Digitisation Fundamentals and their Application

Robin Davies
Michael Nixon
Welcome to DHSI 2012!

Thanks for joining the DHSI community this year – our 10th year, but our 11th offering!

In this booklet, you will find essential course materials prefaced by some useful information about getting settled initially at UVic, finding your way around, getting logged in to our network (after you’ve registered on the Sunday afternoon, and received your login information), and so on.

Given our community’s focus on things computational, it will be a surprise to no one that we might expect additional information online for some of the classes – your instructors will let you know – or that the most current version of all DHSI-related information may be found on our website at dhsi.org. Do check in there first if you need anything that’s not in this coursepak.

And please don’t hesitate to be in touch with us at institut@uvic.ca or via Twitter at @AlyssaA_DHSI or @DHInstitute if we can be of any help ....
Daily Schedule

Sunday, 3 June 2012 [DHSI Registration]

4:00-6.00
DHSI Registration
At UVic Housing / Residence Services Office (Craigdarroch Building)
See the University of Victoria @ Google Maps

After registration, many will wander to Cadboro Bay and the pub at Smuggler's Cove.

Monday, 4 June 2012

8:00 to 8:30
Last-minute Registration
MacLaurin Building, Room A100
See the University of Victoria @ Google Maps

8:30 to 9:30
Welcome
MacLaurin A144

Classes in Session (Class locations for the week are as listed below)
1. Text Encoding Fundamentals and their Application (Clearihue A102)
2. Digitisation Fundamentals and their Application (Clearihue A015)
3. Introduction to XSLT for Digital Humanists (Clearihue A103)
5. Geographical Information Systems in the Digital Humanities (Human and Social Development A170)
6. Physical Computing and Desktop Fabrication for Humanists (MacLaurin D016)
7. Digital Pedagogy in the Humanities (MacLaurin D110)
8. Creating Digital Humanities Projects for the Mobile Environment (Human and Social Development A270)
9. Designing RESTful APIs (Application Programming Interfaces) (MacLaurin D115)
10. Digital Humanities Databases (MacLaurin D114)
11. Augmented Reality: An Introduction (MacLaurin D109)
12. Issues in Large Project Planning and Management (Hickman 120)
13. Digital Editions (Clearihue A012)
14. Out-of-the-Box Text Analysis for the Digital Humanities (Clearihue A105)
15. Understanding the Pre-Digital Book (McPherson Library A003, A130)
16. Online Tools for Literary Analysis (MacLaurin D010 (M-W), D111 (Th-F))
17. SEASR Analytics (MacLaurin D107)

Noon to 1:15
Lunch break / Unconference
MacLaurin A144
(Unconference discussions through the week are coordinated by Deb Raftus; discussion topics, scheduling, and room assignments from among all DHSI rooms will be handled at this meeting)

1:15 to 3:50
Classes in Session

4:00 to 5:00
Institute Lecture: Laura Mandel (Texas A&M)
MacLaurin A144

5:00 to 6:00
Light Reception
University Club

Tuesday, 5 June 2012

8:00 to 9:20
DHSI Colloquium: Textual Analysis
MacLaurin A144

9:30 to Noon
Classes in Session
<table>
<thead>
<tr>
<th>Time</th>
<th>Event</th>
</tr>
</thead>
<tbody>
<tr>
<td>Noon to 1:15</td>
<td>Lunch break / Unconference, various locations</td>
</tr>
<tr>
<td>1:15 to 3:50</td>
<td>Classes in Session</td>
</tr>
<tr>
<td>4:00 to 5:30/6.00</td>
<td>DHSI Colloquium: E-Publishing and Digital Editions</td>
</tr>
<tr>
<td>8:00 to 9:20</td>
<td>DHSI Colloquium: Digital Pedagogy</td>
</tr>
<tr>
<td>Noon to 1:15</td>
<td>Lunch break / Unconference, various locations</td>
</tr>
<tr>
<td>1:15 to 3:50</td>
<td>Classes in Session</td>
</tr>
<tr>
<td>4:00 to 5:30/6.00</td>
<td>DHSI Colloquium: Archives and Databases</td>
</tr>
<tr>
<td>8:00 to 9:20</td>
<td>DHSI Colloquium: Gaming, Gamification, and Media Studies</td>
</tr>
<tr>
<td>Noon to 1:15</td>
<td>Lunch break / Unconference, various locations</td>
</tr>
<tr>
<td>1:15 to 3:50</td>
<td>Classes in Session</td>
</tr>
<tr>
<td>4:00 to 5:30/6.00</td>
<td>DHSI Colloquium: Mapping and Visualization</td>
</tr>
<tr>
<td>8:00 to 9:20</td>
<td>DHSI Colloquium: The Way Forward</td>
</tr>
</tbody>
</table>

**Wednesday, 6 June 2012**

- **DHSI Colloquium: E-Publishing and Digital Editions**
  - Gord Barentsen, "Coding Digital Texts in TEI-Compliant XML, Case Study: The Folger Library Edition of *Troilus and Cressida*"
  - Carol Lea Clark, "The Future of the eBook: A Medium for Scholarly Innovation"
  - Constance Crompton, Daniel Powell and Ray Siemens, "The Devonshire Manuscript Defused: Modeling the Social Edition"
  - Jon Detombe, "Digitizing and Deciphering the Courten MS"
  - Elias Fahssi, "Working with Patience: Textual Editing, Digital Humanities, and Undergraduate Research"
  - Linnet Humble, "eBook Production at Canadian UPS: Problems and Possible Solutions"
  - Anna Gibbs and Maria Angel, "Digital Writing and the Literary Ethos"
  - Jeffrey Witt, "Rethinking the Critical Edition: The 'Editio Critica Electronica' of Petrus Plaouf"

**Thursday, 7 June 2012**

- **DHSI Colloquium: Digital Pedagogy**
  - Almond Aguila, "The Pedagogy of Facebook"
  - Kathi Inman Berens, "Failure is Frictive: Coding and Pedagogy"
  - Eugenie Duthoit, "Re-thinking the Use of Digital Tools to Assist the Pedagogical Translation from Latin"
  - Chris Friend, "Bringing Technology to Student Writing: How DH Practices Can Enhance Composition Pedagogy"
  - Peggy Jubien, "Reexamining Our Tools: Linking Educational Technology to the Socio-Political Dimensions"

**Friday, 8 June 2012 [DHSI + Beyond Accessibility]**

- **DHSI Colloquium: The Way Forward**
MacLaurin A144
Christopher Church, Hannah Farber and Scott McGinnis: "Computing and the Practice of History: Creating an Institutional Framework for the Digital Humanities at UC Berkeley"; Kristin Cornelius, "Disseminating the Humanities"; Brooke Lestock and Sarah Storti, "The Praxis Program and Prism: Rethinking Graduate Training in a Digital Age"; Daniel Powell, Alyssa Arbuckle, Alyssa McLeod and Shaun MacPherson, "Digital Humanities and the Alt-Ac 'Track': Views from the Grad School Trenches"; Robin Wharton, "How to Make a Digital Humanist"

9:30 to Noon
Classes in Session

Noon to 1:15
Lunch break / Unconference, various locations
Registration for conference, Beyond Accessibility: Textual Studies in the 21st Century (8-10 June)
MacLaurin Building, Room A100
See the University of Victoria @ Google Maps

1:15 to 2:15
Institute Lecture: Adriaan van der Weel (Leiden)
Co-sponsored by Beyond Accessibility
MacLaurin A144

2:15-4:30 (or so)
DHSI Wrap-up Session, Show and Tell
MacLaurin A144

Saturday, 9 June 2012 [Beyond Accessibility]
Beyond Accessibility: Textual Studies in the 21st Century
Please visit the conference website

8.30-5:00
Room 1 (MacLaurin D101)
Room 2 (MacLaurin D103)
Plenary Session (Hickman 105)
Breakout / Meeting Room (Hickman 120)

6.30-8:30
TBA, Reception / Dinner

Sunday, 10 June 2012 [Beyond Accessibility]
Beyond Accessibility: Textual Studies in the 21st Century
Please visit the conference website

9.00-12:45
Room 1 (MacLaurin D101)
Room 2 (MacLaurin D103)
Plenary Session (Hickman 105)
Breakout / Meeting Room (Hickman 120)

Contact info:
institut@uvic.ca P: 250-472-5401 F: 250-472-5681

RELATED LINKS: university of victoria humanities computing and media centre electronic textual cultures lab
Regional Map of Greater Victoria

Map Scale

2 km
2 miles
NEW UVic wireless configuration utility

The UVic wireless configuration utility will automatically configure the "UVic" wireless network on your Windows XP SP3, Windows Vista, or Windows 7 computer.

**Download now**

*Note:* The UVic wireless configuration utility is still experimental; use this application at your own risk. UVic is not responsible for any damage caused by the use of the wireless configuration utility. Please report any problems to the Computer Help Desk.

If the above doesn’t work, please follow the manual instructions listed below. After the initial configuration, you should automatically connect to **UVic** (the secure wireless network) when you are on campus.

1. Before you start this procedure, ensure the following:
   - Your wireless card and its drivers have been installed and you have rebooted your laptop since the installation.
   - Your laptop is powered on and booted up.
   - You are in an area with wireless coverage.
   - You have a NetLink ID and password.
   - You are using Windows to manage your wireless connections. If you are using a third-party application (sometimes network adaptors come with their own applications), you may experience problems during the configuration process.

2. Temporarily connect to the Internet using **UVic Open**, an Ethernet port, or your home network. Download the security certificate by right clicking thawte Primary Root CA and saving the thawte.cer file to your computer. Once the file is saved to your computer, locate the file, double click on it, select Install Certificate..., and follow the Certificate Import Wizard instructions.
3. Once you have successfully installed the certificate, open your **Start** menu and click on **Control Panel**.

4. Click on **Network and Internet** or **Network and Sharing Center**.
5. Click on **Network and Sharing Center**.

6. Click on **Manage wireless networks**, located on the left menu.

7. Click **Add**.
8. Click **Manually create a network profile**.

9. Enter the following information:
   - Network name: **UVic** (case sensitive).
   - Security type: select **WPA2-Enterprise**.
   - Encryption type: automatically sets to **AES**.
   - Security Key/Passphrase: (leave blank).
Ensure both checkboxes are selected (by default, the second box is not). Click Next.

10. Click Change connection settings. For now, ignore the pop-up window in the bottom-right corner.

On the Connection tab, ensure the Connect to a more preferred network if available checkbox is not checked.

11. Click the Security tab. Ensure the authentication method is PEAP. Then click Settings.
12. Check the box beside thawte Primary Root CA in the list of Trusted Root Certification Authorities.

If you cannot find the correct certificate listed, please return to step 2 to download the certificate.

At the bottom of the dialogue, ensure that the Authentication Method is Secured.
password (EAP-MSCHAP v2). Click Configure.
13. Deselect the checkbox for Automatically use my Windows logon... and click OK.

14. Close the remaining windows. In the bottom-right corner of your screen, you should see a small window pop-up informing you that Additional information is required to connect to UVic. Click on it to provide additional information.

15. Enter your personal NetLink ID followed by @uvic.ca in the User name field, and your NetLink ID password in the Password field. Click OK.

You should now be connected to the UVic secure wireless network.
After the initial configuration, you should automatically connect to UVic (the secure wireless network) when you are using UVic’s wireless network.

1. Before you start this procedure, ensure the following:
   - Your wireless card and its drivers have been installed and you have rebooted your laptop since the installation.
   - Your laptop is powered on and booted up.
   - You are in an area with wireless coverage.
   - You have a NetLink ID and password.

2. At the top-right corner of your screen there should be the AirPort icon (a semi-circle). If you do not see this icon, your AirPort card or AirPort software may not have been installed properly.

3. Click on the AirPort icon (it may be partially darkened) to reveal a menu. Ensure your AirPort is On.

4. Scroll down the AirPort menu and select Join Other Network ...

5. In the window that opens, enter the following information:
   - Network Name: UVic (case sensitive)
   - Security: WPA2-Enterprise
   - User Name: your NetLink ID
   - Password: your NetLink ID password
   - 802.1X: Automatic

Click Join.
6. If you see a message about Mac OS X wanting to access your Keychain, click *Always Allow*.

7. A *Verify Certificate* window will open saying that the certificate is not trusted.
   - Click *Show Certificate*.
   - Check the box that says *Always trust "sac1cled050..."* (the exact name may vary) and click *Continue*.
   - If you are prompted for your computer password, enter it and click *OK*.
You should now be connected to the UVic secure wireless network. To disconnect from the wireless network, click on the AirPort icon and click Turn Airport Off. Next time you connect to UVic, you should not need to enter any additional credentials.
Connect to UVic: iPhone or iPod Touch

After the initial configuration, you should automatically connect to UVic (the secure wireless network) when you are using UVic’s wireless network.

1. Before you start this procedure, ensure the following:
   - Your device is using firmware version 4.0 or higher.
   - Your device is powered on and booted up.
   - You are in an area with wireless coverage.
   - You have a NetLink ID and password.

2. From the Home screen, press the Settings button.
3. Press the Wi-Fi option.
4. Under the Choose a Network... heading, select UVic.
5. Enter your personal NetLink ID followed by @uvic.ca in the Username field. Enter your NetLink ID password in the Password field. Press Join.

6. If prompted, press Accept to verify the thawte Primary Root CA certificate.
Your device should now be connected to the **UVic** secure wireless network.
DIGITISATION FUNDAMENTALS AND THEIR APPLICATION
Table of Contents

SECTION 1: PROJECT PLANNING...........................................................................................................5
Planning for Digitisation Projects........................................................................................................6
  Making the plan.................................................................................................................................6
  What will the final project look like?................................................................................................7
Assessing your Digitisation Project ......................................................................................................8
What is your presentation medium?....................................................................................................8
What is your long-term plan for this project?......................................................................................9
How will you coordinate and track your decisions?...........................................................................9
How will you store or present your data?...........................................................................................9
The Eight Basic Principles of Project Management.............................................................................10

SECTION 2: DIGITISING IMAGES .......................................................................................................12
  Best Practices for Digitising Images ................................................................................................13
  Basic Rules for Digitising Images ....................................................................................................13
  Common Terms for Imaging .............................................................................................................14
  Common Image File Types .............................................................................................................16
  Imaging for the Web .........................................................................................................................18
  Imaging for Print .............................................................................................................................19
  Imaging for Archival Purposes ........................................................................................................20
Resources for Digital Imaging.............................................................................................................22
  General Info .....................................................................................................................................22
  Metadata Resources.........................................................................................................................23
Imaging Tutorial 1: Flatbed Scanning .................................................................................................24
Imaging Tutorial 2: Digital Photography ............................................................................................27
Imaging Tutorial 3: OCR Scanning ......................................................................................................30
Imaging Tutorial 4: Using Photoshop ................................................................................................31
  Cropping and Rotating ....................................................................................................................31
  Adjusting Your Image – Colour and Contrast ................................................................................31
  Using Levels to Adjust Brightness and Contrast ............................................................................32
  Using the Levels Tool to Improve Colour Balance ........................................................................34
Imaging Tutorial 5: Using Layers in Photoshop ................................................................................35

SECTION 3: DIGITISING AUDIO .........................................................................................................38
Best Practices for Digital Audio............................................................................................................39
  Basic Rules for Digitising Audio Files ............................................................................................39
  Common Audio File Types .............................................................................................................40
  Transferring Analog Audio Sources to the Computer ....................................................................42
  Audio Standards for Recording Audio Files ..................................................................................43
  Archival Standards for Digitising Audio Files ..............................................................................43
  Optimizing Your Computer for Recording Digital Audio ...............................................................44
Resources for Digital Audio ..............................................................................................................45
  Software Packages ..........................................................................................................................45
  Audio Standards Websites .............................................................................................................46
<table>
<thead>
<tr>
<th>Audio Tutorial 1 – Getting Started</th>
<th>47</th>
</tr>
</thead>
<tbody>
<tr>
<td>Input</td>
<td>47</td>
</tr>
<tr>
<td>Recording Audio</td>
<td>47</td>
</tr>
<tr>
<td>&quot;Ripping&quot; Audio from CDs</td>
<td>50</td>
</tr>
<tr>
<td>Opening a File</td>
<td>51</td>
</tr>
<tr>
<td>Saving a Project</td>
<td>51</td>
</tr>
<tr>
<td>Exporting a File</td>
<td>52</td>
</tr>
<tr>
<td>Process</td>
<td>53</td>
</tr>
<tr>
<td>Selecting and the Editing Tool</td>
<td>53</td>
</tr>
<tr>
<td>Copying and Pasting</td>
<td>54</td>
</tr>
<tr>
<td>Cutting</td>
<td>55</td>
</tr>
<tr>
<td>Deleting</td>
<td>55</td>
</tr>
<tr>
<td>Undo, Redo, and Versioning</td>
<td>56</td>
</tr>
<tr>
<td>Output</td>
<td>56</td>
</tr>
<tr>
<td>Burning CDs or DVDs</td>
<td>56</td>
</tr>
<tr>
<td>Audio Tutorial 2 - Advanced Audio Editing</td>
<td>58</td>
</tr>
<tr>
<td>Pasting</td>
<td>58</td>
</tr>
<tr>
<td>Mixing</td>
<td>60</td>
</tr>
<tr>
<td>Silencing Unwanted Sounds</td>
<td>61</td>
</tr>
<tr>
<td>SECTION 4: DIGITISING VIDEO</td>
<td>64</td>
</tr>
<tr>
<td>BEST PRACTICES FOR DIGITAL VIDEO</td>
<td>65</td>
</tr>
<tr>
<td>Digital Video</td>
<td>65</td>
</tr>
<tr>
<td>Shooting Video</td>
<td>65</td>
</tr>
<tr>
<td>Resources for Digital Video</td>
<td>68</td>
</tr>
<tr>
<td>(Some) Codecs Described</td>
<td>69</td>
</tr>
<tr>
<td>SECTION 5: WEB DESIGN</td>
<td>70</td>
</tr>
<tr>
<td>Best Practices in Web Design</td>
<td>71</td>
</tr>
<tr>
<td>The World Wide Web and Digitisation</td>
<td>71</td>
</tr>
<tr>
<td>The Basics of HTML</td>
<td>71</td>
</tr>
<tr>
<td>The Building Blocks</td>
<td>72</td>
</tr>
<tr>
<td>Getting Started</td>
<td>72</td>
</tr>
<tr>
<td>Tag, you're it</td>
<td>73</td>
</tr>
<tr>
<td>Angle brackets</td>
<td>74</td>
</tr>
<tr>
<td>Nesting tags</td>
<td>75</td>
</tr>
<tr>
<td>So many tags, so little time</td>
<td>77</td>
</tr>
<tr>
<td>More Tags</td>
<td>77</td>
</tr>
<tr>
<td>The framework of a web page</td>
<td>79</td>
</tr>
<tr>
<td>Validating Your Code</td>
<td>80</td>
</tr>
<tr>
<td>Coding Exercise</td>
<td>81</td>
</tr>
<tr>
<td>Building Websites</td>
<td>82</td>
</tr>
<tr>
<td>Web design: theory and practice</td>
<td>82</td>
</tr>
<tr>
<td>Paper Prototyping</td>
<td>82</td>
</tr>
<tr>
<td>Using Templates</td>
<td>83</td>
</tr>
<tr>
<td>Web Tutorial 1: Image enhanced web design</td>
<td>87</td>
</tr>
</tbody>
</table>
SECTION 1: PROJECT PLANNING

Prepared by Robin Davies, Gerry Watson, Karin Armstrong
Planning for Digitisation Projects

A plan is essential for any digitisation project. A project plan will help you anticipate problems and needs, track your progress through the material, guide future development, and coordinate resources and staff. Your plan can be as flexible or rigid, detailed or vague as you like, but no matter what form it takes, in the long run making a plan will save you time, headaches, and resources.

Project planning can take many forms and encompass varying ranges of complexity. For our purposes, we'll keep it simple. Even if you are an experienced project manager, there are a few things to keep in mind when planning for any digitisation project.

Making the plan

There are a number of software packages that can help you build your project plan, track resources, and offer a host of other tools and solutions for large project management. While these software tools can be useful, their effectiveness extends only so far as the vision and detail you supply. First, establish broad goals. Aim for an idea of what you want the final product to look like; then you can begin filling in the details. Once you have addressed some of the concerns below, you will be able to consciously plan your project. Remember, you don't have to focus on every single detail the first time through. It is sometimes best to think through your project at different scales, using an iterative approach.

For example, begin by thinking of major goals and the final product, and then break the project into major sections or phases; each phase can then be broken down into smaller goals, and those goals into individual tasks.

This is also a good point to get your team involved in the growth of the plan, as they will have insights into how the work will or should progress. This approach allows your team to take ownership of their specific tasks, and take part in setting deadlines. Extending the decision-making process to your team takes the pressure
off of you to micro-manage the project, and allows them some control over, and responsibility for, their own work.

The plan can take many forms and serve multiple functions. The important thing to remember is to make a plan, revise it as you need to, and try to stick with it. This simple process will help organize your thoughts and give you a clear framework for progress in meeting your project’s unique needs.

**What will the final project look like?**

First, decide what you want to accomplish in general terms. Do you want an online exhibition or portfolio of art, literature, cartography, or historical materials? Is the end product of your digitisation going to be some form of print? Do you want to build a research tool as a software package or an online archive for public use? Is the project for private or institutional use only? All these questions guide what the product should be, and lead to vastly different requirements in terms of digitisation, resources, equipment, staff, storage, and a host of other minutiae that should at least be acknowledged before you begin. Even with a vision and a plan, things will inevitably go wrong, but at least you can be prepared for contingencies, which will enable you to react more appropriately and quickly.
Assessing your Digitisation Project

A vital part of any digitisation project is project assessment. Some initial decisions will streamline the process, help you decide how to proceed, and most effectively dictate the use of available resources. Here are some questions to consider:

What is your presentation medium?

Will your final project include audio, images, text, video, HTML, or some combination of these? Making distinctions between applicable media will help to define and focus your learning task. With each medium, one must gain relevant vocabulary and knowledge, and learn applicable skills.

For example, if your project involves scanning images for use on a web page, it is important to know that you may be limited in your choice of file types (JPG, GIF, and PNG). You’ll also need to understand pixels, compression, and bandwidth.

This knowledge will be integral in the day-to-day work of your project, and will also prove useful in the initial stages of your project as you make staffing and equipment decisions. The equipment you choose will dictate the standards you set. Whether you use equipment or facilities already in place at your institution, or secure funding to purchase dedicated machines, you should educate yourself in the limits and benefits each option represents. When purchasing new equipment, do not rely on what the salesperson says. Research the scanner, computer system, camera, and printer that you will come to depend on. Only you truly understand what you want out of these machines. If the funding is available, it is best to purchase a machine that does more than what you need. At the outset of your project, it can be tough to foresee all your needs, so choose a piece of equipment that does what you need, but is also capable of more advanced functions should they become necessary.

If you know you want to scan in archival images at 400 dpi, it might be a good idea to buy a scanner that can go far beyond this and that can also scan slides and microfilm, in case you need access to those media. Similarly, don’t buy a printer that produces a minimum quality document when someday you may need it to produce
high quality advertisements for your project. If you have the option, make wise purchases based on your project's immediate needs, but with an eye to solving future problems as well. Be prepared!

**What is your long-term plan for this project?**

Decide what your final output will be: web display, an archive, digital projection, storage at home, long-term pedagogy, and so on. What are the procedural implications of this projected outcome? Knowing how you will access and present your final data will go a long way in deciding your digitisation process.

For example, if you are planning to archive your data for long-term storage, be certain to choose the best storage format possible, one which will ensure long-term viability and versatility. Generally, this means large, lossless file formats like TIFF (for images) and AIFF (for audio). Or, for a project destined for web or classroom use, choose a smaller, portable, perhaps lossy format such as JPG (for images) and MP3 (for audio).

**How will you coordinate and track your decisions?**

Project management is critical to the success of larger initiatives. It is recommend, for example, to maintain an up-to-date project summary document which outlines the chronology of decisions made and work completed helping to keep the project on-task and moving forward. Other valuable management practices include establishing consistent naming conventions for documents - title, location, date captured, index - building a well-organized directory structure, and keeping track of metadata.

**How will you store or present your data?**

Storage: burning data to CD or DVD is an inexpensive and convenient choice, providing a spacious, transportable and safe storage solution. A USB flash storage device is also a possibility. Safer still would be a portable or internal hard disk, or a third party server. Ideally, you should try to establish several locations for data
storage, with an established and consistent backup process. Overall, your best bet is to diversify. With larger archival projects, long-term planning will involve establishing a safe, permanent location for your master data.

Presentation: most commonly, a project will be displayed on the web or in a classroom. This often requires versatile file formats and reduced storage space. In this case, aim for clarity and convenience rather than posterity.

**The Eight Basic Principles of Project Management**

When writing your project plan and while your project is underway, remember these simple principles of project management:

No major project is ever installed on time, within budget, with the same staff that started it. Yours will likely not be the first.

Projects progress rapidly until they become 90 percent complete; they then remain 90 percent complete forever.

One advantage of fuzzy project objectives is that they let you avoid the embarrassment of estimating the corresponding costs.

When things are going well, something will go wrong. When things just can’t get any worse, they will. When things appear to be going better, you have overlooked something.

If project content is allowed to change freely, the rate of change will exceed the rate of progress.

No system is ever completely debugged; attempts to debug a new system inevitably introduce new bugs that are even harder to find.

A carelessly planned project will take three times longer to complete than you expected; a carefully planned project will only take twice as long.

Project teams detest project reports, because these reports vividly manifest their lack of progress.

([http://www.daclarke.org/Humour/projects.html](http://www.daclarke.org/Humour/projects.html))
SECTION 2: DIGITISING IMAGES

Prepared by Robin Davies, Cara Leitch, Gerry Watson, Karin Armstrong, Anne Correia, Trish Irish, and Graham Lyons
**Best Practices for Digitising Images**

When beginning a digital imaging project, the first goal is to determine how your images will be used. Images designed for the web, for print, or for archival purposes are all very different. Below are a few basic guidelines for any project, and some specific pointers for particular kinds of projects. Some basic concepts are discussed only briefly here; please refer to the online links for more information.

**Basic Rules for Digitising Images**

**Rule 1:** You should always save a master image in a *lossless* format.

This means either TIFF or PNG should be set as your default image format when capturing or scanning images. While your final project may require JPGs, every time you alter and save a JPG, information is lost and quality decreases. Your image quality will be much better if you edit your master image as a TIFF and then convert to JPG in the final stages of your project.

**Rule 2:** Start with a high quality image, and then decrease the resolution (and perhaps increase the compression) to suit your project.

You cannot take a small JPG and turn it into a high-quality print. But you can take a high-quality printable image and turn it into a small JPG. It’s far easier to scan at a higher resolution than you think you need, and then manipulate the image for your project. Target a higher level of image quality for your scans than you may need for your final project.

**Rule 3: Manage your images.**

It’s hard to remember the difference between imgJune.jpg and picture_5.png. Make the effort to name your images meaningfully from the beginning. You may also want to add an index number for easy reference in large projects. Take the time to consider how to organize your images for your project. A great image is useless if you can’t find it.
**Common Terms for Imaging**

**CMYK**
Like RGB (see below), CMYK, which is an acronym for Cyan, Magenta, Yellow, and Key (black), represents the ratio and relative intensity of colours in an image, specifically for print display. Most colour printers use this scheme. The use of Black here means the more expensive inks can be used sparingly to obtain dark colours.

**RGB**
Red-Green-Blue colour attempts to represent - for a computer monitor or data projector - a colour in numerical form. By defining a level of intensity (regardless of how that intensity is measured) for red, green, and blue, one can create a full colour palette for digital display.

**HSB**
The Hue, Saturation, and Brightness model is perhaps the most human way to describe colour. Hue refers to the value of the colour, as might be chosen from a colour wheel or rainbow. Saturation is the intensity of the colour; colours will low saturation will appear gray, faded. Brightness measures the amount of "light" shining on the colour; very low brightness settings make all colours look black.

**Grayscale**
Grayscale is a mode of displaying images with no colour, in shades of gray. Removing the colour information from an image can make for useful photographic effects, and can decrease the file size of the image. We often erroneously call grayscale images black and white images. In addition, when someone asks for a black and white image, they often are really looking for a grayscale image tinted with a sepia colour.
Pixel

The term pixel is a computer abbreviation for picture element, the smallest element of a digital image or digital display.

dpi

Dots per inch is a term often used to quantify the (maximum) resolution capabilities of printers, scanners and monitors. The more dots per inch, the higher the resolution: 600 dpi would mean 600 x 600 = 360000 dots per square inch. It is important to understand this term when converting from analog to digital, or from digital to analog. Within the digital system (monitors, web display, digital projection), the term "inches" is irrelevant, but we must keep in mind how the "dots" on the printed page will translate to "pixels" on the video screen. In general, remember that a higher dpi means a larger image on the video screen or in the web browser, and a larger the file size for storage.

ppi

Pixels per inch. This term is directly related to dpi - in fact, the two terms are often used interchangeably. Pixels per inch refers to the relationship between video display (digital) and print display (analog). While dpi indicates the number of ink dots for every inch of physical space on the printed page, ppi denotes the equivalent level of detail on a video screen, the number of pixels for each inch of display. However, it is important to remember that the physical display size of an image on a video screen is partly dependent on the resolution of the screen. A 400 pixel-wide picture displays across half of an 800x600 display, but only about a third of a 1280x854 display. The higher the ppi, the larger the image will appear on a monitor. For example, if you were to scan a 4x6 photo with different dpi settings, your image size would display differently on the same monitor; a 4x6 photo scanned at 600 ppi/dpi would display a digital image twice as large as the same photo scanned at 300 ppi/dpi.
Pixelation

When you enlarge a small image, you may, depending on the resolution, encounter pixelation. This term describes the effect of individual (usually square) pixels being visible to the eye. This problem often occurs when digital images are archived incorrectly, stored with a lossy file type like JPG or GIF, or when an image is frequently resampled or colour-altered. You’ll likely run into this problem if you try to appropriate a web image for use in a print situation.

Scaling or Resampling

Resampling refers to changing the size of an image for digital display purposes (monitor, digital projector, or the web). Resampling alters an image significantly, engaging various compression algorithms to fit an image to certain dimensions, such as 800x600. Because each pixel is changed in order to fit new parameters, you might encounter a loss of quality, often because the resampling process adds or removes data from the image through interpolation.

Resolution

The term resolution is applied to digital images, print objects, and monitor displays; it reflects, in general, the “granularity” of an image. Resolution is expressed in dpi for print media, in ppi for visual display, and in pixel dimensions for monitor resolution: 800x600.

Common Image File Types

The list below includes only some of the many image file formats available.

BMP

Bitmap. This is used for bitmap graphics, usually on the Windows OS. BMP is a lossless file format, so files can be very large.
**GIF**

Graphics Interchange Format. This is a very common format on the web, usually for logos or illustrations with few colours. Because of the way this format handles colour variation, photographic colour transitions are often jagged. The GIF format provides for multiple frames, hence the "animated GIF".

**JPG**

Joint Photographic Experts Group is the standard file format for compressed photographic files. When you create a JPG or convert an image from another format to JPG format, you are asked to specify the quality of the image you want. Since the highest quality results in the largest file, you can make a trade-off between image quality and file size. However, the JPG format is *lossy*, meaning that every time you save a file, some information is lost through compression. JPEG loses its E when seen as a file extension - lossy compression!

**PNG**

Portable Network Graphics. Since the PNG compression is fully lossless, and since it supports up to 48-bit colour or 16-bit grayscale, restoring and resaving an image will not degrade its quality, unlike standard JPG.

**PSD**

The Photoshop proprietary format, containing information about layers and other details specific to that software. This format is high quality, and useful if you want to be able to change part of a complex graphic, but the file size is large and not supported on the web, or in many imaging programs beside Photoshop. There are some free software tools which can understand some elements of the PSD files.

**RAW**

A completely uncompressed capture produced by a digital camera with all filters and compression tools disengaged. For professional photographers, this file format maintains the highest fidelity to the photographic subject, and retains a high image
quality after manipulation in Photoshop or other software. Drawbacks of this format include substantial file sizes and a lack of compatibility across platforms and media.

**TIFF**

The Tag Interchange File Format is the most widely supported lossless image format. This is the format to use for any high quality master image, and is especially relevant for archiving. TIFF files can be quite large.

**Vector Images**

All the above bitmap image formats store information about every single pixel within an image. This is excellent, in particular, for photorealistic images where resolution and high-definition are essential. However, many animation and publishing software tools use vector-based formats. In a vector image, information is stored which will allow the user’s system to actually redraw the image. To store an image of a square, for instance, a vector system would write a "recipe" for recreating the image, making note of the location, width, height, and colour of the square, instead of storing the blue data individually for each pixel within the square. Software tools such as Flash and Illustrator create vector images.

**Imaging for the Web**

Putting digital images on the web is often the main goal of a project. However, wherever possible, you should always treat web images as the final product of your project, not as the initial starting point (see Rule 1 above).

In order to be easily viewed on the web, most images are served in JPG or GIF formats, which are lossy and compressed. Your goal is to balance image quality with image size. Quality depends on your audience’s systems, and the purpose of the image. A logo should be very small, while a manuscript page must be quite large.

When you are preparing images for the web, disregard all measurements that come from print. Don’t look at the scale in inches or even in dpi. Use pixels instead, since
this is how an Internet browser measures an image. Scale your image to a maximum of 1024x768 pixels, the size of an average computer screen.

For web imaging, here are three software tools to consider: Photoshop, Fireworks, and Graphic Converter. Photoshop is the industry standard; it is excellent for fine detail manipulation and image creation. Fireworks is much more specialized, and concentrates solely on preparing web graphics. Graphic Converter is a Mac-only and less expensive program, and is very useful for preparing large folders of images using batch conversion.

If you have copyright concerns about your images, you may wish to use Adobe Acrobat’s PDF format, which can make it less easy for a user to save or modify an image. In addition, the PDF format allows a user to zoom in on detail, and offers a better web-to-print experience than other compressed formats.

**Imaging for Print**

In order to print a digital image, you need to ensure that you have a high-resolution image. There are two factors to consider: the physical print size and the resolution. A printer has a maximum dpi, meaning how many dots per inch it can print. If your project is using a single printer, it is worth finding out its maximum dpi and targeting that quality. However, if you are preparing images for general printing, the standards below offer a good compromise:

- Enlarge your image to a suitable print size. For snapshots, a 4x6 inch image is sufficient. Some projects may require a full-page size of 8x10 inches.

- Set your resolution to 300 dpi. This will give a good quality print on almost any printer. Some older printers have a lower maximum dpi, and will not be able to take advantage of your higher resolution.

- Choose a format for your image. This choice will depend largely upon your audience. If your project is a CD, a TIFF or PNG image is perfect.
This will allow your audience to print multiple images on a page, and arrange or edit the images. If you are serving a printable image on the web, a PDF or high-quality JPG offers good compression and quality. Keep in mind that many cameras produce JPG images, so upgrading to a lossless format for printing won't add any quality.

**Imaging for Archival Purposes**

To ensure the longevity of your archival project, the following needs must be addressed:

**Metadata**

For an image to remain useful, it must have information attached to it, such as the place, the date, the subject, and the original medium. Ideally, this information should be searchable and easily accessible. See the online links for more information.

**Storage Medium and File Formats**

Every year, new media are developed and new computer hardware is sold. Eventually, older media can be hard to access. The best way of ensuring that your images remain accessible is to use an open standard format instead of a proprietary format. In addition, a review process should be in place to ensure that all images are migrated to new media and/or formats when needed. Save image files in an uncompressed format like TIFF if possible. In addition, and in general for digitisation projects, don’t be too quick to throw away technology perceived to be obsolete. For example, there are many audio recordings found only on vinyl, and turntables are required to appreciate these recordings.
Quality

Archival images must be of a very high quality; the purpose is to ensure that these images are available and can be easily adapted to many purposes without having to return each time to the physical format (microfilm, slides, paper, etc.). The highest quality can be obtained by scanning at the highest possible resolution.
Resources for Digital Imaging

**General Info**

**Image Quality Calculator**

[http://images.library.uiuc.edu/calculator/](http://images.library.uiuc.edu/calculator/)

Use the Image Quality Calculator to help determine the best resolution for the document you are scanning. While we recommend scanning all documents at a resolution of 300 dpi, you may make a different decision based on the needs and constraints of your particular project.

**Moving Theory into Practice: Digital Imaging Tutorial**

[http://www.library.cornell.edu/preservation/tutorial/contents.html](http://www.library.cornell.edu/preservation/tutorial/contents.html)

**Capture Your Collections: Planning and Implementing Digitization Projects**

[http://www.chin.gc.ca/English/Digital_Content/Capture_Collections/](http://www.chin.gc.ca/English/Digital_Content/Capture_Collections/)

These tutorials will guide you through the process of digitisation; the site includes information about best practices in digitisation project management.

**Scantips.com**

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

Developed by Wayne Fulton as an on-line support tool for his independently-published handbook on scanning, this site provides insightful beginner-level information on scanning and the basics of digital imaging. Check out, in particular, his meditation on image resolution.

**A History of Colour Models**

Printing Resolution Calculator
http://www.scantips.com/calc.html

An on-line tool developed by Wayne Fulton that helps you to decide on scanning resolution, bridging the divide between digital and analog display, and taking into account issues such as cropping and margins in its calculations.

Metadata Resources

Dig35
http://www.i3a.org/technologies/digitalimaging/metadata/

This site focuses on defining metadata standards.

DCMI
http://www.dublincore.org/

Dublin Core Metadata Initiative. An open forum to develop the Dublin core metadata standards, which are used by the Canadian Government:
http://dcpapers.dublincore.org/ojs/pubs/article/viewArticle/1046

MDC
http://www.loc.gov/standards/mdc/elements/

Library of Congress Metadata resources.
Imaging Tutorial 1: Flatbed Scanning

1. Scanners often come with custom software to facilitate the scanning of images into the computer. Sometimes these software tools are separate applications in the computer, but sometimes they exist as plug-ins in standard image editing tools such as Photoshop.

2. Regardless, the software which communicates with the scanner will likely provide you with control over resolution, scanning area, colour or brightness settings, and the type of media you're scanning (photos, drawings, slides). Often, you'll need to access the software's advanced options to control many of these settings. The scanning software might look like this:

![Advanced Properties Window](image.png)

3. Perhaps the most important variable to control is the resolution, which could be set at 400 dpi, as shown above, for high-quality work. The software shown above provides control over brightness and contrast, but it is wisest to capture the image as is, and leave adjustments for later.
4. The last step before you scan your image is often a "preview" mode. This process may take a few minutes if the scanner needs to warm up, but when it is finished you will be able to see a preview of the content currently on the scanner, and the area to be scanned enclosed in the dashed lines. You can resize the scanning area by clicking on the red boxes in each corner. Some tools will provide the capability to select numerous areas of the same content, and save those areas as separate images.

5. Depending on the settings, the scan may take quite some time. Scanning at a very high resolution forces the scanner to look very carefully at every part of the content. When the scanner is finished, the image will appear in the editing software, such as Photoshop, or it may be saved to disc in which case you'll open it manually. The first thing you should do before editing is to save a high-quality copy of the image, using a format such as TIFF, and a meaningful name.
6. You now have a master copy of your image that meets archival standards. This copy should always be left alone and unmodified. Whenever you want to make a new image from the master, the first step is to save it in a new location or with a new file name to avoid over-writing the master image file.


**Imaging Tutorial 2: Digital Photography**

Sometimes the nature of the material to be digitised is not conducive to flatbed scanning. Consider oversize materials like maps, which if sent out to be digitised professionally can be costly. You may also have old manuscripts or fragile archival materials which require gentler treatment. In such cases, it may be useful to try digital photography techniques, which can be effective in the right environment. The right combination of camera, settings, and lighting can yield surprisingly good results. Some things to remember when using this method:

- ✓ Try to control lighting conditions as much as possible.

- ✓ If your document is fragile, use archival techniques in handling and storage.

- ✓ Be patient. Digitising a couple hundred pages of a book in this manner can take hours when done properly.

- ✓ Experiment with different combinations of cameras, lighting, resolution, and file formats until you find one that works.

- ✓ The settings to take a good image of one document may not be perfect for a different document. Make sure you do a trial run for each new document you want to digitise.

- ✓ Depending on the material you are photographing, you may be able to see text or images from the other side of the page. Place a piece of black poster board or construction paper underneath your page to minimize how much of the “phantom” text or image appears in your photo. If your page curls or is otherwise difficult to lay flat, lay a sheet of clear plastic or glass overtop.
The following steps will guide you through the process of photographing content:

1. Mount or position the camera at an appropriate distance from the document. A stable camera will produce far better images in low lighting situations. Be sure you can see everything you need to in the camera view screen.

2. Set the camera to an appropriate resolution. Many cameras have the ability to take pictures with different quality settings. Be sure the camera is will be using a high enough amount of pixels to represent your image.

3. If the material you are photographing is fragile or sensitive to light, turn the flash off and rely on available light. Bring in lamps as required.

4. If necessary, manipulate the zoom lens until you have targeted the document properly. It’s often best to move the camera instead of zooming, particularly if you’re working with a handheld (not fixed or mounted) camera.

5. Take the picture.

Once you have an image on the camera, you will need to transfer it to your computer. The simplest way to do this is connecting your camera with a USB cable.

To transfer images from the digital camera via a USB cable:

6. Plug the USB cable into the computer and also into the appropriate port on the camera.

7. Turn the camera on. You may need to switch the camera to viewing mode as opposed to shooting mode.

8. Software on the computer may open automatically and detect items to be imported. If not, start the appropriate software. Some cameras come with software which needs to be installed on the computer before you connect the camera. Many software tools, such as Apple’s iPhoto, understand how to speak to numerous camera models.
9. When importing pictures, give your "roll" or folder of imported images a name and a description. This will help you to keep track of large amounts of content. Most photo software will organize your imported images by date.

10. Decide whether you want to delete the originals from the camera. If you have taken a lot of photos, you may want to erase them to free up memory for your next batch of photography.

Many cameras make use of removable media for storage, such as an SD card; some computers have ports for these or other specific media. Some cameras and computers allow data to be transferred wirelessly, and/or uploaded directly to online hosting sites such as YouTube or Flickr, simplifying this step of the digitisation process.

You can perform some basic photo editing in software such as iPhoto, or the software which comes with your camera, (cropping, red-eye reduction,) but it is sometimes necessary to export your photos and edit them in a fully-featured program like Photoshop.

When exporting photos:

11. Select the photo (or photos) you wish to export from the capture software.

12. Remember to select a high-quality format such as TIFF, give your file a meaningful name, and leave your image full-size. Many capture tools provide export settings for the web, email, print, or other targets. See if you can find the "expert" or "manual" settings to specify the quality you need.
Imaging Tutorial 3: OCR Scanning

Optical Character Recognition is a process that can be applied to digital images of text documents in order to extract the text from the image. The process operates on the principle that a high-contrast image of black text on a white background, with a standard font, can be analyzed to discern characters, building text files for the document. The first step is to acquire appropriate images of the document that you want digitised via scanning or photography as described in previous tutorials.

Once the acquisition is complete, you will have a number of images open in Photoshop (or a similar editor). You will need to save these individually as TIFFs to an appropriate folder with appropriate file names. You now have the images you need to begin OCR.

Now that you have a clean batch of images prepared, you can begin the OCR process. A detailed and up-to-date description of this process can be found here:

http://community.nuance.com/wikis/omnipage/default.aspx
Imaging Tutorial 4: Using Photoshop

Cropping and Rotating

Images acquired through scanning or photography will often require cropping and rotation adjustments.

1. Cropping capabilities are most easily accessed via the Crop tool.
2. Select the portion of the image you would like to keep by clicking and dragging your mouse across the image.
3. The cropping box which results can be rotated and resized. Press return on the keyboard once the changes are correct. You can always undo and retry the operation if it doesn’t look quite right.

Adjusting Your Image – Colour and Contrast

Photoshop features many ways to adjust the light balance and colour contrast of an image, procedures used most often to increase the clarity of an image. You will find many options and features for adjusting your image under the Image->Adjustments menu.
The *Levels* tool is a versatile way to adjust the colour and contrast of your image, correcting exposure problems. Its advantages include visually displaying a histogram of the levels of colour and light in your image data, (allowing you to make informed decisions beyond the naked eye,) and immediately *previewing* the effects of your choices.

**Using Levels to Adjust Brightness and Contrast**

Photoshop creates a histogram for the image, charting its tonal levels.

The example above features an underexposed image, which means it is darker than one might like. In particular, it is difficult to discern details such as facial expressions. In order to correct for these deficiencies, we need to rebalance the colour distribution.
By moving the centre slider left, (establishing a new mid-range,) and reducing the total colour range, (by moving the right slider,) we are able to lighten the image considerably. This gives us much more facial detail, for example. To compensate for the foggy overall composition, we can turn now to the *Brightness/Contrast* tool.

In this case, increasing the contrast and decreasing the brightness gives a sharp, crisp image, with much more detail than the original.
Before:

![Before Image](image1)

After:

![After Image](image2)

**Using the Levels Tool to Improve Colour Balance**

The *Levels* tool can also help with colour balance. Click on the pull-down menu above the histogram. This allows you to adjust - with the three adjustment sliders - the range and centre of Red, Green, and Blue, respectively. Make sure you adjust the Brightness/Contrast in combination with the levels tool, or you will likely wash out all the vibrant colours in your image. The *Hue/Saturation* feature can be used to achieve similar results. The best way to understand these tools is to explore and experiment a little!
Imaging Tutorial 5: Using Layers in Photoshop

Using layers in Photoshop is like placing an overhead transparency over an image. You can write on the transparency, or print an image on the transparency, place it over your original image and it will look like one image.

1. Start by browsing the web for a photo of uVic or Victoria. Try to find a high resolution image - Google will help you with this. Open the image in Photoshop; this will be your background image. Next we’ll layer some additional elements on top of the background.

2. In the Layers toolbar, click the new icon. You may want to give the layer a useful name. Be sure this new layer is selected. Choose the marquee tool and draw a rectangle over the image where you might like to place some text.

3. Click on the background colour for the document, and use the magnifying glass tool from the Colors box to select a colour from the background image.
4. Use the *Paint Bucket* tool to fill the rectangle with the colour you’ve chosen.

5. To see how layers work, click on the eye icon next to the layer you are working with. Poof! Your rectangle is gone. Click on the eye to see your layer again. Try adjusting the Opacity setting for your layer to make the box slightly see-through.

6. Select the *Text* tool, and draw a box inside the rectangle you just created. Add some text inside the image, adjusting font and size to be sure the text fits inside the rectangle. Note that using the *Text* tool created an additional layer in your document.

7. Visit [dhsi.org](http://dhsi.org), and grab the logo from the top left corner of the page. Drag or save the logo from the browser window right to the Desktop. You may need to right-click on the image to do this. Open this image in Photoshop. Select the image using *Edit->Select All*, and copy it. Go back to the postcard image, and paste the logo on top. Note that this creates a new layer as well.
8. Note that, in the image above, the logo has been resized, and edited slightly to fit better with the rest of the image. In addition, the Blend Mode of the logo has been set to Multiply. Normally, layers have a Blend Mode of Normal, as seen in the Layers panel, which means layers on top completely cover layers below. This pop-up menu has a whole list of possible Blend Modes, which help layers merge their pixels with the pixels belonging to the layer below. The Multiply setting has the effect of making the resulting pixels always darker, and making the content on top slightly transparent.

9. Once you’re satisfied with your modified image, choose the Save As command to produce a new image for distribution. You may, of course, choose different settings and formats depending on the destination of your media.
SECTION 3: DIGITISING AUDIO

Prepared by Robin Davies, Allison Benner, and Brian Millward
Best Practices for Digital Audio

Whether you are starting with a vinyl record, a cassette tape, or a fresh recording your own voice, the following guidelines will help ensure your digital audio is suitable for your intended purpose. We will use software called Audacity to demonstrate. Audacity, though free, performs similar operations to industry standard software such as Avid’s Pro Tools, or Apple’s Logic.

There are four main factors you should to keep in mind when digitising audio files/media:

- The most common audio file formats
- How to connect analog audio sources to the computer
- Audio recording standards
- Optimizing your computer for recording digital audio

Basic Rules for Digitising Audio Files

The first step in any digital audio project is to determine how the audio will be used. For example, a digital audio recording for streaming on the web has very different requirements than a recording for a CD or DVD. We will begin by reviewing the basic rules for digitising audio files.

Rule 1: Always save a master audio file in an uncompressed format. We encountered this quality issue when dealing with images as well. Likewise, when you compress an audio file, you affect its quality in the process of compression. As a result, when you reformat an audio file from one compressed format to another, there will be greater quality differences than if you convert an uncompressed audio file into a compressed format.

Rule 2: Start with the highest quality audio file/media possible and then make the proper adjustments necessary for your project. Start with an original vinyl or reel-to-reel recording, as opposed to a poorly created MP3 file.
Rule 3: Manage your audio files. It is important to use a file naming convention to ensure that files are correctly labeled, and to avoid accidentally erasing or copying over a file when editing your audio files. Also, high quality audio files will take up large amounts of space. Be prepared for this.

Common Audio File Types

Many audio projects begin with an analog audio source which needs to be converted into a digital audio file. You will have to select a file type that is best suited to your needs. Each file format has a slightly different purpose. WAV and AIF file types are uncompressed digital audio; they are best suited for creating audio CDs, Blu-Ray soundtracks, or an archival master file. Compressed audio formats - MP3, RM, WMA, OGG - are best suited for computer and web-based presentations. For a brief definition of a full range of audio-related terms, visit the RANE Pro Audio website at: [http://www.rane.com/par-num.html](http://www.rane.com/par-num.html).

Macintosh AIFF (.aif)

The Audio Interchange File Format is the standard for high quality audio files on the Macintosh computer. It supports audio data of various bit depths, and in mono or stereo. If you want to transfer an audio file from a Mac to a PC, the AIF file format works because the format can be read by PCs using standard audio software like Audacity.

Microsoft Wave (.wav)

The WAV format is a Microsoft Windows-based audio file format that stores sounds as waveforms. The wav format also supports audio data of various bit depths, and in mono or stereo. If you are working with Windows-based PCs, you will probably use this format most of the time, because it is the most widely supported audio file format. If you want to transfer an audio file from PC to a Mac, the WAV file format works because Macs with standard audio software like iTunes can read it.
MPEG Audio (.mp3, .mpg, .mpeg)

MP3 is the audio file format which was the catalyst for the arrival of iPods and the fall of the traditional music industry. Officially known as MPEG-1 Audio Layer 3, MP3 is a form of lossy compression. MP3 music files can be one tenth to one twelfth the file size of the CD format WAV or AIFF file, while sounding exactly the same to most ears. This compression is achieved by filtering out elements of the audio not necessarily detected by the human ear using a technique called perceptual encoding. Some hardware CD and video players understand the MP3 format, and other file formats more commonly found only on computer systems, but the MP3 is most commonly found online and in portable music players. Audacity can export MP3 files using the LAME MP3 encoder, or you can use software such as iTunes to make MP3s for you.

RealMedia (.rm)

The RealMedia format has traditionally been one of the more popular formats used on the Internet for streaming media. This compression format allows you to stream audio and video files over the Internet, and listen to or watch them as they are downloading to your computer. The RealMedia format is a proprietary compression format that belongs to RealNetworks. For more information please visit www.real.com. In order for people to be able to play .rm audio files, they will require a RealPlayer media player.

Windows Media File (.wma)

The Windows Media File format is similar to RealMedia, in that it is a proprietary compression format used for streaming audio files over the Internet. Like RealMedia, .wma files require a special media player. For more information about the wma format and to download a free media player, go to http://windows.microsoft.com/en-US/windows/products/windows-media
**Ogg Vorbis (.ogg)**

This compressed audio format was designed as a free alternative to the MP3. Audacity can import and export this format. Ogg Vorbis files tend to take up a little less disk space than MP3s for similar compression quality, and Ogg Vorbis is free from patents and licensing restrictions. However, Ogg Vorbis files are not as commonly used as MP3 files.

**FLAC (.flac)**

FLAC stands for Free Lossless Audio Codec. This is a relatively new lossless compressed audio format that is a free alternative to MP3. Audacity can import and export this format, and many players support this file format. For more information on FLAC, and to download a free Windows installer, go to [http://flac.sourceforge.net/](http://flac.sourceforge.net/)

*Transferring Analog Audio Sources to the Computer*

Even though there are many different types of analog sources (tape players, turntables, amplified voice, and radio), the way they are connected to your computer is straightforward. With both PCs and Macs, the line from the amplified sound source that connects to your computer must be plugged into a port on your computer. You may find a number of audio ports on the computer:

- **Microphone** – this is where you would connect your microphone. You might also use a USB microphone. This port is specially designed to boost the quieter signals which are usually produced by microphones.

- **Input (Line-in)** – this where you would connect your stereo, mixer, or the headphone output of another computer or portable audio device.

- **Output** – this is where you would connect your speakers or a line to a stereo system or other audio recording device.
Audio Standards for Recording Audio Files

To achieve the highest quality audio files, you need to record the audio at the highest sampling and bit rate possible. The sampling rate defines how often the computer records the sound, and is measured in terms of samples per second ($Hz$ or $kHz$). The larger the sampling rate, the higher the frequency of sound that can be captured, and the better the sound quality. However, the depth of each sample also affects the quality of sound. Just like the sampling rate, the higher the bit depth, the better the quality of sound you can record. The basic guidelines to ensure CD quality audio for music or speech are to sample at 16-bit, 44.1kHz stereo.

Archival Standards for Digitising Audio Files

The questions to take into consideration in the creation and preservation of digital audio for archival purposes are similar to those for the digitisation of text and images: quality, stability, longevity, and accessibility. It is now possible to sample audio at rates of up to 192kHz, but are the storage needs created by the use of such a high sampling rate worth the information stored? You may choose to make a large collection of audio files accessible to the public in a compressed format such as MP3 (though, of course, you have an uncompressed master file!). Perhaps you have an auditory preference for a new, lossless file format, but is that file format widely used and supported? Will others be able to appreciate the file?
Standards for the creation and preservation of digital audio remain open to debate. However, the most widely used uncompressed file format used for archival master files is the WAV file, sampled at a minimum of 44.1kHz, 16-bit (though 48 kHz or higher, and 24-bit or higher is recommended). For accessibility purposes, the most common compressed file format is MP3.

**Optimizing Your Computer for Recording Digital Audio**

Most audio recording and editing software is disk-based - you record and edit directly onto the computer’s hard-drive. This allows you to edit large amounts of data and to retain extensive *edit history*, the ability to undo and redo, also known as non-destructive editing. However, this also means that your hard drive needs to have sufficient temporary storage and free space to store large amounts of data. In addition, a higher amount of RAM in the computer will allow for faster processing and editing, because RAM is much faster than accessing the HD. Therefore, to ensure better performance and to avoid problems associated with *incontiguous* files, you should run your operating system’s defragmentation program before starting to record audio, and close unnecessary applications.
Resources for Digital Audio

**Software Packages**

Audacity is free, open source software for sound recording and editing, available for many operating systems.


Cakewalk audio software offers a wide range of software packages for professional and amateur musicians.

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

Dart software produces digital audio recording and restoration software for all levels of users.

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

Waves creates digital audio recording and restoration tools.

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

Adobe Audition (formerly Cool Edit) offers advanced audio mixing, editing, and effects processing capabilities.


No list of digital audio resources would be complete without a mention of Digidesign's Pro Tools software (which is now managed by Avid), the industry standard for recording and production.

**Audio Standards Websites**

http://www.bnoack.com/

This website provides comprehensive technical information on both audio and video standards, as well as a vast inventory of available audio- and video-related products to support those standards.

www.mpeg.org

MPEG.ORG claims their website is the most complete and comprehensive index of MPEG resources on the Internet. However, even though the website is an excellent resource the MPEG.ORG mostly focuses on the MPEG-1 and MPEG-2 standards.

www.pro-mpeg.org

The Pro-MPEG Forum is an association of broadcasters, program makers, equipment manufacturers, and component suppliers with interests in realizing the interoperability of professional television equipment, according to the implementation requirements of broadcasters and other end-users.

http://www.vorbis.com/

This website is dedicated to the Ogg Vorbis audio compression format. Here you can find information about Ogg Vorbis and download the encoder. Ogg Vorbis is comparable to other formats used to store and play digital music, such as MP3, but it is completely free, open, and unpatented.

http://flac.sourceforge.net/

This website provides information on the FLAC (Free Lossless Audio Codec) audio compression format. You can also download a free Windows installer for FLAC on this website.
Audio Tutorial 1 – Getting Started

We will use Audacity for this tutorial. Please keep in mind that each of the functions covered here are applicable to other audio editing software. The goal of this tutorial is to demonstrate how to acquire audio files, how to perform basic editing, and finally how to burn audio files to CD. In order to help you understand the many different functions within the program, we have broken the tutorial into three main categories: Input, Processing, and Output. We will begin by looking at all the different ways to input audio media into the software.

Input

To illustrate the input mode, we begin by connecting a microphone to the computer. There are several types of microphones available. You will need to make sure that the microphone you want to use is designed for a computer.

Once you have verified which type of computer microphone you have, plug it directly into the computer’s microphone jack, or plug it into a USB port. You may also be able to use the computer’s internal microphone. Once you have connected the microphone, open your audio editing software and begin recording. For an example using Audacity, refer to the section on recording audio within this tutorial.

In order to digitise vinyl records and audio cassette tapes, you will first need to connect the audio device to your computer with an audio cable. Refer to the Best Practices for Digital Audio for greater detail on this point.

Once you have connected your analog audio device to your computer, you will need to open your audio editing program. From your audio editing program you can begin to record.

Recording Audio

Before recording audio in Audacity, it is important to save a new project and to set the preferences for any files you will record or import in your project.
Create a project

Under the File menu, select Save Project As... and choose a location and filename for your project. Remember that tools such as Audacity make non-destructive changes to your source material, by creating new audio files when you make modifications to existing ones.

Set your preferences

1. Open the Preferences dialog box (from the Edit or Audacity menu).

2. Under the Audio I/O tab, check that you have selected the correct input and output devices. If you are recording from an analog device, select your sound card. If you want to make a stereo recording, select the appropriate number of Channels.
3. Next, under the **Quality** tab, select the sampling rate and the bit depth for your project. Any files you record within the project will automatically have these settings.

![Audacity Preferences](image1.png)

4. Finally, check that the gain (input level) on the mixer of your soundcard is set to an appropriate level. Depending on your soundcard drivers, you can do this using the volume control settings on your computer, via a Control Panel or System Preferences.

![Volume Control](image2.png)

To adjust the gain within Audacity, use the onscreen sliders. The control on the left, marked by a speaker icon, controls your output/playback level, while the control on the right, marked by the microphone, controls your input/recording level.
Record

5. Audacity, like most other audio software, uses a multitrack tape deck metaphor for an interface. Click on the red **Record** button to begin recording.

6. Click on the blue **Pause** button if you need to pause the recording, particularly if you’re giving a vocal performance. Any blank sections in the recording still take up space on the HD, and will likely have to be removed later. Note that when the Pause button is pressed it’s impossible to use many of the other features of Audacity.

7. When you are finished recording, press the **Stop** button. When you stop recording, the start position will reset to the beginning of your audio track.

8. To review your recording, press the green **Play** button.

JUST FOR FUN:
Ever wanted to hear yourself singing in harmony with yourself? You can do this in Audacity.

Record yourself singing the melody line. Then, go to the **Quality** tab under **Preferences** in the **Edit** menu. Select **Play other tracks while recording new one**. Press **Record** and sing the harmony in sync with the melody. Press **Play** and see if you like what you hear. If you do, export the file as a wav; the two parts will automatically be mixed into a single sound file.

"Ripping" Audio from CDs

Many audio editing software packages can import directly from an audio CD. However, when using Audacity, you will need to use *iTunes* or other CD **ripping** software program to extract CD tracks into a format that Audacity can read, such as wav or aiff.
Opening a File

9. To open a file, you can drag and drop the audio file into the Audacity window, or you can go to the Project menu, select Import Audio, and choose the file you want to open. Audacity can import many file formats.

10. The audio file will display as a waveform in an audio track. If you have made a mono recording, there will only be one window. If you have recorded in stereo, the top window corresponds to the left channel, and the bottom window corresponds to the right channel.

Saving a Project

While editing all the files that are part of your project, you should frequently (as with any other computer task) save your work, in this case with the Save Project function under the File menu. Note that most of the audio data for an Audacity project is not stored in the project file itself, the .aup file, but in a separate folder with the same name as the project plus the _data suffix. When you open the project file, all the audio content associated with the project will automatically open as well. However, until you export sound files that you have recorded, they will not be readable by other players or software programs, so don’t lose track of the .aup and _data folder!
**Exporting a File**

To save a master copy of the file that is readable by other software programs, you need to *export* the sound file from Audacity.

Audacity lets you export a file in a variety of formats, including wav, aif, mp3, ogg, and flac. As stated in the best practices, you should first export your project as an uncompressed master by selecting the **Export as WAV** or **Export as AIFF** under the **File** menu. After you have created a master audio file, you can export the project in a compressed format using the **Export as MP3** or **Export as OGG** options under the **File** menu. When you export a file, Audacity will ask you to specify where you want the file to be stored.

**TIP:**
If you want to export only a part of your project in a format readable by other programs, select the desired portion, go to the **File** menu, and use the **Export Selection as WAV**, **Export Selection as AIFF**, **Export as MP3**, or **Export as OGG** options.
Process

This part of the tutorial will show you how to perform basic editing functions on your audio files. Although editing is usually viewed as a complex process, the following tutorial has been structured to look at each editing function separately. We will also review the benevolent **Undo** command. The most important thing to remember when editing is that most software packages follow the convention of “select then do”.

**Selecting and the Editing Tool**

The easiest way to select a portion of a file for editing is to use your mouse. Simply click and drag in the desired direction until the portion of the sound file you want to select is shaded. If you want to extend the shading beyond your original selection, press the Shift key and drag the selection to the right or left of its original location.

If you are making a selection within a stereo recording, both the left and right audio channels are selected by default. If you want to control selections in the left and right channels independently, go to the drop-down menu under the arrow immediately to the left of the audio track, and select **Split Stereo Track**.
Copying and Pasting

Just like in a word processor, the copy function allows you to copy a selection without modifying the original file. To demonstrate this, we will quickly review the step-by-step process.

1. Open a project containing an audio file.
2. Create a selection. (To listen to the selection, press the space bar.)
3. From the Edit menu, choose Copy or click on the Copy button.
4. Once you have copied the selection it will remain on the clipboard until you copy another selection. The next step is to paste it into a new location within the project.
5. Move your cursor to where you would like to insert the copied clip.
6. From the Edit menu, select Paste. The selected clip is now inserted at your cursor.
7. Press play to hear the changes.
Cutting

As with the copy function, cutting is similar to a word processing program. The Cutting function allows you to cut a selection out of the audio file by removing it and placing it on the clipboard. To demonstrate this, we will quickly review the step-by-step process.

8. Create a selection.

9. From the Edit menu, choose Cut.

Use the Paste command to place it at the end of the audio file.

Deleting

Deleting a selection permanently removes it from the project. To delete a selection, press the Delete Key on the keyboard.
Undo, Redo, and Versioning

You can undo any edit operation by choosing Undo in the Edit menu. You can redo any undone edit by choosing Redo in the Edit menu. Some audio editing software offers an editing History, similar to the history you may have seen in Photoshop. This allows you to see a list of the processes you have performed while editing, and jump back to a specific location in your work. Audacity will let you step back numerous times. However, it's often beneficial, particularly with a large project, to save separate versions of your work when you reach significant milestones. Choose the Save Project As... command, and make a copy of the file, or duplicate your file on the Desktop. Your computer will automatically put a new date on your work, so you can always go back and see what you were working on several days or weeks ago. Often, projects will move in directions that hindsight dictates were incorrect. This technique of versioning allows you to evaluate your progress.

Output

Burning CDs or DVDs

Many audio editing software programs allow you to burn an audio CD. Audacity does not include this feature, but software tools such as iTunes, Nero, or Toast make burning easy.

1. Insert a blank or unclosed CD or DVD into your computer.

2. With some burning software, you may be prompted to choose the type of disc you want to burn. If you only want to play your disc in a computer, (or some media systems, such as DVD players) you may choose to burn a data disc. However, if you want your CD to play in a conventional CD player, you need to burn an audio or music CD.

3. Next, add or delete the tracks/files you want to burn to your media, and preview how the tracks sound back to back. iTunes accomplishes this by treating the files destined for a CD like any other playlist.
4. Before you burn the disc, decide whether or not you will want to burn more tracks or sessions to this disc. It's very possible you'll only be burning a few tracks or files at a time. If you close the disc right away, it's like renting a large moving truck and only putting one chair inside. The rest of the space is wasted.

5. Don't forget to label your media once it's burned!
Audio Tutorial 2 - Advanced Audio Editing

The following Tutorial will cover advanced audio processing tools. Some features, such as pasting, mixing, and noise reduction, can be performed within Audacity. If you require more advanced, automated editing functions, you may wish to investigate a more specialized audio editing package, such as Sound Forge or Audition.

**Pasting**

1. If you have already saved an Audacity project file (.aup), open it. Any sound files you previously saved in the project will automatically open, and you may proceed to step 3 below. If you are starting afresh, create an empty project.

2. Assuming you are starting from an empty project, import two sound files into the project, either by dragging them into the Audacity window, or by using the *Import Audio* function under the *Project* menu.

3. Select a portion of one of the audio files.

4. Use the **Copy** function under the *Edit* menu, or the Copy icon on the toolbar at the top right-hand portion of the screen.
5. Place your cursor at the location in the other audio file where you want the selected portion to appear. Select **Paste** from the **Edit** menu, or press the Paste icon on the toolbar.

6. The selected portion will now be inserted at the location you selected. If you are satisfied with the result of your paste, choose **Save Project** under the **File** menu. If you are not, press **Undo** under the **Edit** menu. Keep in mind that within an Audacity project file, you can still undo a change even after you have saved it.
Mixing

If you are working with several audio files within an Audacity project, one way to approach your audio editing is via the Cut, Copy, and Paste functions described above. Alternatively, you can use the mixing features within Audacity to work with these files. If you want any of the sound files to overlap, for example by combining two parts of a harmony, or combining a vocal line with background sound effects that you have recorded separately, mixing is essential.

Suppose that you have two sound files. The first is a recording of one person talking, and the second is a recording of a different person talking. You want to combine these files such that the first person’s utterance slightly overlaps the second. Mixing can help you do this.

1. Open or save an Audacity project and record or import the sound files you want to mix.

2. Activate the **Time Shift** tool by pressing on the button in the upper-left-hand corner of the Audacity window. This tool will allow you to move each audio track around within its own space, allowing you to adjust the position of each sound file relative to any others in your project. In the example below, the beginning of the second sound file has been shifted to the right, and it slightly overlaps with the first sound file.

**TIP:**
If you have more than one sound file open within an Audacity project, and you press **Play**, both files will play simultaneously. To listen to one of the files without the interference of the other, press **Mute** on the window immediately to the left of the audio track. If you want to hear the files playing together again, simply press **Mute** again on the sound file you previously silenced.
3. When you are pleased with the way the two sound files sound relative to each other, save the project and export the project as a wav. The exported wav file will be a mix of the two sound files.

**TIP:**
If you are working with several audio files in a project, and you want to see them all at once on the screen, select Fit Vertically under the View menu.

**Silencing Unwanted Sounds**

If your sound file contains an unwanted noise such as loud breathing, the sound of a table leg squeaking, or something falling on the floor, and if this sound occurs in isolation, you can either cut it out completely or replace this portion of the sound file with silence.

To use this feature, use your mouse to select the unwanted sound. Next, under the Generate menu, select Silence. A small window will appear specifying the length of time of your selection. Press Generate Silence, and your selection will be replaced with silence.
If you have performed this operation, you may find that the result is an unnatural drop in ambient noise. To minimize this effect, select a portion of the file before the silence, go to the **Effects** menu, and select the **Fade out** function. Then select a portion of the file after the silence, go to the **Effects** menu, and select the **Fade in** function. In your selections of the portions to fade in and fade out, keep in mind the general principle that you should fade in quickly, and fade out slowly. This is the way most real-world sounds occur.

**Click Removal**

If your recording contains unwanted clicks and pops, you can use the **Click Removal** feature in Audacity to remove them. To take advantage of this feature, follow the steps below.

1. Open a file that you want to edit for clicks and pops. Using your mouse, select the portion of the file that you want "cleaned."

2. Under the **Effects** menu, select the **Click Removal** function. A window will appear that will allow you to preview what your sound file will sound like following click removal, based on the default settings on the sliding threshold scales. Press **Preview** to see if you like the result. If you detect no noticeable difference, or if the result sounds distorted, adjust the threshold accordingly.

   When you are satisfied with the result, press **Remove clicks**.
Noise Reduction

Audacity has a **Noise reduction** feature within the **Effects** menu. While this feature is not always effective in removing noise without introducing artefacts or distortion, it can work well with some types of noise, such as fans, tape noise, or hums.

1. Open a sound file you want to rid of background noise. Using your mouse, select a portion of the file that contains an example of the background noise *on its own*. Often you can find a good example at the beginning of the sound file.

2. Under the **Effects** menu, select the **Noise removal** function. A window will appear outlining a two-step process. In the first step, Audacity will use the noise selection to create a noise *profile*. To activate this feature, press **Get Noise Profile**.

3. Press **Preview** to hear the results of the noise removal based on the sample you have provided and the noise removal level indicated by the sliding scale in the middle of the Noise Removal window. If you find the result distorted, try sliding the gauge towards less noise removal. If you are satisfied with the anticipated result, press **Remove Noise**.

These features which remove clicks and noise are handy. However, the best way to deal with noise is to have a clean recording in the first place. If you have control over the sound quality and recording environment, do your best to keep things quiet and tidy from the outset.
SECTION 4: DIGITISING VIDEO

Prepared by Robin Davies, Allison Benner, Derek Finstad and Laurel Fulford
BEST PRACTICES FOR DIGITAL VIDEO

**Digital Video**

Digitising video involves three steps: capturing the video (transferring video from a video camera or another source to the computer), manipulating the video, and exporting the video to a file that can be viewed with a computer-based player.

Digital video is data displayed as pixels which constitute the individual frames of a video sequence - moving pictures. The quality of the video is determined by two main factors: frame rate and bits per pixel. A high frame rate produces a smooth moving image, while a low frame rate produces a choppy moving image. The number of bits per pixel determines the colour quality.

While uncompressed file formats and lossless compression methods are often used for digital audio, lossy compression is the norm in digital video, even for archival purposes, because the large amount of information contained in digital video files makes lossless compression impractical. The most widely used compression for video is MPEG, which is available in MPEG-1, MPEG-2, and MPEG-4 schemas.

**Shooting Video**

Here is a short list of things to keep in mind when shooting video. Some of the information provided is exclusively for shooting video for the web.

- Use a tripod, especially if you plan to use a zoom lens.

- Use close-ups if your video will live online. Close-ups are better for the often small format used by mobile devices, YouTube, and similar sites.

- Frame the shot; leave a little breathing, looking, and talking space around your subject.

- Don’t move the camera too quickly. Avoid the *Blair Witch Project* effect.

- Backgrounds should not be busy or similar to your subject.
Backgrounds should not be exceedingly light or dark; some cameras will compensate for these backgrounds by lightening or darkening your subject.

Make sure the timecode is disabled in the camera, unless you want the timestamp to appear in your video permanently. It is possible to erase this timestamp in your final product, but it is very tedious.

Leave extra time at the beginning and end of each shot; this time can be removed when you edit the video, but it is difficult to add.

Avoid jarring jump cuts. A jump cut results when there is a gap in the flow of action, often introduced during the editing process. To avoid this, make sure you have footage that can be inserted to create continuity.

Minimize changes and movement from frame to frame. This is an internet trick, since video compression often functions by storing only information that is changing. More complicated video will take longer to compress, and will likely be a larger file.

Bright whites, yellows, and blues (and other hot colors or glares) are not as attractive as darker tones in video for the Internet. Try to make sure your subject is wearing dark, solid colors, as opposed to crazy patterns.

Set your audio levels as close as you can to what you want in your end product. Aim for as hot a signal as you can get without clipping or distorting the sound on the tape. Do some tests in advance to ensure the audio setup is functioning correctly. It’s very challenging to compensate for missing or bad audio after the shoot!

Save camera tricks for the editing stage. Then you won’t have to include them if you change your mind.
When shooting a close-up of someone speaking, remember that excessive hand gestures can be distracting. Try to frame your subject’s head and shoulders. If the subject is standing, provide a chair for steadiness. If the subject is seated, encourage a straight back, thus simulating a standing posture. Otherwise, the subject may look slouched.

Let the camera capture the initial video at a high frame rate. Video on traditional television plays approximately 30 frames per second. Video on the internet sometimes plays as low as 8 to 15 frames per second. When you are shooting, it is preferable to let the camera film at the standard 30 frames per second; when you have edited your video, you can let the computer compress it to fewer frames per second. If you film at a low frame rate, there is no way to make it look better if you ever decide to use your footage somewhere other than the 'net.
Resources for Digital Video

http://www.mediaconverter.org/
This is an all-purpose, online media convertor for audio, video, images, text files, and more!

http://www.mediacollege.com/
This site provides information and tutorials on digital audio and video recording and processing.

Check here for information about the current video compression standards used by Blue-ray Discs, YouTube, and others.

http://websitehelpers.com/video/
Here’s a great discussion about web video formats.

Many video examples for iPhoto, iMovie, and Garage Band can be found here.
(Some) Codecs Described

DV-NTSC

Using DV, which some camcorders still use, results in very clear footage with large file sizes. NTSC is the standard for television in Canada and America. Your storage is limited to the number of tapes you can purchase.

Animation

This codec will compress your animation files but will not damage the visual effect. The codec works best on images with large areas of solid colour, as found in animations. File sizes are still quite large.

Flash Video

This .flv codec is widely used online by sites such as YouTube. There are many free players for .flv content, and software such as iMovie will export in this format. The .flv format belongs to Adobe (formerly Macromedia).

H.264

You may come across media encoded in H.261 and H.263 formats. H.264 is the latest in this line of codecs, and uses a process similar to MPEG4 to reduce the data size of videos. This codec is used for viewing videos over the Internet, as the file size and data rate are small. However, as the compression system is quite complex, a fast CPU is required to decode the data. As a result, a machine with a fast internet connection may be able to acquire the data quickly, but be unable to decode the data to display it.
SECTION 5: WEB DESIGN

Prepared by Michael Nixon, based on material prepared by Cara Leitch and Gerry Watson
Best Practices in Web Design

This section will provide an overview of the design and construction of simple web pages, introduce some more advanced techniques, and examine how you can incorporate different types on digital materials into your web site. We will begin with the construction of a simple web page, then move on to a discussion of website design and the use of templates. Next, we will tackle some more advanced design topics.

An increasing number of projects, both digital and non-digital, now incorporate web components. It is important to know not only how this technology works, but also the opportunities website design affords in terms of presentation and interactivity.

The World Wide Web and Digitisation

The World Wide Web refers to the collection of inter-related documents and resources that are made up of computers, servers, and other equipment that allow the communication of information across the globe. Hailed by some as “the Gutenberg press of our time,” the web presents a number of interesting opportunities and challenges for scholarly projects and digitisation.

While it can be tempting to put your information online “as is,” there are major benefits to starting your web project by planning out a proper website. You should begin by planning out how you will organize your information and how users will navigate your site. The next step is to design the site so that it is not only aesthetically pleasing but also usable, intuitive, and scalable. For projects in the digital humanities, where communication of information and scholarly collaboration are so important, a well designed and accessible website can prove to be an essential tool.

The Basics of HTML

HTML, or Hypertext Markup Language, is a coding language that is used to tell web browsers what you want to include in it, how you want it to be displayed, where to
find external resources like images and how to display them, and where to send people when they want to go elsewhere. As you will see, it is a little more complicated than this, but this is a good way to begin thinking of HTML. The most modern form of HTML is XHTML, which is XMLized HTML. This means it's a little more strict in terms of requiring closing tags and a few other formalities.

**The Building Blocks**

Every web page consists of a few basic elements: **content** in the form of text or multimedia objects (images, videos, etc.), information about this content that will be translated into **instructions** about how to present the content, and **information** about the standards and markup used in the page.

**Content** consists of all the material that needs to be presented on the screen, from the text of a link, to blocks of text, videos, or images. Some of this content is held internally in the website markup, such as text and links, and some of it is opened by the browser as it loads, such as images and videos.

The **instructions**, written in HTML (and often referred to as code, tags, or mark-up) surround the content and tell a web browser more about the content and how to display it.

**Information** about standards and markup is required to build standards-compliant websites.

**Getting Started**

The first step to building a webpage is using a good editor. Initially, web design was done by hand, writing code in a simple text editor. Some web designers write code by hand while others use a more sophisticated IDE (integrated design environment) that helps to fill in pieces of code.

For this course, we will focus on the basics by using a moderately sophisticated text editor.
To start a new document, begin with the proper starting template. You can get one of these from the Web Standards Project:

http://www.webstandards.org/learn/reference/templates/xhtml10t/

Web pages use a Document Type Definition (DTD) in order to specify the version of HTML or XHTML that they rely on for validation. XHTML 1.0 Transitional is one of the most commonly used.

Once you copy and paste the template from the website above, you can begin to make changes. Add the following code within the <body> tags:

```html
<p>Hello World!</p>
```

You can then save it on your computer as a starting point.

Congratulations! You’ve just made and saved a web page that we will use for the exercises to follow.

**Tag, you’re it**

A tag is mark-up that surrounds (or wraps) a content element, giving the browser specific instructions about how to display the content. Tags are the parts of code that make websites look much scarier than they actually are. All tags have this basic syntax:

```html
<tag attribute="value">content</tag>
```

The **tag** is the instruction that tells the browser to perform a certain action with the **content**.

The **attribute** modifies the **tag** in a certain way, which is specified by the **value**. All tags have a certain set of attributes in common, but many have attributes specific to themselves. Each attribute has certain values or ranges of values that can further specify the instruction. To illustrate this point, let’s modify our "hello world!" line. Add the highlighted code to the line so that it looks like this:

```html
<div class="special"><p>Hello World!</p></div>
```
To clarify the above example, **div** is the name of the tag and has the **class** attribute with a **value** of “special”.

**Angle brackets**

Before we move on, you may be wondering what the angle brackets are for. When the browser reads through your `index.html` file, it is looking for angle brackets to tell it which bits of text are instructions, where these instructions end, and what they apply to. The angle brackets let the browser distinguish between the tags and the content.

You may have noticed above that the content was wrapped by the tag in this format:

```
<tag attribute="value">content</tag>
```

The first instance of the tag, along with its attributes and values, is enclosed in angle brackets so that the browser can recognize that there is an instruction to be applied to the content, which begins after the first closing angle bracket. After the content, there is the closing tag enclosed with angle brackets again, but preceded by a forward slash. This is the second half of the tag and it tells the browser where to stop applying the instruction. Without the slash, the browser will not know where to end the tag and will also think the closing tag is really the start of a new instruction altogether.

A more simplified form of the tag structure would be:

```
<start instructions>content</end instructions>
```

Even experienced web page designers sometimes forget an angle bracket or forward slash, which can cause problems with how things look. If you have problems, check to make sure your syntax is correct. Every tag needs to be opened and closed in order for the browser to understand the instruction properly.
Nesting tags

Often, you will want to do multiple things with one section of content. Imagine that we want to take the phrase “Hello World!” and make the whole phrase emphasized, and render only the word “World” in bold. The code to do this could look like this:

```html
<p><em>Hello <strong>World</strong>!</em></p>
```

While this looks a bit daunting, it is easy enough to follow when you talk it out, step by step:

“Hello World!” is wrapped in one `<p>` tag to tell the browser that this is a paragraph; it will likely be given vertical space away from other text to demarcate it.

There is a `<em>` tag to make it all emphasized (usually italicized). Only the content “World” is wrapped in the `<strong>` tag to make it appear bold.

Try adding this to your code and see what it looks like in your brand new page. Don’t forget to save and refresh.

Notice that each tag is opened and closed in the same order. These are properly nested tags. Be careful always to close your tags in the proper order. Sometimes you can get away with improperly nested tags, but it can cause major display problems, is confusing to read, and most importantly, it isn’t good practice or valid HTML.

There is another lesson to learn here: keep things simple whenever possible. Use the minimal amount of mark-up within the document, and apply style information inside of Cascading Style Sheets.
Do you think you could change the font color to red? How about green? Or change the font size? How about making all the content italicized with the <em> tag? Try adding these features one at a time and notice how your changes are reflected in the web browser.

**TIP:**
Instead of <strong> and <em>, web designers used to use <b> and <i>. Actually, some still use <b> and <i> even though their use is discouraged. Don’t be one of those web designers. The following will explain why:

Physical Styles  [<big>, <blink>, <b>, <font>, <i>, <small>, <s, strike>, <sub>, <sup>, <tt>, <u>]

Physical styles explicitly describe what the final appearance of the contained text should look like. If the rendering device does not have the capability to produce the indicated Physical style (such as a browser for the visually impaired), this formatting may be lost.

Virtual Styles [<abbr>, <acronym>, <cite>, <code>, <dfn>, <em>, <kbd>, <q>, <samp>, <strong>, <var>]

Virtual styles purposefully do not include any final rendering hints in their definitions. These styles describe instead how the contained text is used in the context of the document. The main benefit of such elements is when a document is experienced using a method OTHER than a standard screen-based visual environment; the INTENT of the content can be preserved. (“More About Character Formatting in HTML”


There are also specialized tags that can replace some of these instructions. One such tag is a heading tag <h1></h1>, outlined in more detail below, but let’s use it to clean up our "Hello World!" line:

1. Delete the size attribute and value from the font tag and remove the <strong> and <p> tags. Don’t forget the closing tags.

2. Next add the <h1></h1> tags before the <font> tag. Remember the rules on nesting here. Where does the end </h1> have to be?
3. Save the file and then view your changes.

*So many tags, so little time...*

This is a good time to mention that there are many different tags that you can use to do a variety of different things, from organizing the internal structure of your web page, to linking to other sites, changing the format and style of your text, building tables and web forms, and so on. Frankly, if you are new to web design, it is overwhelming and frustrating to work with all these tags. If you want to look at more tags and see what they do, you can look at the HTML documentation at the W3 Schools website ([http://www.w3schools.com/](http://www.w3schools.com)). If you prefer printed documentation, O'Reilly prints a small HTML handbook would be quite useful in this regard.

*More Tags*

**headings** - `<h1>content</h1>` automatically makes the content a certain level of heading from 1 to 6. The smaller the number, the larger the heading will appear. For example:

```
<h1>This is bigger</h1>
<h3>Than this would be</h3>
```
ADVANCED TIP:
Can you guess what the problem with this way of thinking is? Consider the following text:

News of the World
Strange Happenings at UVic
Rabbits Suspected!

There are three headings. The first heading is larger than the next two. Headings two and three are the same size. If you use heading tags as physical styles, you would write the following code:

```html
<h1>News of the World</h1>
<h2>Strange Happenings at UVic</h2>
<h2>Rabbits Suspected!</h2>
```

However, remember what we said about the value of using virtual styles that describe how text is used in a document. “Rabbits Suspected!” is Heading 3 and should be tagged

```html
<h3>Rabbits Suspected!”</h3>
```

You can then use CSS to control the font size of each heading.

**paragraph** - `<p>content</p>` denotes that the content is a text paragraph and should be treated as such. A new set of `<p>` tags means a new paragraph.
lists - you can make them ordered (<ol>) or unordered (<ul>) and each list item is displayed as indented with numbers for ordered lists and bullets for unordered. For example, this code:

```html
<ul>
  <li>1st item</li>
  <li>2nd item</li>
  <li>etc…</li>
</ul>
```

Results in this list:

1st item
2nd item
etc…

These are just three examples for you to try out. We will explore more as we go along. In some ways, the W3 site (noted above) is more helpful, as it has the tags broken into categories based on what they do.

**The framework of a web page**

Now that you know basic HTML, we can take a closer look at the framework of a web page. There are two main parts to a web page that need to be understood, the `<head>` and the `<body>`. The **head** contains code that gives the browser and anyone looking at your source file information about the web page and its content. The **body** contains all the content, structure, and formatting for the website, which is what we’ve been changing so far.

The head is not displayed when a browser loads the page and holds information such as the `<meta...>` tag, which should be on your screen and describes what data type the document is and how it is encoded. You should also see the `<title>` tag.
which contains the text that shows on the title bar of your web page as well as in bookmarks.

To change the title of your page, simply change the text wrapped by the `<title>` tag. Refresh your web browser and notice how the title bar of the window now contains the text you just typed.

The body section of your website is simply the place to put all content that you want to be displayed on the screen when your web page is loaded. This section can be as long or short as the content you have to fill it, but remember that the viewer will only scroll down or click onwards if they are interested by what appears on the screen.

**Validating Your Code**

The best way to ensure a stable foundation for your web page is to design it to certain standards. This means choosing a set of rules that you can then check your site against for compliance. In HTML, the rulebooks come in the form of Document Type Definitions, or DTDs. A DTD is the digital rulebook that specifies which tags can be used where, what attributes they can have, and which features are supported and how. By choosing a DTD, you are declaring that your site plays by a certain set of rules, making it more consistent and reliable. The DTD that you choose will dictate certain aspects of your site, but also gives you a way to validate your code. Notice the very first line of index.html that starts with `<!DOCTYPE`. . . >This line tells your browser which standard the web page should comply with, and where to find the appropriate DTD. For now, let’s validate index.html:

- First, make sure the file has been saved.
- In your web browser, load [http://validator.w3.org/#validate_by_upload](http://validator.w3.org/#validate_by_upload)
- Upload your file and click “Check.”
- You can click on the errors to pinpoint where they are as well as learn what the specific problem was.
• Now, remove a forward slash from one of the tags wrapping "Hello World!" and validate again. What happens?

• Fix the error you just created as well as any others and save index.html

• You now have valid HTML.

**Coding Exercise**

Before we move on, let's take a look at some less than perfect code and see if you can find the errors without the help of a validator. Below, there are five things to be fixed and one irregular tag—good luck!

```html
<h1>Jabberwocky</h2>
<p><em>Lewis Carroll 1871</p>
<br />
<p>'Twas brillig, and the slithy toves<p>
<p>Did gyre and gimble in the wabe:</p>
<p>All mimsy were the <span style=funny>borogoves</span>,</p>
<p>And the <b>mome</b> raths outgrabe.</p>
```
Building Websites

All web sites start with a single page. What do you do, however, if you have two major topics and want to give them equal emphasis? You make a new page and link to it with a special tag. This is the easiest way to divide your content and navigate to and from important sections. You can repeat this step as many times as necessary. Before you do, however, it is advisable to plan out (on paper, if you prefer) how all these pages will work together.

Just like any digitisation project, your digitisation web site requires a plan before you go any further. You need to think of a design and style that suits your topic and audience. You need a layout that has the appropriate balance of visual, spatial, and functional components. You also need to think about growth. As you digitise more material, can your site grow to accommodate it? Will you be able to easily incorporate new functions and features? How easy will it be to maintain this site? (This is especially important if someone else will be responsible for site maintenance.) Will you be able to easily update both the content and style of the site? Before we dive into the basics of HTML, we will consider all these questions.

Web design: theory and practice

The first step in developing a new website is thinking about what it is you want to present, how you want it to look, and how you want your users to be able to interact with the information. You can think of design at the micro level of links, words, and typefaces or on the larger scale of overall look, white space, and flow of the page or site. The best-looking websites have often paid close attention to creating a balance between the two levels.

Paper Prototyping

The first step to good web site design is to start with a paper prototype. Think of one sheet of paper as the browser window and begin sketching in the major divisions of
your site, such as the main content area, the header, the footer, and navigation. Think about where you want each element to appear and how they will fit together.

It is important to catch and retain a viewer’s interest. When a first-time visitor comes to your site, he or she should be able to tell at a glance who is responsible for the site, what the web site is about, and what sort of information it contains. Be sure to include the following:

- The name of your project/institution: Think of this as the title of a book. Combining the name of your project or organization with an eye-catching graphic will give the viewer a visual to associate with your name.

- An interesting overview of the purpose of the site or the goals of your organization: This is your site’s subtitle. It should be informative enough to answer the question What is this site about? However, it should be short enough to be read at a glance.

- A simple, catchy and informative navigation bar: This is your site’s table of contents. Keep the information simple and concise so that the viewer will know where he or she is going when they click on a link.

These are all things to remember when sketching out your site. Sometimes the hardest part of web design is coming up with a good hook—fortunately, inspiration is never far away.

The best source of inspiration for your web design is other web sites. Keep track of sites you like (or dislike!) and don’t be afraid to use your browser’s View Source function to have a look under the hood of your favourite page. Looking at another designer’s HTML is a great way to learn more.

**Using Templates**

Once you have decided upon a design and layout for your website, you must then go about the construction of a template for your pages. A template is like a cookie cutter for web sites. You don’t need to build a new page from the ground up. Instead,
you build a template that incorporates all your design and functionality and use it to build the different sections of your website.

Basically, you build one skeleton page that meets your every need and save it as your template. Then, when you want a new page you can save the template to the proper location with a new filename and add any unique content.

Once you have a template and have started making new pages, it is difficult to go back and make a change in those sections that are common. Say you have a title section in your template and after you make a 100-page web site, you need to change the title. You will have to make the change in 100 separate files, which is very annoying, to say the least. While you can do a massive search and replace, this has its risks. The easiest way is to get it right the first time.

Please open the provided template. Using this template, you will build a web site consisting of four pages: index.html, images.html, video.html, and extra.html. First, you need to customize the template and make sure it will work the way you want it to. Once it is ready, you can begin building the website.

1. The first step is to give your site a title. Change the title of your site in the <head> section of the template. You will see: <title>Replace this text with your title!</title>. Just replace the text with a title of your choosing.

When looking at the template, notice first that it has a several sections, each marked off by a <div> tag that has an id attribute and a unique value. Div stands for division or divider and is used heavily with CSS-enabled websites. For now, they let us see where each section of our web page starts and ends.
2. Next, find the `<div id="header">` section and type in the title of your site, using the font or heading tags we saw earlier. Your code could look something like this:

```html
<div id="header">

<h1>Welcome to (your name)'s site!</h1>

</div>
```

3. The third step is to build the navigation scheme for the site. In this case, we will use a simple horizontal navigation bar made up of link tags and a vertical dash, also called a pipe (push Shift + \ to get a pipe). The navigation section will go in the div with id="nav" and it should look something like:

```html
<div id="nav">

<a href="index.html">Home</a> | <a href="images.html">Images</a> | <a href="video.html">Video</a> | <a href="extra.html">Extra</a>

</div>
```

The `<a>` tag can be used in a number of ways, but linking like the above is the most common. Take note that `href` is an attribute and its value is the target of the page you are linking to. This target can be a file path or an URL and can be either an absolute path or a relative path (see Appendix C to learn more about paths). The text wrapped by the `<a>` tag is the text that will be the link text for people to click on (blue and underlined) when your page displays.
4. Finally, you are ready to save a copy of template.html four times, renaming it each time to **index.html, images.html, video.html, and extra.html**. In each page you should add a heading in the `<div id="content">` section that specifies which page it is, for example in audio.html: `<h3>This is the audio section</h3>`

5. Now, open **index.html** in the browser and start clicking on the links in your browser. You should have a uniform title and navigation bar on each page. If any links are broken, go back to your code and make sure you have the file names typed properly and double check the tags, ensuring that you aren't missing a quotation mark or forward slash.
Web Tutorial 1: Image enhanced web design

Images are the most common elements in web sites, especially in sites for the digital humanities. An excellent way to display a document is to serve high quality digital images, allowing the viewer access to the original while at the same time preserving it. This section will detail some of the concerns surrounding images in digital humanities web sites as well as how to mount an online exhibition and outline some of the standards to use.

Permissions and Copyrights

The major issues surrounding offering images online are copyrights and standards. Ultimately, however, if you put an image online, people will be able to take it and do whatever they want with it. The easiest way to limit what happens to an image is to control access. Unfortunately, the goal of many digitised projects is to offer public access to the documents. In these situations, your institution likely has lawyers and copyright rules to help direct your actions. You also need to be sure of the permissions you have to use the document in question. Can you reproduce the document? Is public access to the document allowed? These are questions and concerns that will often need to be dealt with on a document-by-document basis.

One way to manage use, and at least limit the reproduction of your images is by choosing standards for your web images that limit the potential for misuse. Think about the different formats of images available to you and what you know about screen resolutions. Standard screen size is 800 X 600 pixels and a 72 dpi image looks just fine in a web browser, while also keeping file size down. You can also make this image a JPEG, a lossy format that does not reproduce well. If someone then takes such an image and tries to do anything but reproduce it on a screen they will be disappointed with the results. So, choose the appropriate quality of image to address any concerns associated with the document. If there are no restrictions on a document, then offer high quality formats. If someone really wants to use an image, they will contact you for permission and possibly for a high quality version of the
image. You should also remember that high-resolution images create large file sizes, which can result in slow load times for your site for some users.

**Size, shape and scale**

Before deciding how to lay out a page for images you need to think about how you want the images displayed, how many there are, and what format to use. Adobe Photoshop is the industry standard for image manipulation and editing, and has a number of batch processing tools that make life easier as well. Photoshop plays a major role in getting your image from a large tif file to an Internet-ready jpeg.

First you need to decide what you want to do with the picture. Is it a stand-alone image as part of a web page, or will you offer multiple images linked through thumbnails—or is it something in between?

**Preparing single images**

If you are adding a single image to your site you need to size it to fit the purpose of the page and the space you want it to fill, and save it in the appropriate format. In this example we are going to make an image that will be quite large, but the steps are the same if you need it to fill a smaller space—it just takes a little experimentation to find the right size.

- Open Adobe Photoshop
- Once Photoshop is ready, click to **File -> Open.** Find the provided folder named **images.** In this folder you will see several other folders. Open up the folder called **Master** and choose the image entitled **tutorial.tif**
- When editing from a master copy, your first step should always be to save a copy of it in a new location, to avoid inadvertently saving over the master file. Click **File -> Save As** and in the images folder choose the **edits** folder and then click **Save.**
Next you need to open the **Image Size** window. Click on **Image -> Image Size**. In this window you will see, amongst other things, the dimensions of the image in pixels as well as its actual size. First, make sure the **Constrain Proportions** option is selected, which keeps your images from being skewed. To make the image web-ready, change the width in pixels to 800 and notice how the height changes in proportion to this new value. You also should change the resolution to 72, as this is the standard for web images.

Click **Ok** to confirm these changes. The image should now be resized for you. Click to **Save As** again and make sure to change the **File Type** in to JPEG and give it an appropriate name and check the location before clicking **Save** to continue. Save the image in **images -> web_large**.
You should be prompted with a JPEG Options window where you can further tailor the image quality. Notice how you can manipulate the image quality slider. As you make the quality level higher, the file size gets bigger. Set it at high quality in this case and click Ok. You now have a JPEG image that is ready to be added to a web page.

Preparing Thumbnail Images

You may also want the option of thumbnail images linking the viewer to a larger version of the image. If this is the case, you should set standards of quality and size for both your thumbnail and larger images before you begin adding them to your site. Usually you want all your thumbnails to be a set height and width so you can build a decent layout with them. A good range is to size them from 75 to 100 pixels wide, and 100 to 130 high, but it depends on the images and their orientation. For book pages, or portrait style pictures, usually 90 wide by 120 tall is optimal. In any
case, the first step is to make the larger version of the image ready for web, as we’ve already covered, and saved in a directory for your large images.

Making thumbnails raises a few problems. Your images are usually going to be slightly different proportionally, meaning that if you resize the images to exactly 90X120 you will get skewed thumbnails. There are several ways to get around this problem, but the best way is to first **crop** your image to the approximate shape of your thumbnail, and then use the **Canvas Size** tool to trim it to fit. This exercise will take you through all the steps to prepare a thumbnail image by hand.

1. Open the master copy of the same image we used in the last exercise and save a copy of it in the edits folder as we did before. Remember this is a crucial step: the loss of a master TIFF file means going back and digitising the document all over again. The next step is to **crop** the image.

In image manipulation **cropping** refers to an action much like cutting a picture with scissors. You select the area you want to save, and with one click of the mouse, Photoshop will trim away the excess, leaving you with your selection.

2. To **Crop** the image select the cropping tool from the tools palette (or hit C). Next, decide which area of the image you want as the thumbnail. Starting from the upper left, click and hold the left mouse button and drag the pointer towards the bottom right. Make sure you select more than you need at the bottom, so there is something to trim later. You want your selected area to be about twice as high as it is wide. Once you have the area selected that you want to become your thumbnail image, release the mouse button. Make sure this is the right selection and then click the checkmark at the top of the screen to confirm.

3. You now should have an image that is roughly the right shape for your thumbnail, but you need to resize the image using the same process as in the last exercise. Click to **Image -> Image Size** and resize the image to 90 pixels wide.
4. Next you need to click to **Image -> Canvas Size** to trim the height of the image. This process can be done in other ways, but the **Canvas Size** tool offers the most precise method. In the **Canvas Size** window, first set the **Anchor** point so that the image will be trimmed on the bottom and right edges, as shown below.

![Canvas Size Window](image)

You will need to change the unit to pixels using the drop-down menu to the right of Width and Height, then change the height to 120 pixels and click OK.

This basically trims the image at 120 pixels measuring from the top. Now you have a perfectly handcrafted thumbnail image.

5. Save the thumbnail in the folder **web_thumbs** with the same name as the larger version you made in the last exercise. Make sure it is saved in JPEG format. Keeping your pictures organized this way means you only have to remember one filename; the location of the file tells you what size it is.
This process takes some time, but ensures that your thumbnails are perfect and contain the proper content. If you are less picky or have a uniform batch of files, then you can use a batch process that simply resizes all the images to 90 pixels wide and 72 dpi resolution, and then trim them to 120 pixels high. There are also features such as **Save for Web** within Photoshop and certain online tools or software that will manage, resize, and display your images for you.

**Adding Images to the Site**

After you have your images all resized and saved, you can begin adding them to your web site. If the image is being added to textual content, then you need to determine where you want it and how you want the text to align around the image. If you are going to use a thumbnail layout, then you need to determine an appropriate grid structure or some other layout for the thumbnails. These are all considerations that you may want to discuss with your team if image handling becomes a function of your web site. This exercise runs through a simple method of adding images to the website we've already constructed.

1. Open Dreamweaver
2. Open your images.html file and find the content section.
3. In the content section of the page type:

   ```html
   <img id="test" src="images/web_large/yourPic.jpg" alt="describe the content of the image" />
   ```

   This tag adds the image found at the path in `src="images/web_large/yourPic.jpg"`. That’s all there is to adding an image into the page. Within the img tag you can resize the picture using `width=""` and `height=""` attributes.

4. Save images.html and preview it in a web browser. If your image doesn’t load, the problem is most likely an incorrect path to the file; double-check the `src=""` and try again.
To add a thumbnail linking to the larger image:

- First change the src="path" to point to the thumbnail image we made.
- Now, wrap the <img> tag in a link that points to your larger image.
  
  Remember the link tag?

  <a href="path_to_file">Content</a>

Your “path_to_file” becomes the location of the larger image and “Content” is actually going to be the <img> tag. Your code for a thumbnail link should look something like:

<a href="images/large/bigpic.jpg">
  <img src="images/thumbs/littlepic.jpg" />
</a>

- Save images.html and preview it in a web browser. You should be able to see the thumbnail image. Clicking on it should open the larger image. If something goes wrong, the problem is most likely an incorrect file path, as noted above.

TIP:
Every image in your site should be given an alt value in the form of text that describes the image. This text is displayed if the user decides to disable image viewing in their browser. More importantly, this text is necessary to be compliant with web standards for the visually impaired.
As you can see, the hard work in mounting images online is the prep work with the actual image files. Once you’ve made those, you can set up a template and begin loading in images. Remember that, for large layouts of images and thumbnails, you may need to do some serious positioning and formatting with CSS. There are also tools and scripts that can be used to make this whole process easier—talk to your team about streamlining and automating the process. These tutorials have shown you some solutions, but there are many. As you gain more experience, you will get better at determining which option is best for your project.

**Examples of image-oriented sites**

Early English Books Online:

Early English Books Online - Text Creation Partnership:
[http://dlxs.odi.ox.ac.uk/e/eebo/](http://dlxs.odi.ox.ac.uk/e/eebo/)

Internet Shakespeare Editions:
Web Tutorial 2: Multimedia

Unfortunately, adding video or audio to a web page is not as simple as adding images, mostly because browsers all support different ways of adding these files. There are two main ways to include multimedia objects in your site.

This first way is to make the files available for download so that the viewer can then open the files if they have the appropriate player. This means easier coding and less organization, but also distances the viewer from your site, taking away your control over the media as well as taking it out of context. Linking to the media is done with the same tag as linking to a different web page:

```html
<a href="path_to_media/filename" alt="" title="click this link to experience multimedia" pluginspage="http://www.apple.com/quicktime/download/">Download the file</a>
```

This will open a new, blank page and prompt the browser's download window to appear. It's up to the user how to proceed from this point.

The second approach is to embed the media in your web page, which is much more difficult, but has advantages. If you decide that embedded media is the way to go, then the only way to make life easier for you and your viewers is to make several decisions before you add video to your site.

First, you need to choose a widely used video format and player that is likely to be supported by your viewers' web browsers. If your viewers don't have the appropriate plugin or software, add a link so they can easily retrieve it. Second, keep the size of the video to a minimum or offer multiple links, so viewers can choose how large a file to play, depending on their circumstances. High quality video is of no use if it takes more than a few minutes to download.
**Adding Video to Your Web Site**

For this tutorial we are going to embed a QuickTime Movie into our website.

1. Open your editor

2. Open your *video.html* page and find the content section

3. In the content section type: `<p>My video should appear below... I hope!</p>.

4. Next, copy and paste the following code into your file:

   ```html
   <object classid="clsid:02bf25d5-8c17-4b23-bc80-d3488abddc6b"
   codebase="http://www.apple.com/qtactivex/qtplugin.cab"
   width="300" height="300">
   <param name="src" value="file_name.mov">
   <param name="autoplay" value="true">
   <param name="controller" value="true">
   <embed src="file_name.mov" width="300" height="300"
   autostart="false" loop="true"
   pluginspage="http://www.apple.com/quicktime/download/"
   />
   </object>
   ```

5. Correct both src attributes to indicate an actual movie file. This gives you a basic embed tag with width and height attributes, which you may need to change if the video appears too small on the page. You can also add `autostart` and `loop` attributes, each of which have a true/false variable turning them on or off. There are other attributes as well, but this is the basic embed tag.
Doesn’t that look like fun? Fortunately you can copy this from the text file called movie_code.txt in the video folder. All the extra code around the embed tag is required for Internet Explorer to do what takes one tag in other browsers, but it is a popular browser, so you need to support it. Internet Explorer does not understand the embed tag and so ignores it, while other browsers ignore the object tag in favour of embed. The long number that is the value of the object tag’s classid is an ActiveX control for QuickTime movies; this would need to be changed if you were to use a media format that requires a different player. Fortunately, QuickTime can play several formats of sound (including MP3) and video, and is widely supported, making it a good choice most of the time for web media.

It is important to note that movies are usually large files, require the loading of a plugin, and can drastically increase the load time of a page. Usually the best way to go is to use a link to a movie that then begins to play (autoplay="true"). For audio, make it the viewer’s choice to listen or not, and be clear about what they will hear. A surprisingly large number of people find it disconcerting when a website starts talking to them.

**Adding Sound to your Web Site**

As we have just seen with video, there are two ways to add multimedia to a web site: linking and embedding. Many people find it irritating when they visit a site and sound plays automatically. Linking to an audio file puts your visitors in control of when (or if) they will listen to the sound file. For this course, we will link to audio files.

**To Link to An Audio File**

1. Make a regular anchor tag and change the href to point to the audio file:

   `<a href="audio/my-file.mp3">download my music</a>`
Web Tutorial 3: CSS (Cascading Style Sheets)

Why CSS?

One of the goals of standards-compliant web design is to separate content from presentation. This means that your HTML file should contain only content and information about that content, while the presentation side of things is handled by your CSS file. For instance, the tags <p></p> in your HTML file tell your browser that the content between those two tags is a paragraph. Corresponding information in your CSS file tells your browser how to display a paragraph.

Class conflict and invalid id's

CSS depends upon HTML code that has been properly structured and organized. The more effort you put into organizing your code, the more power over it you can have with CSS.

Think back to the template, remember all those <div> tags with id and class attributes. These attributes are like names for your tags so that your CSS can find them in the code as you would a friend in a crowd. The only difference between an id and a class is how they are used.

An id is a unique name, specific to one single tag and, by extension the content governed by that tag. Here's an example you've already seen:

<div id="header">Content</div>

There should never be another div with an id="header" though it is up to the coder to keep track, you can give multiple things the same id, but it isn’t good practice and results in invalid HTML.

A class is the catchall. You give a class distinction to tags that are similar in some way or that you want to look the same. Say you have many <h3> tags in your page, but some you want to bold and others you want to be italics. How could you use a class value to make it clear which should be which? What if you want one of them to be blue, but all the rest to be red? Your code could look something like this:
Here you have 3 different heading tags, each with different id or class values. Because of these different values you can now tell the CSS to target them in different ways. You can say to it every time you see an `<h3>` tag with a class of “italic” make the content italicized affecting the first and third lines in the example, targeting “bold” or “funky” `<h3>` tags would have a different effect. The id on the first `<h3>` will allow you to make just that one blue, and you can target all other `<h3>` tags and make them red.

**CSS and the joy of web design**

CSS is simple, that is why it is so great and why many people really enjoy the styling side of web design. Before we dive into our creative sides, we first need to understand how CSS works.

You already know that CSS works by targeting HTML tags that have different id and class distinctions. Once you have targeted the proper tags, you then have to give CSS instructions on what to do.

There is a simple syntax for CSS:

```
Target {
    Property: value;
    Property: value;
    . . .
}
```

The target is what you want to change; everything inside the curly braces says what to do to the target. Once you have targeted a specific tag or a class of tags you can begin deciding what to do to them in the form of the property, of which there are
many, and the value for the property. A property is basically an instruction, similar to an HTML attribute in that it specifies what aspect of the tag to change, be it color, font, size, position, and numerous other options. The value specifies what is to be done to the property. If you want to change the color of text to red then you use color: red; if you want all the font to be a certain size then you’d use font-size: 18pt; (or font-size: large; - there are many options).

This may seem a little strange at first, but once you get the hang of CSS and start to learn all the ways that it can affect your website you will begin to understand just how powerful a tool it can be.

A lesson in targeting

In order to make life simpler there are a number of ways to target your HTML code.

- You can target by the type of tag - h3 { color: red; }
- You can target by id using a # - #header { font: verdana; }
- You can target by class using a . - .bold { font-weight: bold; }

Or you can use combinations of these to select targets by context:

\[
\text{div}\#\text{header} \text{ h3.bold} \{ \text{font-weight: bold; } \}
\]

Sometimes it helps to talk it out. In the above example we are looking for all <h3> tags with a class="bold" that are within the div that has id="header" and we want to make the font-weight for all of these bold.

Using this syntax, and by adding id and class values to your code, you can manipulate the finer details of design for the screen, print and other media.

Adding CSS to your web site

Okay, so you have a grounding in CSS and know how it relates to your HTML code, but here’s the clincher – how do you let the browser know that it should use the CSS
file to style your site? This is done with a simple instruction line in the <head> section of your web page. Take a look at the header in template.html:

```html
<head>
<meta http-equiv="Content-Type" content="text/html; charset=iso-8859-1">
<link rel="stylesheet" type="text/css" href="CSS/main.css">
<title>Template</title>
</head>
```

The HTML links to the CSS with the line:

```html
<link rel="stylesheet" type="text/css" href="CSS/main.css"/>
```

Take notice of the href attribute as it shows the location of the CSS file relative to the web page. This line tells your browser to use the instructions in the CSS as it loads the page.

Let’s add a pre-made CSS file to your site. In the directory course/HTML/ copy the CSS folder into your working site directory, making your relative link active. Open your website in a web browser and see how it looks. This is made possible through the interaction of CSS with the consistent structure of your HTML that you developed in your template.html. With different style-sheets you can get very different looks.

**For More Information About CSS**

[http://www.csszengarden.com](http://www.csszengarden.com)
CSS Zen Garden

CSS Layout Techniques

[http://www.w3schools.com/css/default.asp](http://www.w3schools.com/css/default.asp)
CSS Tutorials
SECTION 6: WEB 2.0

Prepared by Michael Nixon
While the term remains somewhat controversial due to being criticized as a vague buzzword, “Web 2.0” still retains value as a description of the latest trends to influence businesses and their use of the internet. In 2001, the so-called dot-com bubble burst, and many businesses with shaky profit models collapsed. Based on the observation that many web-based companies survived and had important elements in common, O'Reilly started a conference series in 2004 to highlight the characteristics. Rather than describing an updated technical specification, which 2.0 usually indicates, Web 2.0 focuses on specific characteristics that define many successful web ventures:

- rich user experience
- user participation
- dynamic content
- metadata
- web standards
- openness & freedom
- collective intelligence

To help you understand this movement, each of these elements will be described in terms of their contribution to the web you use, along with examples and ways your own work can incorporate them.
Tim O'Reilly described a hierarchy of four levels for websites:

1. applications that work equally well offline or online (e.g. Google Maps)
2. applications that operate offline but gain features online (e.g. Google Docs or iTunes)
3. applications that can operate offline but gain advantages from going online (e.g. Flickr and its community-contributed tags)
4. applications that exist only on the Internet (e.g. eBay, del.icio.us, Wikipedia)

He summarized the “core competencies” of web 2.0 business as follows:

1. Services, not packaged software, with cost-effective scalability
2. Control over unique, hard-to-recreate data sources that get richer as more people use them
3. Trusting users as co-developers
4. Harnessing collective intelligence
5. Leveraging the long tail through customer self-service
6. Software above the level of a single device
7. Lightweight user interfaces, development models, AND business models

One of the reasons for an educator to understand these trends comes from Prensky's identification of generation-Y (born in the 1980's and early 1990's) as “digital natives”, in contrast to “digital immigrants” who were born earlier. He
argues that continual immersion in contemporary technologies changes the way
digital natives learn. Rather than simply shortening their attention span, it changes
the way they prefer to learn to methods that include hypermediation, random
access, and play. According to educational marketers, understanding how students
access information directs us to make it available in more applicable ways.

**Rich User Experience**

In order to keep visitors invested in a site, delivering a complete experience similar
to a desktop application became crucial. Browsers could deliver “applets” using Java
and DHTML in the early 1990’s, but these were clunky, being slow to load and
poorly integrated into the browsing experience.

*Google’s Gmail and Google Maps are the canonical examples of web-based applications
with rich user interfaces. The combination of technologies used in them was named*

“The combination of technologies used in them was named

“AJAX” in an essay by Jesse James Garrett and includes:

- standards-based presentation using XHTML and CSS;
- dynamic display and interaction using the Document Object Model;
- data interchange and manipulation using XML and XSLT;
- asynchronous data retrieval using XMLHttpRequest;
- and JavaScript binding everything together.

This explains the rise in importance of standards-based design, openness, and the
importance of several particular technologies (notably XML and Javascript).

Primarily, this impacts you in the range of web applications you have access to, and
the ease with which they can be used.
**User Participation**

One key thing that users of a site can add is their own unique data. Sites succeed that involve their users who in turn add value to it. Whether this involves breath-taking photos such as Flickr’s content or the amateur video hosted by YouTube, visitors now seem comfortable adding their own content. Another way users are involved is in testing. Given that many web applications stay in a perpetual “beta” state (again, Google provides the most obvious examples), new features can be released to a large audience to test their reception.

Taking advantage of the large number of users who aren’t in the conventional majority has been referred to as harnessing the “long tail” of the Internet. Small sites make up the bulk of the internet’s content; narrow niches make up the bulk of internet’s the possible applications.

The focus on user participation means that researchers and students can often taken advantage of a wide range of contributed data in a particular subject. Based on the community’s value of freedom, this is often available for reuse. As well, you can in turn share some of your own content.

**Dynamic Content**

In order to keep sites up to date, you used to have to access the pages marked up using HTML and add more content, then save and up the pages. The growth of database-backed sites has led to the ability to edit content “on the fly” through a form, often without using HTML yourself. Sites that do this are generically called “content management systems” and include blogs, wikis, and all sorts of other ways to edit content in an easy way.

**Metadata**

As the web grows more data-centric and various sources release their information through the use of APIs (programming interfaces that control how you can access the data), it’s grown more important to accurately describe your data. This is
especially true because the most frequent browsers of your web site are search engines which try to use metadata to determine the overall importance of your content. This growing trend means that formal metadata schemas such as Dublin Core have grown in popularity.

**Web Standards**

The web has changed a lot since the 1990’s when Internet Explorer dominated the market and frames were the latest innovation in browsing technology. The arrival of a competing browser whose adherence to specifications made Internet Explorer look dated was one factor that helped things change. Now that the web has so many “moving parts” in the form of interacting technologies, it’s become important that web programmers and designers work to standards. For page coding, this means XHTML, an XML-ized HTML, and CSS, the style sheet language for visual effects. For client-side interaction, this means unobtrusive Javascript packaged in libraries, working in the background and showing up as little as possible in the actual html file.

On the one hand, the free-for-all that was creating websites in the 1990’s made the barrier to entry really low. On the other, many of the interactions that occur now can only happen because we keep to standards-based design.


**Openness & Freedom**

People used to be reluctant to share ideas in order to keep an edge over competitors. Now people and companies tend to loosen hold over these resources because they reap more benefits in doing so. By allowing others to share ideas, their products are able to gain significant improvement and scrutiny through collaboration. This was equally true for code specifications or collections of data. Now in order to reach more users, sites are making their information and tools available for general use. One particular motivator of this was the rise of the “free
and open-source software” movement, which holds that software should be
developed for the good of the community and released for general adaptation and
modification. This spirit of production is line with most research endeavours, and
means that the production of knowledge and community enrichment is a generally
appreciated goal.

Besides making tools and data freely available, another outcome has been the ability
to migrate data in and out of sites. By not creating an environment that traps users’
data in an attempt to keep them on the site, sites gain respect and the ease with
which users can return with their data means that they’re likely to have them do so.

**Collective Intelligence**

The rise in using user-created content has led people to try use this force to
accomplish goals. For example, Amazon has a “Mechanical Turk”
[http://www.mturk.com/mturk/](http://www.mturk.com/mturk/) project that recruits people to perform “human
intelligence tasks” like group photos. Less commercially, this refers to the fact that a
group sharing knowledge has a larger combined knowledge base. This means that
many of the community that have formed on the web tend to collectively know a lot
about the subject that binds them, and can provide an excellent resource. This takes
a different and more persistent form that the latent potential of a mailing list, owing
to the ongoing conversations that are stored in many of the new media forms.

**Examples**

**Social Networks**

A social network service is a web site that builds an online community of people
who share interests by offering ways of sharing information about themselves and
explores that of others. These sites often provides ways of interaction such as e-mail
and instant messaging. The most popular services are strongly themed and often
focus on classmates and friendships; examples include MySpace, Facebook and
Twitter. Social network services rely on user participation, since all their content is
user created. They are the culmination of almost all the Web 2.0 trends. These
services provide a means of becoming connected to colleague and student alike, but beware the privacy concerns that arise from such open access.

**Mashups**

A Mashup is a web application that combines data from multiple sources into a single integrated tool. It is made possible through access to open APIs and freely available data sources. To obtain content, people use web feeds (e.g. RSS) or information from third party databases such as Amazon, eBay, Flickr or YouTube. This first occurred when people used Google Maps to plot interesting data like real estate or 911 calls.

The advent of mashups has led to a “remix culture” which thrives on the recombination of cultural artifacts.

Besides actually writing programming code to create your own mashup, you can also take advantage of tools such as the Google App Editor [http://code.google.com/appengine/], Yahoo Pipes [http://pipes.yahoo.com/pipes/], or Mozilla Ubitquity [https://mozillalabs.com/ubiquity/].

Creating a mashup is an ideal way to take advantage of existing tools; whether you have your own source of data, or you’d just like to see how a third party data set could be visualized, mashups let you publicize that intersection.

**Blogs**

The idea of a web log or online diary has been around since before Web 2.0, of course. What’s new is how commonly they’re used and the maturity of the platform. All blog software typically offers chronological linking and sorting, categorization, and is easily searchable. The popularity of blogs means that people with insights from our research areas or favourite hobbies are writing about it, and what’s more, able to receive comments and interact in a public space. Blogs have really replaced the static “personal home page” with a dynamic, social space. Of course, they need to be kept updated like anything else.
Popular examples of no-cost blogging sites are Google’s Blogger [http://blogger.com] and WordPress [http://wordpress.com]. Blogs can be used in the classroom to provide a quick feedback loop between creative writing and feedback, or to maintain discussions outside the classroom. Teachers can even maintain class blogs to keep students updated on homework and assignments.

**Feeds**

Feeds became popular around 2003 when they were primarily associated with blogs. By subscribing to an RSS (the syndication format) feed, you could keep up to date on headlines from a blog. Now feeds are being used to syndicate the format from all sorts of media, from pod- and vid-casts to social networks (e.g. Facebook, Twitter) and search agents. You can use a feed reader (a special application, either offline or embedded in a web page) to manage your feeds and monitor your favourite sites.

**Wikis**

By now, everyone’s familiar with the most popular and – in the educational world – notorious wiki, Wikipedia. While the online encyclopedia that’s edited by the community introduced us to the format of easily linked and marked up web pages, they’ve expanded to be used for a wide variety of uses. The primary focus of wiki is collaboration, and harnessing the so-called wisdom of crowds. A wiki is a type of free on-line writing space that allows users to add, modify and update its pages. If something is missing or incorrect in a wiki and permissions allow you to edit the wiki, you can easily add your thoughts or make changes to the wiki. It is essentially a fully editable web site.

Many classes use wikis to manage student or project collaboration. One of the important decisions to make is whether you’re creating a canonical reference that will be built upon by successive classes or whether you’ll create a new one for each class to chronicle their learning journey. There are many wiki alternatives for download, and it’s worth checking whether your institution hosts a specific one that you could test out.
Folksonomy

This term, coined by Thomas Vander Wal, refers to a new form of organic categorization that comes from internet users who encounter new information. Users add a keyword or descriptive phrase, and from then on, all users can find the item using the same keyword. This spreads the categorization workload around to all the participants in the community, and creates a flat hierarchy that is more representative once a consensus model emerges based on the group interaction. On the other hand, misspellings can challenge the system with orphaned content, and lazy tagging can make matters worse.

PodCasts

Podcasting is very similar to a radio broadcast with the main differences being transmission via the Internet and producing pre-recorded content. Podcasts can be enhanced through the use of images, video as well as links to web site content. Broadcasting over the internet offers students and teachers access to a worldwide audience, with authentic feedback from around the world.

Usually, the multimedia files are uploaded to a blog via a podcasting service and posted to the internet. This content can then be syndicated through RSS feeds and downloaded or streamed by the listener.

After creating the audio file on your own computer (using, say Audacity and the skills learned in the Audio section of this course), you can create a podcast using a hosted service such as podomatic [http://www.podomatic.com/].

Creative Commons

The wide variety of work available on the internet can create a lot of confusion around ownership rights. It's hard to tell what's “public domain” and when the copyright holder should be contacted before use. The ease of downloading means that that is the simple route is often the default one. In the spirit of openness, the Creative Commons (http://www.creativecommons.org/) – a non-profit group devoted to expanding the range of available creative work – created some clear and
easy to use licenses based on copyright. These licenses clearly the ways in which a work (be it an image, video, song, or writing) can be reused. By using a Creative Commons license, you can clarify how your work can be reused, whether freely remixed without attribution or only with proper attribution.

**Summary**

The Web 2.0 movement and the trends involved have served to make more tools freely available to teachers, researchers, and learners. These tools can help them collaborate with their peers, visualize information in new ways, or delve into niche communities. The biggest challenge can be keeping track of the most important tools and information. Web feeds help on this front, especially once you develop a list of canonical news sources that keep you up to date on subjects of interest. Hopefully this section has helped you understand what's going on with Web 2.0!