

[2] Encyclopaedia Britannica. Encyclopaedia Britannica Inc., CD-ROM ed., 1998. Search query: emblem.

[3] Daly, Peter M. Literature in the Light of the Emblem. Toronto, Ont.: University of Toronto Press, second ed. 1998.

[4] *See* Daly. Literature in the Light of the Emblem. 67.

[5] Schöne, Albrecht. Emblematik und Drama im Zeitalter des Barock. München, Ger.: C.H. Beck'sche Verlagsbuchhandlung, 1964.

[6] *See* Michel, Alain "Rhetorique et Philosophie de l'emblème: Allégorie, Réalisme, Fable". Jones-Davies. Ed. Emblèmes et devises au temps de la Renaissance. Paris, Fr: Université de Paris-Sorbonne, Centre de recherches sur la Renaissance, 1981, 23-31.

[7] *See* Katz, Wendy R. "John Aiken's 'On emblems': The sources of an Emblematic Dialogue." Emblematica. AMS Press, Inc., New York, Volume 7 (1993): 145-155.

[8] *See* Schöne. Emblematik und Drama im Zeitalter des Barock. 75.

[9] *See* Daly. Literature in the Light of the Emblem. 41.

[10] *See* Daly. Literature in the Light of the Emblem. 9-26.

[11] *See* Schöne. Emblematik und Drama im Zeitalter des Barock. 42-45.

[12] *See* Daly. Literature in the Light of the Emblem. 29.

[13] *See* Caldwell, Dorigen. "Reflections and Conclusions on Imprese." Emblematica. New York: AMS Press, Inc., Volume 11 (2001): 225-239.

[14] *See* Daly. "Modern Advertising and the Renaissance Emblem: Modes of Verbal and Visual Persuasion". Höltgen, Karl Josef, Peter M. Daly and Wolfgang Lottes Eds. Word and Visual Imagination: Studies in the Interaction of English Literature and the Visual Arts. Erlangen, Ger.: Universitätsbund Erlangen-Nürnberg e.V., 1988, 349-370.

[15] *See* Schöne. Emblematik und Drama im Zeitalter des Barock. 67-69.

[16] Der Spiegel Extra: Das Kultur-Magazin. Spiegel-Verlag Rudolf Augstein GmbH & Co. KG, Hamburg, Ger., Heft 1, January 1997, 24-25.

[17] Der Spiegel. Spiegel-Verlag Rudolf Augstein GmbH & Co. KG, Hamburg, Ger., No. 34, 20.08.2001, 10-19.

[18] Der Spiegel. Spiegel-Verlag Rudolf Augstein GmbH & Co. KG, Hamburg, Ger., No. 7, 13.02.1995, 8-9.

[19] *See* Schöne. Emblematik und Drama im Zeitalter des Barock. 110.

[20] *See* Wang, Xuemao et al. "Internet-A global perspective and its impact on library services." Fletcher, Patricia Diamond and John Carlo Bertot Eds. World Libraries On the Information Superhighway: Preparing for the challenges Of the New Millennium. London, Eng.: Idea Group Publishing, 2000.

[21] *See* Miall, David S. "The Library versus the Internet: Literary Studies under Siege?" PMLA. Volume 116, Number 5, October (2001): 1405-1414.

[22] Gibaldi, Joseph. MLA Handbook for Writers of Research Papers. New York: The Modern Language Association of America, fifth edition, 1999.

[23] *See* Alonso, Carlos J. "The Internet Sublime." PMLA. Volume 116, Number 5, October (2001): 1297-1301.

[24] Canadian Intellectual Property office, an agency of Industry Canada. A GUIDE TO COPYRIGHTS. PDF Format, 2000. <http://cipo.gc.ca>.

[25] Bielefield, Arlene and Lawrence Cheeseman. Technology and Copyright Law: a Guidebook for the Library, Research, and Teaching Professions. New York: Neal-Schuman Publishers, Inc., 1997.

[26] *See* Borland, John. Broadcasters launch online TV legal war. CNET News.com <http://www.zdnet.com/filters/printerfriendly/0,6061,2430035-2,00.htm>.

[27] *See* The Motion Picture Association of America (MPAA). PLAINTIFFS' PROPOSED FINDINGS OF FACT AND CONCLUSIONS OF LAW WITH RESPECT TO ISSUANCE OF PRELIMINARY INJUNCTION†. District Court For The Western District Of Pennsylvania, Icrave Press releases, <http://www.mpaa.org/Press/iCrave_Findings.htm>.

[28] *See* Maney, Kevin. Ruling stops iCraveTV transmissions. USA Today <http://www.usatoday.com/life/cyber/tech/cth248.htm>.

[29] *See* LaPlante, Gina. iCraveTV.com Gets Sued for Cyber Piracy. Internet Law Journal <http://www.tilj.com/content/ipheadline02090001.htm>.

[30] BERNE CONVENTION FOR THE PROTECTION OF LITERARY AND ARTISTIC WORKS (Paris Text 1971). <http://www.law.cornell.edu/treaties/berne/overview.htm>.

[31] Cornish, Graham P. Copyright: interpreting the law for libraries, archives and information services. London, Eng.: Library Association Publishing, 1997.

[32] Baudouin, Renaud et al. Eds. Québec Civil Code. Montréal, PQ., Wilson & Lafleur Limitée, 1994.

[33] Canadian Copyright Act. <http://lois.justice.gc.ca/en/C-42/text.html>.

[34] Romano, Frank J. PDF Printing and Workflow. Upper Saddle River: Prentice Hall PTR, 1999.

[35] The World Wide Web Consortium. HyperText Markup Language Home Page. <http://www.w3.org/MarkUp/>.

[36] Buyens, Jim. Running Micrsoft Front Page 98. Redmond: Microsoft Press, 1997.

[37] Goldfarb, Charles F. and Paul Prescod. The XML Handbook. Upper Saddle River: Prentice-Hall PTR, 2001.

[38] *See* Richmond, Alan. Introduction to XHTML: Why do we need XHTML? <http://www.wdvl.com/Authoring/Languages/XML/XHTML/why.html>.

[39] Lew, Michael S. Ed. Principles of Visual Information Retrieval. London. New York: Springer, 2001.

[40] *See* Russel, Daniel S. "The Needs of the Literary Historians" Peter M. Daly. Ed. The Index of Emblem Art Symposium. New York: AMS Press, 1990, 113.
See also Russel. "Perceiving, Seeing and Meaning: Emblem and some Approaches to Reading Early Modern Culture". Peter M. Daly and John Manning. Eds. Aspects of Renaissance and Baroque Symbol Theory, 1500-1700. New York: AMS Press, 1999, 89.

[41] *See* The Unicode Consortium. What is Unicode? <http://www.unicode.org/unicode/standard/WhatIsUnicode.html>.

[42] *See* Kuhn, Markus. UTF-8 and Unicode FAQ for Unix/Linux. <http://www.cl.cam.ac.uk/~mgk25/unicode.html#linux>.

[43] *See* Siemens, R.G. Methodology: Lemmatization. <http://www.chass.utoronto.ca/epc/chwp/siemens2/siemens3.html>.

[44] *See* Richardson, Tim. Software pirates cost industry $11 billion. <http://www.theregus.com/content/4/25188.html>.

[45] *See* BBC News. Music piracy rises worldwide. <http://news.bbc.co.uk/hi/english/entertainment/music/newsid_2039000/2039365.stm>.

[46] López Poza, Sagrario. et al. Eds. Grupo de Investigación sobre Literatura Emblemática Hispánica. <http://rosalia.dc.fi.udc.es/emblematica/>.

[47] *See* Graham, David. " 'Voiez icy en ceste histoire... ': Cross-Reference, Self-Reference and Frame-Breaking in Some French Emblems." Emblematica. AMS Press, Inc., New York, Volume 7, Number 1 (1993): 1-23.

[48] *See* Graham, David.. "Putting Old Wine in New Bottles: Emblem Books and Computer Technology." Emblematica. AMS Press, Inc., New York, Volume 5, Number 2 (1991): 271-285.

[49] *See* Daly. "Sixteenth-Century Emblems and Imprese as Indicators of Cultural Change" Whitman, Jon. Ed. Interpretation and Allegory : Antiquity to the Modern Period. Leiden; Neth:, Brill, (2000): 383-417.

[50] Sajó, Tamás "Introduction."Studiolum 2000. CD-ROM ed. 2000.

Bibliography

Adobe Systems Incorporated. Adobe Acrobat 4.0: Classroom in a Book. Berkeley: Peachpit Press, 2000.

Alciatus, Andreas. Emblematum Libellus. Darmstadt, Ger.: Wissenschaftliche Buchgesellschaft, 1967.

Alonso, Carlos J. "The Internet Sublime." PMLA. Volume 116, Number 5, October (2001): 1297-1301.

Balfour, Joya. The (dis)information highway: conspiracy theories on the Internet. MA-Thesis, McGill University, Montreal, PQ.: 2000.

Benson, Allen C. Complete Internet Companion for Librarians. New York: Neal-Schuman Publishers Inc., 1997.

Bielefield, Arlene and Lawrence Cheeseman. Technology and Copyright Law: a Guidebook for the Library, Research, and Teaching Professions. New York: Neal-Schuman Publishers, Inc., 1997.

Brooks, Charles J. and Marcraft International Corporation. A+ Certification: Training Guide. Kennewick: Marcraft International Corporation, 2001.

Buyens, Jim. Running Microsoft Front Page 98. Redmond: Microsoft Press, 1997.

Caldwell, Dorigen. "Studies in Sixteenths Italian Imprese" Emblematica. New York: AMS Press, Inc., Volume 11 (2001): 1-257, especially "Later Development in Theory." 181-210, and "Reflections and Conclusions on Imprese." 225-239.

Chowdhury, G.G. Introduction to modern information retrieval. London, Eng.: Library Association Publishing, 1999.

Cornish, Graham P. Copyright: interpreting the law for libraries, archives and information services. London, Eng.: Library Association Publishing, 1997.

Daly, Peter M. "Reviews: P. VAN HUISSTEDE and J.P.J. BRANDHORST. *Dutch Printer's Devices*. 15th-17th Century. A Catalogue. With CD-Rom. Nieuwkoop: De Graaf Publishers, 1999. 3 volumes. Pp. viii, 1-653 + 654-1191 + 1195-1639." Emblematica. New York: AMS Press, Inc., Volume 11 (2001): 421-433.

---. "The Bibliographic Basis for Emblem Studies." Emblematica. New York: AMS Press, Inc., Volume 8, Number1 (1994): 151-175.

---. Emblem Theory: Recent German Contributions to the Characterization of the Emblem Genre. Nendeln, Ger.: KTO Press, 1979.

---. Dichtung und Emblematik bei Catharina Regina von Greifennberg. Bonn, Ger.: Bouvier Verlag Herbert Grundmann, 1976.

---. Literature in the Light of the Emblem. Toronto, Ont.: University of Toronto Press, second ed. 1998.

---. "Modern Advertising and the Renaissance Emblem: Modes of Verbal and Visual Persuasion." Höltgen, Karl Josef, Peter M. Daly and Wolfgang Lottes, eds. Word and Visual Imagination: Studies in the Interaction of English Literature and the Visual Arts. Erlangen, Ger.: Universitätsbund Erlangen-Nürnberg e.V., 1988.

Der Spiegel Extra: Das Kultur-Magazin. Spiegel-Verlag Rudolf Augstein GmbH&Co. KG, Hamburg, Ger., Heft 1 January 1997. 24-25.

Der Spiegel. Spiegel-Verlag Rudolf Augstein GmbH & Co. KG, Hamburg, Ger., No. 7, 13.02.1995. 8-9.

Der Spiegel. Spiegel-Verlag Rudolf Augstein GmbH & Co. KG, Hamburg, Ger., No. 34, 20.08.2001. 10-19.

Dickinson, Gail K. Selection and Evaluation of electronic Resources. Colorado: Libraries Unlimited, Inc. 1994.

Fletcher, Patricia Diamond and John Carlo Bertot. World Libraries On the Information Superhighway: Preparing for the challenges Of the New Millennium. London, Eng.: Idea Group Publishing, 2000.

Gibaldi, Joseph. MLA Handbook for Writers of Research Papers. New York: The Modern Language Association of America, fifth edition, 1999.

Goldfarb, Charles and Prescod, Paul F. The XML Handbook. Upper Saddle River: Prentice-Hall PTR, 2001.

Graham, David. " 'Voiez icy en ceste histoire... ': Cross-Reference, Self-Reference and Frame-Breaking in Some French Emblems." Emblematica. AMS Press, Inc., New York, Volume 7, Number 1 (1993): 1-23.

---. "Putting Old Win in New Bottles: Emblem Books and Computer Technology." Emblematica. AMS Press, Inc., New York, Volume 5, Number 2 (1991): 271-285.

Graham, Ian S. HTML 4.0 Sourcebook: A Complete Guide to HTML 4.0 and HTML Extensions. New York: Willey Computer Publishing, 1998.

---. XHTML 1.0 Language and Design Sourcebook: The Next Generation HTML. New York: Willey Computer Publishing, 2000.

Groth, David, David Wall and Michael De Beer. i-Net+ Study Guide. Alameda: Sybex, 2000.

Haggard, Mary. Survival Guide to Web Site Development. Redmond: Microsoft Press, 1998.

Johnson, Jeff. GUI Bloopers: Don'ts and Do's For Software Developers and Web Designers. San Francisco: Morgan Kaufmann Publishers, 2000.

Johnson, Peggy and Bonnie MacEwan, eds. Virtually Yours: Models for Managing Electronic Resources and Services. Chicago: American Library Association, 1999.

Jones-Davies. M.T., ed. Emblèmes et devises au temps de la Renaissance. Centre de recherches sur la Renaissance, Paris, Fr. : Université de Paris-Sorbonne, 1981.

Katz, Wendy R. "John Aiken's 'On emblems': The sources of an Emblematic Dialogue." Emblematica. AMS Press, Inc., New York, Volume 7 (1993): 145-155.

Lagadec, Jean. Le nouveau Guide pratique du Droit. Paris. Fr.: France Loisirs, 1994.

Lew, Michael S., ed. Principles of Visual Information Retrieval. London. New York: Springer, 2001.

López Poza, Sagrario. "Spanish Emblem Studies, Status Quaestionis." Emblematica. AMS Press, Inc., New York, Volume 10, Number 2 (1996): 347-377.

Marcella, Rita and Arthur Maltby, eds. The Future Of Classification. Hampshire, Eng.: Gower Publishing limited, 2000.

McGrath, Sean. XML by Example: Building E-commerce Applications. Upper Saddle River: Prentice-Hall PTR, 1998.

Merz, Thomas. PostScript and Acrobat/PDF: Applications, Troubleshooting and Cross-Platform Publishing. Berlin, Ger.: Springer-Verlag, 1997.

Miall, David S. "The Library versus the Internet: Literary Studies under Siege?" PMLA. Volume 116, Number 5, October (2001): 1405-1414.

Microsoft Corporation. Computer Dictionary. Redmond: Microsoft Press, 1997.

Pike, Mary Ann et al. Using the Internet: The Most Complete Reference. Indianapolis: Que Corporation, 1995.

Poulter Alan, Hiom Debra and Tseng Gwyneth. The Library and Information Professional's Guide To The Internet. London, Eng., Library Association Publishing, 2000.

Baudouin, Renaud et al., eds. Québec Civil Code. Montréal, PQ., Wilson & Lafleur Limitée, 1994.

Romano, Frank J. PDF Printing and Workflow. Upper Saddle River: Prentice Hall PTR, 1999.

Russel, Daniel S. "Perceiving, Seeing and Meaning: Emblem and some Approaches to Reading Early Modern Culture". Daly, Peter M. and John Manning, eds. Aspects of Renaissance and Baroque Symbol Theory, 1500-1700. New York: AMS Press, (1999): 89.

---. "The Needs of the Literary Historians" Daly, Peter M., ed. The Index of Emblem Art Symposium. New York: AMS Press, (1990): 113.

Schöne, Albrecht. Emblematik und Drama im Zeitalter des Barock. München, Ger.: C.H. Beck'sche Verlagsbuchhandlung, 1964.

Schwartz, Richard. Web Site Construction, Tips & Tricks. Albany: OnWord Press, 2001.

Siegel, David. Creating Killer Web Sites: The Art of Third Generation Site Design. Indianapolis: Hayden Books, 1997.

Smith, Mark. Internet Policy Handbook for Libraries. New York: Neal-Schuman Publishers Inc., 1999.

Spolsky, Joel. User Interface Design for Programmers. New York: Springer-Verlag, 2001.

Still, Julie M., ed. The Library Web: Case studies in Web site creation and implementation. Medford: Information Today, Inc. 1997.

Tel-A-Vision Media Networking GmbH. Scanners: photo numérique et Traitement d'image. Paris, Fr.: Sybex, 1999.

Vablais, Cerise and Leininger, Tony eds. How The Web was Won: conquering the Digital Frontier. Redmond, Microsoft Press, 1998.

Vicario, Enrico, ed. Image Description and Retrieval. New York: Plenum Press, 1998.

Whitman, Jon, ed. Interpretation and allegory : antiquity to the modern period. Leiden; Neth.: Brill, 2000.

Digital Resources

Floppy Disks 1.44

Duden: das Herkunftswörterbuch. Floppy Disks, Bibliographisches Institut & F.A. Brockhaus AG 1994 Langenscheidt KG, Aus.

CD-ROMs

Bertelsmann Universallexikon 1997. CD-ROM ed. München, Ger.: Bertelsmann Electronic Publishing, 1996.

Encyclopaedia Britannica. Encyclopaedia Britannica Inc., CD-ROM ed., 1998.

EPCollin's Fremdsprachen-Assistent Version 1.0D. CD-ROM ed. Köln, Ger.: EPC Software GmbH, 1995.

Larousse unabridged French English, English-French Dictionary. CD-ROM ed. Paris, Fr.: Larousse Computer Software, 1995.

Oxford English Dictionary. CD-ROM version 2.00, Oxford University Press, 1999.

Sajó, Tamás. Studiolum 2000. CD-ROM ed. 2000.

Internet Resources

Adobe Systems Incorporated. What's new in Adobe Acrobat 5.0. <http://www.adobe.com/products/acrobat/newfeatures.html>.

Arampatzis A. T. et al. Linguistically-motivated Information Retrieval Technical Report CSI-R9918. Dept. of Information Systems, Faculty of Mathematics and Computing Science, University of Nijmegen, The Netherlands. <http://www.cs.kun.nl/~avgerino/Avi_Arampatzis/publications/HTMLized/encyclop/encyclp.html>.

BBC News. Music piracy rises worldwide. <http://news.bbc.co.uk/hi/english/entertainment/music/newsid_2039000/2039365.stm>.

Borland, John. Broadcasters launch online TV legal war. CNET News.com. <http://www.zdnet.com/filters/printerfriendly/0,6061,2430035-2,00.htm->.

Canadian Intellectual Property office, an agency of Industry Canada. A GUIDE TO COPYRIGHTS. PDF Format, 2000. <http://cipo.gc.ca>.

Department of Justice Canada. Copyright Act. <http://laws.justice.gc.ca/en/C-42/36076.html>.

World Wide Web Consortium. HyperText Markup Language Home Page. <http://www.w3.org/MarkUp/>.

Hartemink, Ralf. INTERNATIONAL CIVIC ARMS. <http://www.ngw.nl/indexgb.htm>.

IBM. Information Brief: CD-ROM Technology. <http://www.pc.ibm.com/us/infobrf/cdtech.html>.

Legal Information Institute. Copyright: an overview. <http://www.law.cornell.edu/topics/copyright.html>.

---. BERNE CONVENTION FOR THE PROTECTION OF LITERARY AND ARTISTIC WORKS (Paris Text 1971). <http://www.law.cornell.edu/treaties/berne/>.

LaPlante, Gina. iCraveTV.com Gets Sued for Cyber Piracy. Internet Law Journal. <http://www.tilj.com/content/ipheadline02090001.htm>.

Maney, Kevin. Ruling stops iCraveTV transmissions. USA Today. <http://www.usatoday.com/life/cyber/tech/cth248.htm>.

Nielsen, Jakob. Top Ten Mistakes in Web Design. <http://www.useit.com/alertbox/9605.html>.

---. Top Ten Guidelines for Homepage Usability. <http://www.useit.com/alertbox/20020512.html>.

Parker, James. A GLOSSARY OF TERMS USED IN HERALDRY. <http://www04.u-page.so-net.ne.jp/ta2/saitou/ie401/index.htm>.

Richardson, Tim. Software pirates cost industry $11 billion. <http://www.theregus.com/content/4/25188.html>.

Richmond, Alan. Introduction to XHTML: Why do we need XHTML? <http://www.wdvl.com/Authoring/Languages/XML/XHTML/why.html>.

Royal Netherlands Academy of Arts and Sciences (KNAW). What is ICONCLASS ? <http://www.iconclass.nl/texts/sysfr.htm>.

Siemens, R.G. Lemmatization and Parsing with TACT Preprocessing Programs. Department of English, University of British Columbia. <http://www.chass.utoronto.ca/epc/chwp/siemens2/index.html>.

Text Encoding Initiative. <http://www.tei-c.org>.

The Motion Picture Association of America (MPAA). <http://www.mpaa.org/home.htm>.

---. PLAINTIFFS' PROPOSED FINDINGS OF FACT AND CONCLUSIONS OF LAW WITH RESPECT TO ISSUANCE OF PRELIMINARY INJUNCTION†. District Court For The Western District Of Pennsylvania, Icrave Press releases, The Motion Picture Association of America (MPAA). <http://www.mpaa.org/Press/iCrave_Findings.htm>.

The Unicode Consortium. Unicode Standard Annex #28: Unicode 3.2. <http://www.unicode.org/unicode/reports/tr28/>.

---. The Unicode® Standard: A Technical Introduction. <http://www.unicode.org/unicode/standard/principles.html>.

---. What is Unicode? <http://www.unicode.org/unicode/standard/WhatIsUnicode.html>.

Kuhn, Markus. UTF-8 and Unicode FAQ for Unix/Linux. <http://www.cl.cam.ac.uk/~mgk25/unicode.html#linux>.

United States Copyright Office. Copyright Basics. <http://www.copyright.gov/circs/circ1.html>.

Velde, François R. Coat of Arms of Jeanne d'Arc.

<http://www.heraldica.org/topics/france/jeannedarc.htm>.

---.PARTS OF A COAT OF ARMS. <http://www.designsofwonder.com/articles/shieldparts.html>.

Evaluated Web Sites

Adams, Alison and David Graham, eds. The Glasgow University Emblem Website. <http://www.emblems.arts.gla.ac.uk/>.

Wade, Mara and Patricia Hardin, eds. The University of Illinois at Urbana-Champaign emblem web site. <http://images.library.uiuc.edu/projects/emblems/>.

The Mnemosyne Partnership. <http://www.mnemosyne.org/>.

Universidade da Coruña. España.
López Poza, Sagrario et al., eds. Grupo de Investigación sobre Literatura Emblemática Hispánica. <http://rosalia.dc.fi.udc.es/emblematica/>.

APPENDIX A

List of Abbreviations

ARPANET	Advanced Research Projects Agency Network
ASCII	American Standard Code for Information Interchange
CD	Compact Disc
CD-R	CD-Writer, also used to designate writable or recordable CDs
CD-ROM	Compact Disc Read-Only Memory
CD-RW	CD-ReWriter, also used to designate rewritable CDs
CERN	Centre European de Recherches Nuclaire
CRT	Cathode-Ray Tube
CSS	Cascading Style Sheets
DHTML	Dynamic HTML
DTD	Document Type Definition
DTP	Desktop Publishing
DVD	Digital Versatile Disc or Digital Video Disc
GATT	General Agreement On Trade and Tariffs
GIF	Graphics Interchange Format
GUI	Graphical User Interface
HTML	HyperText Markup Language
HTTP	HyperText Transfer Protocol
IRS	Information Retrieval System
ISP	Internet Service Provider
JPEG	Joint Photographic Experts Group
LCD	Liquid Crystal Display
NCSA	National Centre for Supercomputing Applications
NSF	National Science Foundation
OCR	Optical Character Recognition
OS	Operating System
PDF	Portable Document Format
PNG	Portable Network Graphics
SGML	Standard Generalized Markup Language
TCP/IP	Transmission Control Protocol / Internet Protocol
TEI	Text Encoding Initiative

TFT	Thin Film Transistor
TIFF	Tagged Image File Format
TRIPS	Trade-Related Intellectual Property issues
TWAIN	Technology Without An Interesting Name
UCC	Universal Copyright Convention
URL	Uniform Resource Locator
VBScript	Visual Basic Script
W3C	World Wide Web Consortium
WORM	Write Once, Read Many
WYSIWYG	What You See Is What You Get
XHTML	eXtensible Hypertext Markup Language
XLL	eXtensible Link Language
XML	eXtensible Markup Language
XSL	eXtensible StyleSheet Language

APPENDIX B

Illustrations

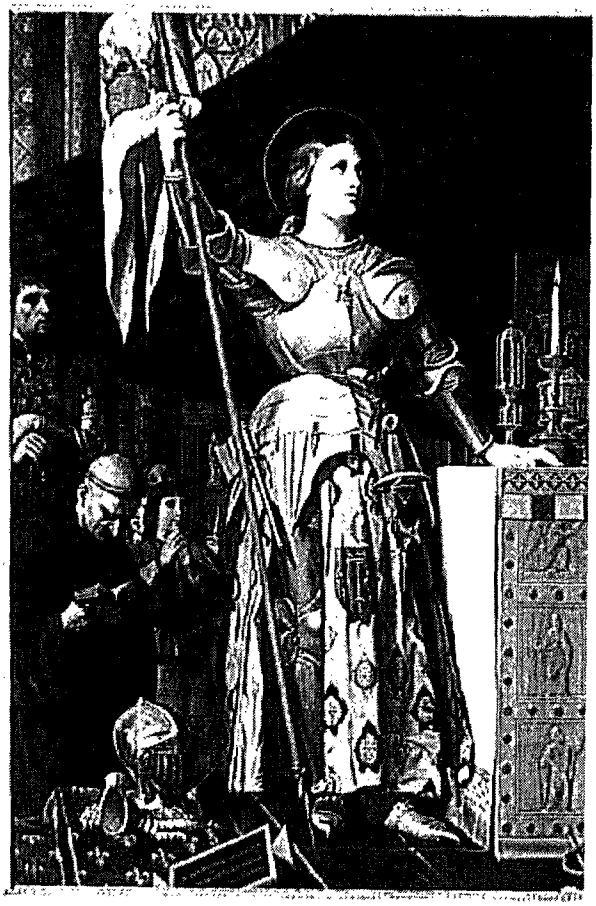

Figure 1-1
Jeanne d'Arc at the coronation of Charles VII, by Ingres. (France, Ministère de la Culture, base de données Joconde.)
Source:
<http://www.heraldica.org/topics/france/jeannedarc.htm>

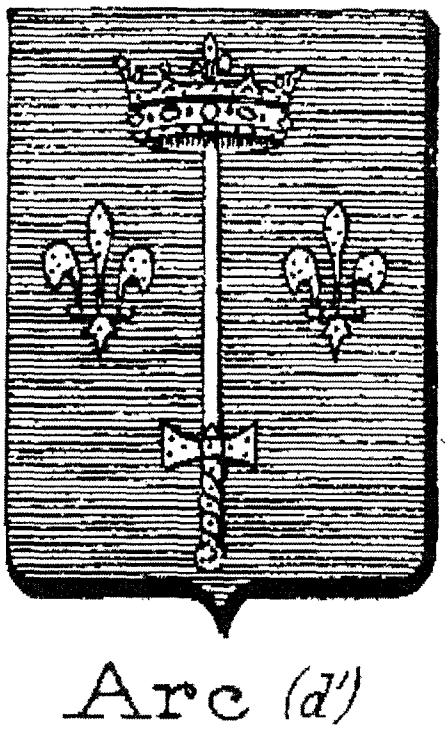

Figure 1-2
Coat of Arms of Jeanne d'Arc.
Source:
<http://www.heraldica.org/topics/france/jeannedarc.htm>

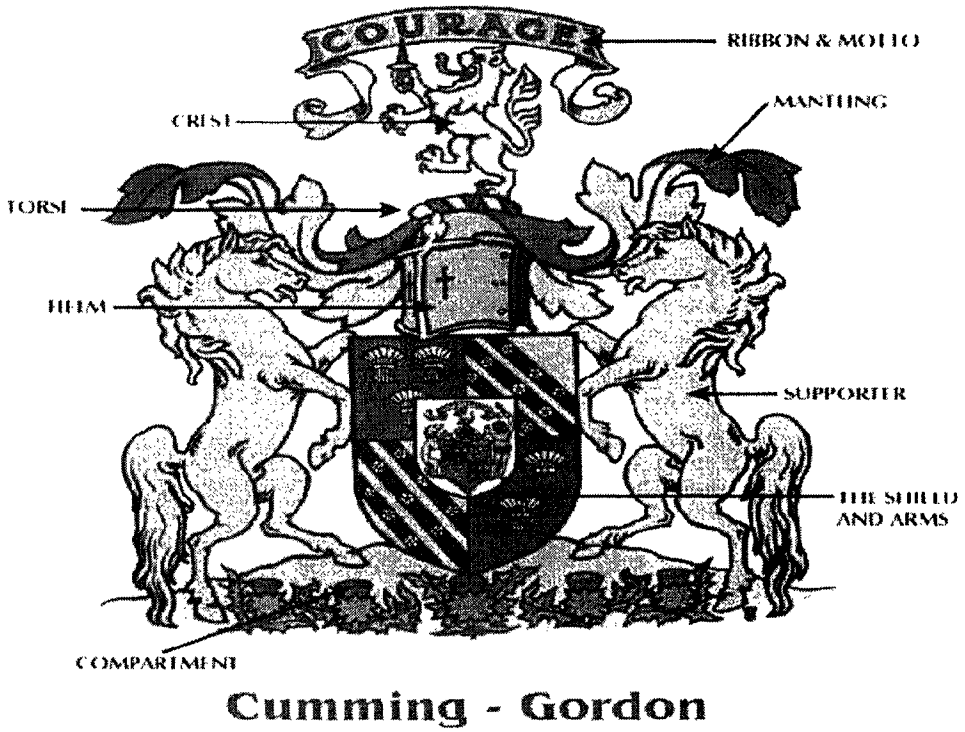

Figure 1-3
Parts of a coat of arms.
Source: <http://www.designsofwonder.com/articles/shieldparts.html>

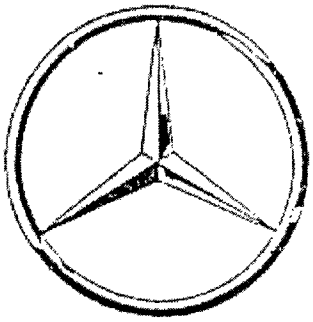

Figure 1-4
The trademark of Mercedes-Benz.
Source: <http://www.mercedes-benz.de/>

Figure 1-5
The trademark of Tac-Software.
Source: <http://www.tac-soft.com/>

Figure 1-6
The trademark of Metro-Goldwyn Mayer.
Source: < http://www.mgm.com/>

Figure 1-7

The trademark of the watchmaker Breitling.
Source: Der Spiegel .No.34/20.08.01, 19.

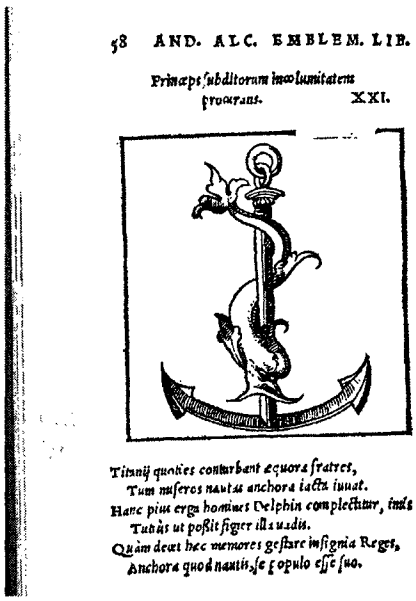

Figure1-8

The anchor and the dolphin.
Source: Alciatus, Andreas. Emblematum Libellus. 58.

Figure 1-9

Oculis vita.
Source: Schöne, Albrecht. Emblematik und Drama im Zeitalter des Barock. 67.

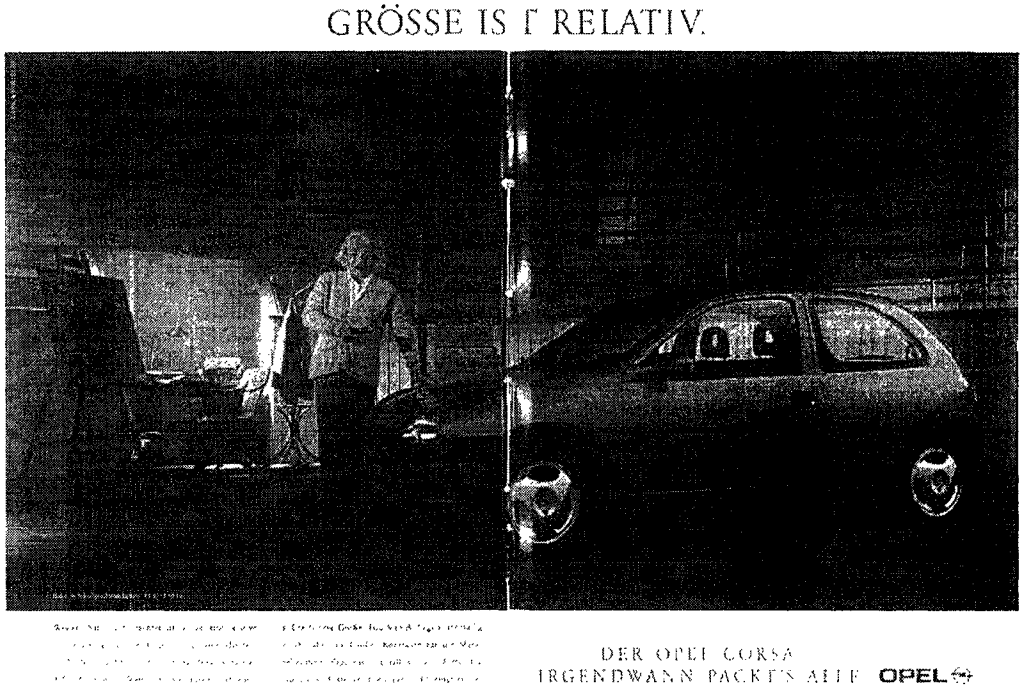

Figure 1-10

Einstein and the Corsa.
Source: Der Spiegel Extra: Das Kultur-Magazin. Heft1, January, 1997, 24-25.

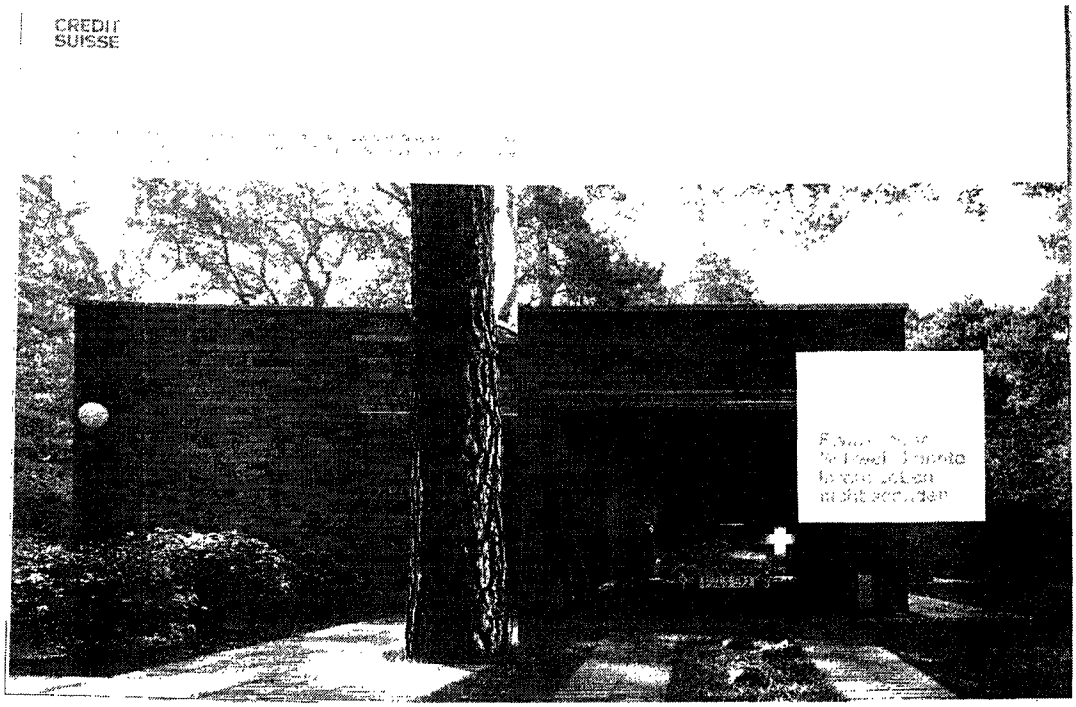

Figure 1-11

Etwas mehr Schweiz.
Source: Der Spiegel. No. 34/20.8.01, 10-11.

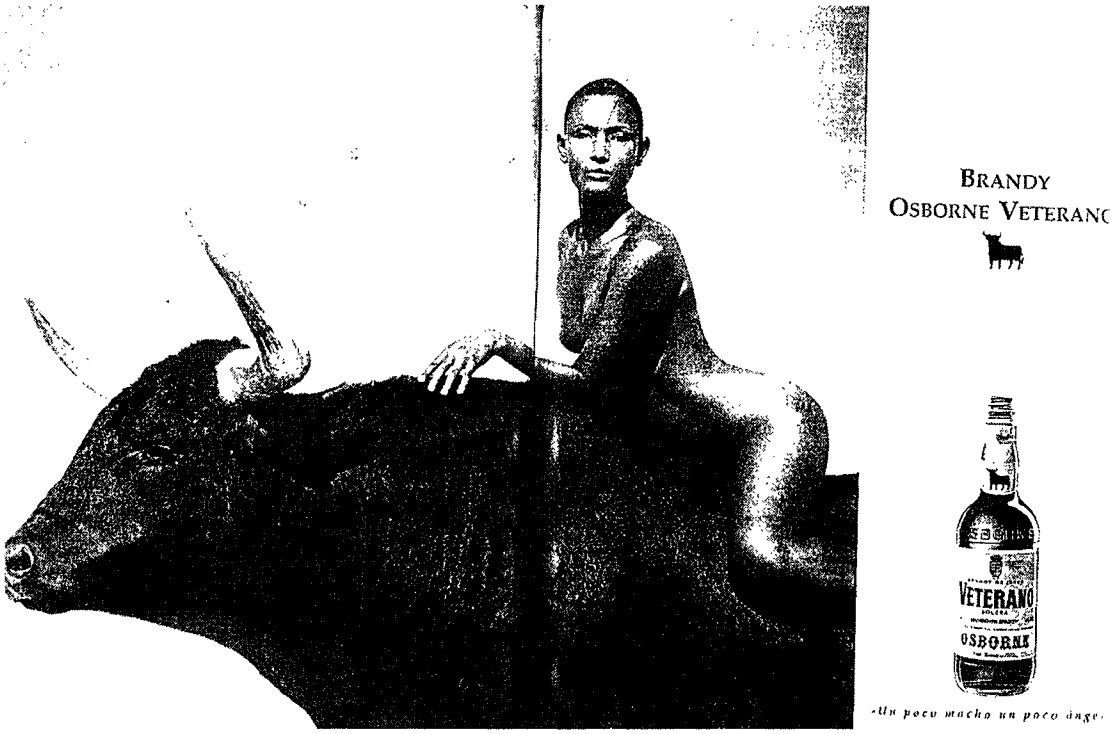

Figure 1-12

The young lady and the bull.
Source: Der Spiegel. No. 7/13.02.95, 8-9.

Figure 1-13

The young lady and the unicorn.
Source: Schöne. Emblematik und Drama im Zeitalter des Barock. 110.

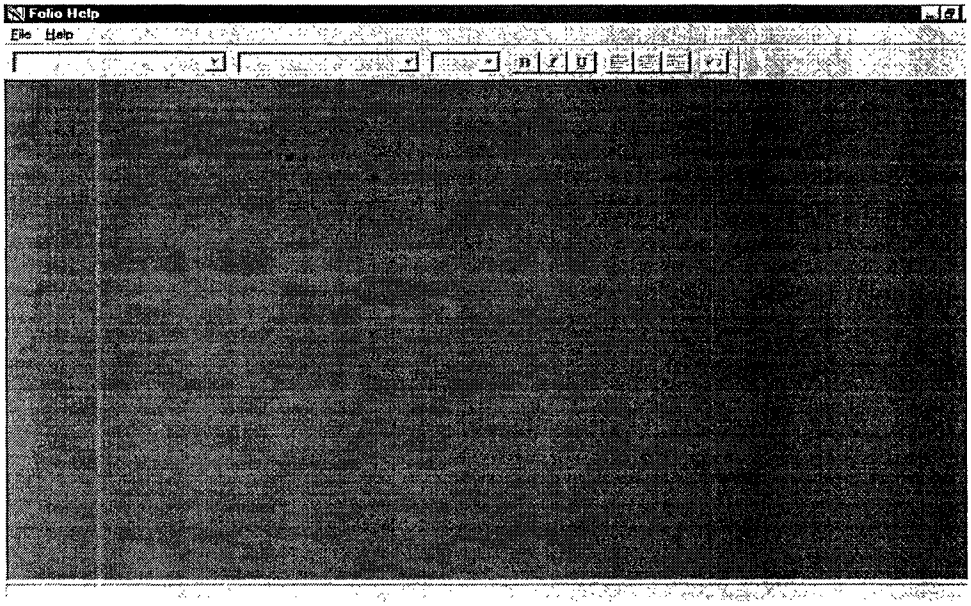

Figure 6-1

The *help* window of Studiolum.

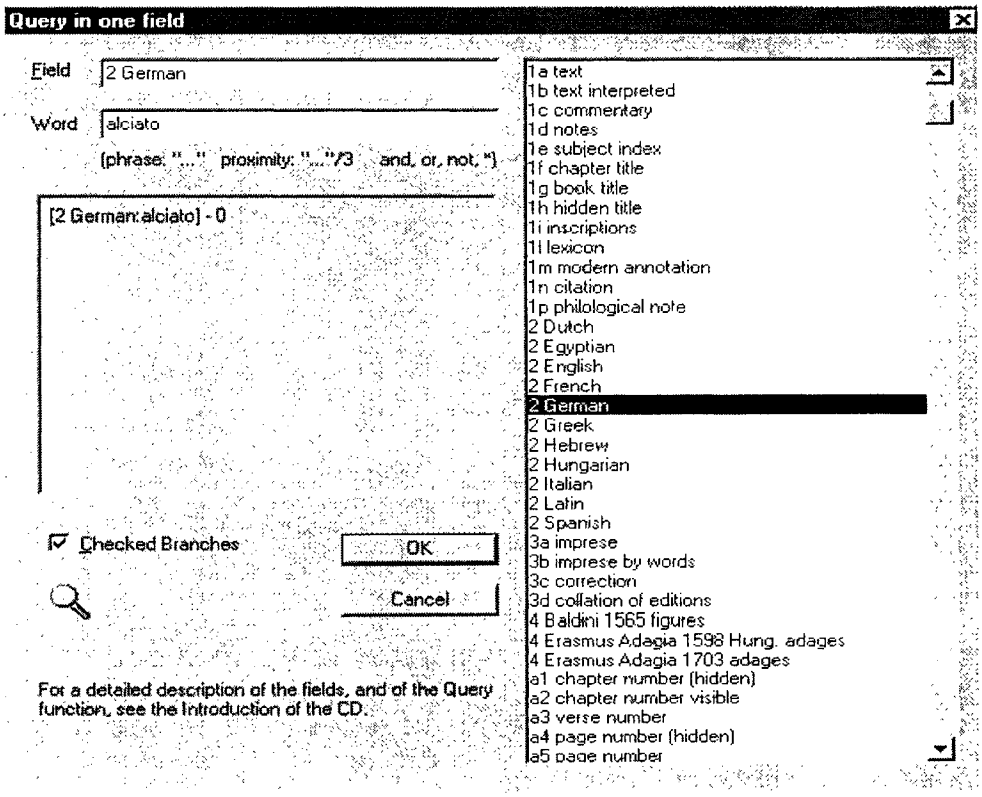

Figure 6-2

The *simple search* window offered by Studiolum.

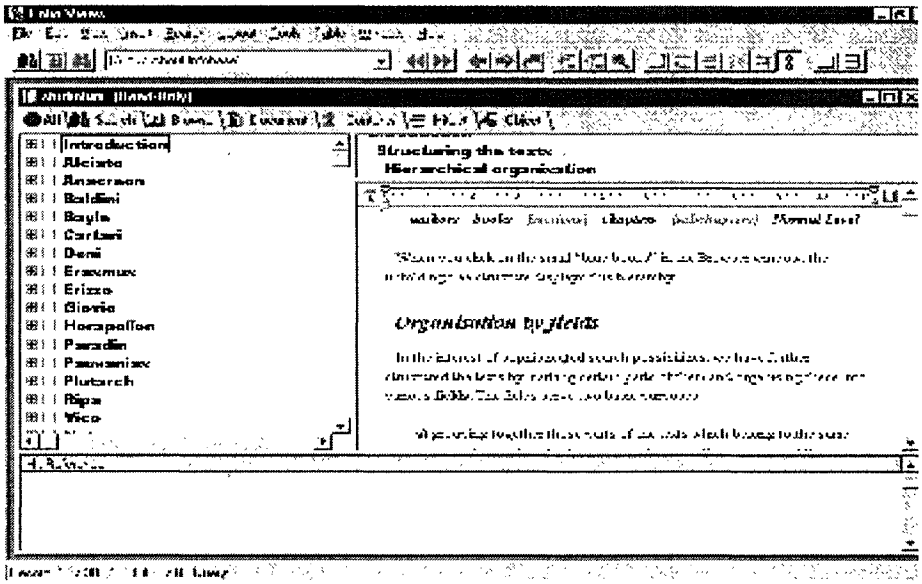

Figure 6-3

The *main interface* of Studiolum.

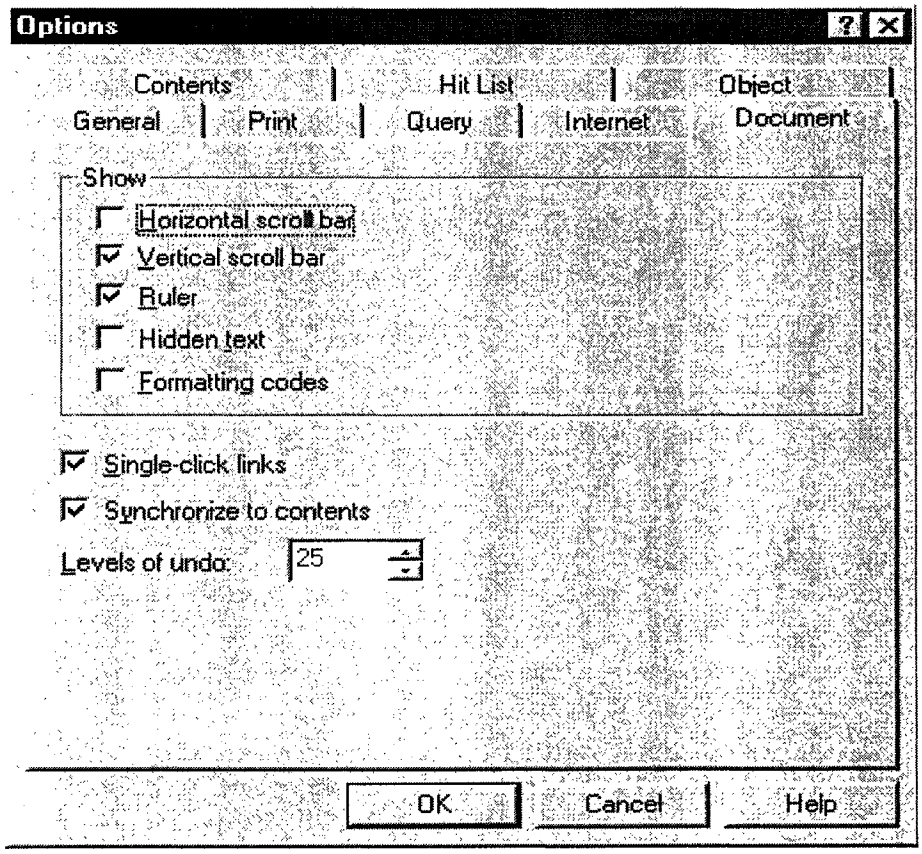

Figure 6-4

The *option* dialog box.