

# Information Everywhere

Understanding how to cope with and use **information** is one of the most important aspects of digital literacies. Many people nowadays believe that digital technologies have brought about a phenomenon known as **information overload**, (Waddington, 1998), a condition characterized by increased levels of stress, confusion and difficulty in making decisions resulting from having 'too much information'.

In this chapter we will argue that the problem of 'information overload' is not so much one of 'too much information', but rather one of defining what we mean by information in the first place, and of understanding how to create it by forming strategic relationships between different pieces of data. While digital technologies have dramatically increased people's opportunities to create information, they also provide extremely sophisticated tools for filtering and channelling information. Coping successfully with information involves understanding both the information creating and the information limiting affordances of digital media.

## INFORMATION AND RELATIONSHIPS

Much of the concern about information overload comes from a fundamental misunderstanding of what information is. Think about walking on a busy city street. All around you things are happening. There are thousands of sights and sounds, text everywhere, from shop signs to advertisements on the sides of passing busses, people all around you talking, dressed in different clothes and wearing different expressions on their faces. Most people who find themselves in such situations do not feel they are suffering from 'information overload' because they do not consider everything that is happening around them to be information. They selectively pay attention to and process the **data** which they judge to be important for them. In other words, they create information from the data that is available.

And so the first distinction we need to make is between 'information' and 'data'. Data are 'facts' (including sights, sounds, colours, words) which exist in the external world. These 'facts' only become information when we create some kind of relationship with them.

Besides data and information, there is also a third category that we need to consider, and that is **knowledge**. Knowledge is what is created when information is integrated into our minds in a way that we are able to adapt it to different circumstances and apply it to analysing and solving problems. Knowledge is created when information is transformed in some way - when, for example, it is combined with other information or applied to a particular task in a useful way.

Let's return to the busy city street and consider how the concepts of data, information and knowledge apply. In that environment, when we read the data made available to us on signs and connect it to the data on the map we are carrying, we have created information; when we remember that information and interpret it to the extent that we can not only get from one place to another without having to read the signs, but we can navigate through the city in innovative and creative ways, discovering 'short cuts' and 'scenic routes', then we have created knowledge.

The biggest problem people have in 'managing information' is not that there is too much information but that we have too much data available to us, and we are sometimes not sure how to decide which of it is worth turning into information and knowledge. In other words, we have not adequately worked out how to filter all of the data that is available to us and in a way that results in useful information. The good news is that while digital media add to our confusion by making much more data available to us, they also provide a host of tools for filtering data, and for forming the kinds of relationships that transform data into information and information into knowledge.

And so, to sum up, information is not about 'facts' so much as it is about the relationships that we create between ourselves (and other people) and those 'facts', and between different 'facts'. In the first chapter we argued, quoting Michael Schrage (2001), that what is often referred to as 'the information age' is more accurately thought of as the 'relationship age', mostly because people seem to use computers as much to connect with and communicate with other people as they do to search for, store and manipulate information. Now we would like to take that idea even further, arguing that even these practices of searching for, storing and manipulating information are more a matter of relationships than they are of data itself. In other words, we would like to argue that information is most usefully seen not as a collection of 'facts', but as a social practice based on establishing relationships.

### Activity 2.1: Reflecting on your information management habits

#### A. WHAT'S THERE

Think about all of the data that currently exists in all of your personal storage areas, including the hard drive of your computer, any space on servers or 'virtual disks' that you use, and any webmail or other messaging services (including Facebook) where you store messages and other data as well as physical spaces like bookshelves and desk drawers.

1. How much of this data would you consider 'information' and how much would you consider 'unprocessed data'? Give examples.
2. How much of this 'information' do you think you have successfully turned into 'knowledge'? In other words, how much of it have you been able to integrate with other information in ways that help you to formulate new ideas or solve problems?
3. How much of this data was 'pushed' onto your computer (or other storage area) without you asking for it? Give examples.

4. How much of this data did you actively go out and retrieve from some other place? Give examples.

#### B. WHAT YOU DO WITH IT

Think about how you manage, organize and use the data in your personal storage areas, including the tools and techniques you use to create information out of data and for limiting the amount of irrelevant data you are exposed to.

1. Do you have a system for organizing your data/information? Is this system useful in helping you find data to create information and find information that you can use to create knowledge?
2. What tools or techniques do you use to access data/information in your own personal storage areas? How effective are these tools and techniques?
3. Do you have a method or methods for identifying and removing irrelevant or unnecessary data from your personal storage areas?
4. Do you have a method or methods for preventing other people from pushing irrelevant or unnecessary data into your personal storage areas?

#### C. HOW YOU FEEL

1. Do you often feel frustrated when trying to locate relevant data? Explain.
2. Do you often become distracted from what you are doing by the other data available to you or being pushed to you?
3. Do you experience frustration at the amount of electronic data you need to process daily?
4. Do you spend a lot of time organizing and processing data which in the end turns out not to be very useful to you?
5. Do you have the constant feeling that there is data in your personal storage areas that you have failed to process correctly or fully? Do you feel like you have processed the wrong things or have failed to process the right things?

### ORGANIZING DATA

The first step in creating information is having data available to us in a way that makes it easy to form useful relationships with it. Throughout history, human beings have come up with various systems of organizing and classifying data. An **organization system** is any system which makes it easy for us to locate the data with which we can form meaningful relationships in order to create information and, eventually, knowledge. Organization systems usually arrange data in relation to other data. They can exist in books or online or even in the arrangement of physical objects (the layout of the streets in a city, for example, can be seen as an organization system).

The most widely used organization system is the **hierarchical taxonomy**. The eighteenth century Swedish botanist Carl Linnaeus is usually considered the father of modern

taxonomies. He developed a classification system for plants and animals that is still used by scientists today. A nested hierarchy which classifies all living things first into one of three 'kingdoms', which are then subdivided into classes, orders, genera and species.

Among the best known classification systems of this type are those we find in libraries such as the Dewey Decimal System, developed by American librarian Melvil Dewey in 1873, which consists of 10 main classes of subjects divided into 100 divisions and 1,000 subdivisions, and the Library of Congress System, developed about a quarter of a century later, which allows for more growth of topics and subtopics by using a combination of letters and numbers.

In order to be used effectively, organization systems also usually require technological tools to act as an interface between the system and its users. One of the greatest advances in information technology in the nineteenth century, for example, was the invention of the filing cabinet and its close cousin, the library card catalogue cabinet. You may think that a filing cabinet is not such a big deal, but at the time it was. The idea that you could store documents vertically in different folders for easy classification and access was a tremendous advance. In fact, when the filing cabinet was invented it won a Gold Medal at the 1893 World's Fair.

The development of hierarchical taxonomies and technological tools such as filing cabinets, which make searching through them more efficient had a profound effect on the way we interacted with data. It changed what we were able to do by making it much easier to locate the data which we needed to create meaningful information. It also affected the nature of the meaning behind these relationships since data were arranged in a way that much of their meaning was derived from their relationship with other data. It changed human relationships and identities by giving to certain people and institutions the power to control knowledge by deciding what belonged in the classification system and where it should be classified. Finally, it changed the way we thought about data, first by encouraging us to think of the natural and social worlds as rational and orderly, and second, by making us think of data itself as 'information' (since the classification system itself seems to create a kind of predetermined relationship between pieces of data).

There are, of course, distinct disadvantages to organizing data in this way. Often it is difficult to know exactly where different items should be placed in the series of categories. Classifying books, for example, is often extremely subjective. Whether or not a particular volume is about linguistics or psychology, or about anthropology or sociology, or about history or politics, or even if a work is fiction or non-fiction is often a matter of debate. And so, placing an item in this system always involves judgment, and that judgment is often a reflection of a particular agenda or ideology (see Chapter 7).

The second problem with such a system is that items can occupy only one place in the system at a time. Library books can only have one call number, and no other book can have the same call number. Documents can usually only occupy one folder in a filing cabinet. Of course, it is possible to make multiple copies of a document and store it in multiple places, but after a while this becomes unwieldy, expensive and, in the end, undermines this whole system of classification, which is premised on the idea that there is a place for everything and everything should be in a particular place.

Finally, this system does not really work the way our minds work. We do not usually think in strict categories and hierarchies. Instead, we think associatively, making connections between topics based not just on whether or not they are hierarchically related to each other, but on whether they are related in a whole host of other ways, many of which may not be immediately obvious to others or even to ourselves.

Unfortunately, even up until today, many people have not gone beyond the filing cabinet method of organizing data, despite the fact that digital tools present all sorts of new possibilities. Many 'new technologies' have incorporated the 'old technology' of the hierarchical taxonomy. The file systems on both Windows and Mac OS, as well as many web portals and directories (see below) rely on this system, based on the metaphor that data are like pieces of paper that can be classified and stored in different folders.

## NETWORKS AND ORGANIZATION

One of the greatest contributions of digital media when it comes to the way we create and use information has been to make available another way of organizing data which is actually closer to the way the human brain works, a method based on **networked associations**. What we mean by networked associations is that pieces of data can now be easily linked with other pieces of data based on all sorts of different relationships other than simple hierarchy. This capacity to link different pieces of data together based on a variety of different relationships, in fact, is one of the most important affordances introduced by digital media.

The internet, with its collection of linked web pages, constitutes a kind of natural system of networked associations that has developed organically from the bottom up based on the kinds of relationships people saw between the data they created and encountered on the internet. While this is an extremely powerful way of organizing data, using it requires a very different set of techniques and tools than those we previously used with static classification systems and hierarchical taxonomies.

Besides linking (see Chapter 3), another new tool for organizing data which digital technologies make available is the ability to 'tag' data with **meta data** or data which describes a particular piece of data with a set of concepts or references to other data. With **tagging**, rather than putting items into different folders, we put different labels onto the items, and then we can search for them later using single terms or combinations of terms. This allows our information to be organized in many different ways in many different places at once.

The good thing about tagging is that it allows people to organize information in the same way that the brain does, using multiple, overlapping associations rather than rigid categories. For example, in Apple's iPhoto I can tag a picture of my puppy wearing a Santa hat with labels like 'puppy', 'cute', and 'Christmas'. Then when I look for that picture I can find it by using any of those key words. More importantly, I can also find other similar photos using combinations of these words and Boolean operators (and/or). For example, I can combine 'cute and Christmas' and find not just my puppy but all sorts of cute things having to do with Christmas like elves and Rudolph the Red-Nosed Reindeer. Or I can combine 'puppy and Christmas' and find all of the pictures that I've taken of my dog on Christmas for the last five years.

Although many software applications support tagging, and you can attach metadata to files in both Microsoft Windows and Mac OS, the place where tagging has become a really central strategy for the organization of data is on the World Wide Web. In the past, whenever we wanted to get data from a public source, like the library, we relied on an expert to organize and classify it for us. The problem was that we were stuck with whatever classification system the expert had settled on, whether we liked it or not. Nowadays, many websites, rather than relying on experts to organize and classify information, rely on users to do this

work. On sites like Delicious, Flickr and Amazon, users attach tags to different files or other data that they either upload there or find there, and other people can use these tags to look for things that interest them. Of course, everybody has a different way of tagging based on their own judgments and their own opinions. I might tag a picture of Lady Gaga with words like goddess and sexy, and somebody else might use words like devil and obscene. So, if everybody is adding their own tags to the same piece of information, why doesn't this result in chaos?

The answer is a concept that the philosopher Pierre Lévy (1997) calls **collective intelligence**, a concept that we will discuss in more detail in later chapters, especially Chapter 11. The idea of collective intelligence is that if lots of people make decisions about how something should be classified or organized and you put all of these decisions together, you end up with a system that reflects the **collective 'wisdom' of the community**.

Many sites that use what has come to be known as **social tagging** present the results in the form of a **tag cloud** so you can visually understand how other people have tagged a particular item or group of items. In tag clouds, the more often a term is used to describe a particular item, the larger the word will appear.

For example, users of the site Library Thing (<http://www.librarything.com/>) tag books they have read with different key words. So, if you look up the book *Collective Intelligence* by Pierre Lévy, you get the tag cloud shown in Figure 2.1.

One advantage of this practice of displaying tag clouds is that it not only gives you a good way to understand what other people thought the book was about, but it also allows you to search for other books you might be interested in based on key words that are associated with this book.

These systems of classifying data that are invented by the people who actually use the data are called **folksonomies**, as opposed to the rigid, hierarchical taxonomies which we discussed above. Taxonomies are 'top-down' - that is, they are invented by experts. Folksonomies are 'bottom-up' classification systems. Some people argue that folksonomies are better than taxonomies because they better reflect the way real people think and classify information in their own minds. Others, however, are more sceptical. The writer Cory Doctorow (2004), for example, points out that collective tagging does not necessarily result in better classification systems for three reasons: 1) people lie (sometimes they just tag things randomly or try to confuse other people); 2) people are lazy (they don't think hard enough when they are doing their tagging); and 3) people are 'stupid' (most people are not very good at thinking up useful tags).

**anthropology** cognition cognitive science collective intelligence  
 computers cultural anthropology cultural studies cyberculture cyberspace emergence  
 essay essays France globalization hypertext intelligence Internet internet studies media  
 media studies media theory new media non-fiction philosophy  
 professional social systems sociology technology web web 2.0

Figure 2.1 Tag cloud  
 (Retrieved February 12, 2011 from <http://www.librarything.com/work/323515>, used with permission)

It should be obvious by now that organizing data based on associative networks as opposed to hierarchical taxonomies dramatically changes what we can do with data and the kinds of connections and meaningful relationships that we can form with it. Hyper-linking and social tagging have also had a profound effect on social relationships, to some degree shifting the power to 'create knowledge' away from experts and towards the people who use the data. While this can discourage the ideological control of knowledge by a few powerful people, at the same time, as we will discuss further in the next chapter, it can sometimes make the ideological agendas behind the various associations formed by hyperlinks and tags less transparent and easy to detect.

## Case Study 2: Search engines

Perhaps the most important and widely used digital tool for turning data into information is the internet search engine. Over the years there have been many different approaches to searching the internet, but nearly all search engines consist of three main components: 1) a **crawler or spider**, which is a software program that travels through the World Wide Web and retrieves data to be indexed; 2) the **indexer**, which arranges what has been harvested into a form that can be searched by the user; and 3) the **interface**, which consists mainly of a group of **algorithms** or sets of procedures by which the index is searched and the results of the search are sorted. All three of these components present special kinds of challenges for the designers (and users) of search engines. When they work well, however, search engines provide the enormous affordance of freeing us from hierarchical taxonomies and allowing us to take advantage of the associative, networked structure of the internet.

Search engines were not always the preferred way for locating data on the internet. In the early years of the World Wide Web and even into the first years of the twenty-first century, 'directories' or 'web portals' were much more widely used. Portals, like Yahoo and AOL, were originally web pages with lists of links arranged in hierarchical taxonomies according to subject along with, as they developed, more things like news stories, weather reports and horoscopes. In fact, the development of the World Wide Web in the 1990s can in some ways be seen as a competition between the two systems of organization discussed above: the hierarchal taxonomy and the associative network.

The problems with using a **directory** to manage data on the World Wide Web are obvious. First, there is just too much data to fit realistically into a directory, and so the links that are included must always be selected by some central authority. Second, the larger a directory gets, the more time and labour intensive it becomes to search. And finally, as we stated above, directories lock users into rather rigid conceptual categories that may not match with the way they divide up data in their own minds.

Search engines also have problems, mostly having to do with the special technological challenges associated with the three components mentioned above. The first challenge is to develop a **crawler** that can harvest the massive store of data on the

web both thoroughly and efficiently. The second is developing a method for indexing the data so that the right kinds of search terms result in the right kinds of results. For example, if you are searching for York, you are probably more interested in York, England than New York, though most of the web pages on the internet containing the word York are about New York. Similarly, if you type in the name George Washington, you are likely more interested in pages about George Washington (the person) rather than pages that just mention him or pages about the George Washington Bridge or George Washington University. Lastly, there is the challenge of developing a set of procedures which will return results in a way which can help the user to judge their relevance to what he or she wants to know or do and facilitate the forming of useful relationships with and among these results. This set of procedures is called an 'algorithm'.

Over the years, different developers have gone about solving these problems in different ways. Perhaps the first great advance in search engine design came with the 1995 launch by Digital Equipment Corp. of Alta Vista, a search engine which, for the first time, made the efficient crawling and indexing of the web possible. The problem with Alta Vista and many other search engines of this period was that they lacked an effective algorithm with which to judge the relevance of results and so were open to abuse by 'spammers'.

'Spam' is a term used for unsolicited and usually unwanted data which is pushed onto your computer. The type of spam most familiar to us is email spam, but another important kind of spam is known as 'search engine spam', which refers to web pages which attempt to fool the indexing systems of search engines and 'impose themselves' into the results of unsuspecting searchers. Back in the 90s the most popular method for doing this was 'keyword stuffing' - filling webpages with popular keywords, often hidden (for example white text against a white background) in order to fool crawlers and indexers. For example, a pornography site might secretly embed the names of popular entertainment figures in order to trick search engines into listing them as the results for popular searches.

One of the most important developments in search technology was the invention of the **PageRank algorithm** in the late 1990s by Larry Page and Sergey Brin, two students at Stanford who went on to found Google. PageRank is based on the central idea we used to introduce this chapter: that information is not about 'facts', but about *relationships*. Thus, the 'information value' of any given piece of data comes from the number and strength of the relationships it has with other pieces of data and with other people. PageRank sorts search results in terms of 'relevance' based on the number of other sites which link to them and the quality of these linkages. In other words, the more sites that link to a given site, the more 'important' that site is deemed to be. Not all relationships are equal, of course. If your brother links to your site, that may help you, but not much because not many sites have linked to his site. If, on the other hand, *The New York Times* links to your site, your site will go up in 'information value' since so many other sites have linked to *The New York Times*.

The PageRank algorithm can be expressed mathematically as:

$$PR(A) = (1-d) + d (PR(T1)/C(T1)) + \dots + PR(Tn)/C(Tn)$$

where

PR(A) is the PageRank of page A,

PR(Ti) is the PageRank of pages Ti which link to page A,

C(Ti) is the number of outbound links on page Ti and

d is a damping factor which can be set between 0 and 1.

This algorithm is not the only way that Google ranks search results, but it is the most important and, when it was developed, the most revolutionary method of returning relevant results, resulting in the meteoric success of Google as the search engine of choice for most people at the time this book was written.

Besides the PageRank algorithm, Google also uses other kinds of **data signals** to determine the relevance of particular websites to users' queries. Starting in 2009, it added personalized search features (see below). What this means is that the results you get are not just affected by other people's behaviour in linking pages to one another, but also by your own past search behaviour - the terms you have searched for before and the results you have clicked on. These and other factors go together to help the search engine determine the relevance of particular sites for you (or for the kind of person the search engine thinks you are). If you frequently search for and click on links related to George Washington University, for example, when you search for George Washington, or even just George, chances are that the university homepage will appear high in your results. Personalized search also takes into account your location, and so the term George is more likely to return George Washington University if you are searching from Washington D.C. than if you are searching from Cairo. Google's personalized search also makes use of data that is present in other Google apps such as Gmail and Google Docs. It searches through the emails you have sent and received, for example, and cross references that data with information on what you have searched for in order to get a more accurate idea about what kinds of search results you will find most relevant. Some people are concerned that such practices might violate users' privacy or might make search results too biased towards users' preconceived opinions or ideas about the world (see Chapter 7).

PageRank and personalized search algorithms, however, are not immune from search engine spammers. Search engine spammers have, over the years, attempted various methods to trick the algorithms of search engines, including creating 'link farms' which have no purpose other than to link to other sites and increase their ranking, and installing malicious **cookies** (see below) on people's computers. Consequently, Google and other search companies constantly 'tweak' their algorithms in order to get around such activity, a process which unfortunately sometimes causes legitimate sites from honest purveyors of data to fall in the rankings.

The lesson that can be taken from this brief survey of search is that no matter how good search engines become at delivering to you the data you want and need to

create useful information, there will always be limitations. There will always be people who will try to figure out ways to beat the system and to push to you data that you do not want, and sometimes attempts to stop them or to 'personalize' your search can end up making the results too narrow.

Ironically, this situation helps to highlight the affordances of directories, those unwieldy and inefficient systems of organization that search engines promised to save us from. For all of their limitations, at least directories come with a different kind of filter: the judgment of the person who constructed the directory. And so, for certain kinds of data retrieval, especially those in which the reliability of data is very important or which focus on rather narrow disciplinary domains of knowledge, directories can be better choices than search engines. A good example of a useful directory is the Librarians' Internet Index (LII) (<http://www.lii.org/>), which promotes itself with the slogan: 'Information you can trust'. The site classifies internet resources by subject, including only those which are deemed by the participating librarians as reliable and relevant to academic or educational pursuits.

One final and very important point about search is that the words people type into search engines themselves become data, especially for people trying to sell things. The words you type into Google's search box as well as the results that you have clicked on, for example, are used to help advertisers to determine which ads might be most appealing to you. Civil libertarians have voiced concern over such practices, citing the invasion of privacy they constitute and wondering about the possibilities of governments using this information to conduct surveillance of citizens.

Aggregated data on search terms is also used by businesses, journalists, anthropologists, historians and other scholars to measure trends in what people are interested in or concerned about. John Battelle (2005), former editor of *Wired Magazine*, calls this aggregated data the 'database of intentions', a record of the thoughts, ideas, desires and fears of human cultures. One important step towards the maintenance of this database is Google's practice since 2001 of publishing yearly aggregated search statistics under the title of 'Google Zeitgeist' (<http://www.google.com/intl/en/press/zeitgeist/>).

## FINDING AND FILTERING

In his book *The Advent of the Algorithm*, mathematician David Berlinski argues that the development of the concept of the *algorithm* at the end of the nineteenth century was the basis for all modern science, especially modern computing. Computers, he says, are physical machines that embody algorithms. Although the mathematical idea of the algorithm is quite complicated, its essence is simply a finite list of precise steps for executing a particular task. A recipe for chocolate chip cookies, for example, can be seen as a simple kind of algorithm.

As we have seen in our case study, algorithms are central to the way search engines work. In fact, they are central to the way all computer programs work. Google's PageRank

algorithm, for example, is a series of steps that a computer program executes to determine the number of links to a given webpage matching a particular search term and the value of those links based on the number of links to those linking pages.

The main actions achieved by algorithms are *finding* data that is potentially useful to us based on input we have provided (e.g., search words) and *filtering out* data that is deemed not to be useful to us. Many of the tools that we use to find and filter data are based on very simple algorithms. For example, an **RSS feed** delivers to us blog postings based on a simple formula that selects only postings from blogs we have subscribed to.

Not only is the algorithm an important mathematical concept for digital media, it is also a useful tool for understanding how to manage information. Many of the algorithms involved in delivering data to us are executed by technological tools, mostly computer programs. We can call these **technological algorithms**. One step in making more efficient use of tools like search engines is to understand the algorithms that they use to determine which results to deliver to us.

There are, however, other kinds of algorithms. Some make use of people or groups of people to find useful data and protect from useless data. We can call these **social algorithms** or **social filters**. By participating in social networks (e.g., Facebook, Twitter), for example, which are made up of our friends, colleagues, and other people we respect and trust, we gain access to the data that they consider to be important. In the course of the day, our friends and contacts browse a range of sources on the internet, coming across a lot of uninteresting junk, and a few interesting pieces of data that are worth sharing. The uninteresting junk is 'filtered' out, and the interesting or useful data is passed on by posting a link to a social network site like Facebook, Twitter, or to a social bookmarking site like Delicious or Diigo. When we log on to these networks, we can filter data based on the recommendations of our friends.

Social algorithms are an example of what are known as **agent-based algorithms**, algorithms that involve the interaction between a set of procedures and some kind of agent who can apply the procedures in an intelligent way based on the demands of different circumstances. In many ways, agent-based algorithms are a more accurate means of finding and filtering data than technological algorithms alone. The mental processes that agents use to make decisions about what we may or may not find useful or interesting are often much more sophisticated and complex than those used by computers, taking into account a lot more information and making use of 'fuzzy logic'. This is why your boyfriend is better at choosing a birthday gift for you than a computer (you hope). And so, social algorithms help us to create information not just by forming relationships with data but also by forming relationships with other people whom we trust.

**Personalized algorithms** are algorithms that change based on your own choices or behaviour. Perhaps the most basic kind of personalized algorithm is the RSS feed. RSS stands for 'Really Simple Syndication'. It allows owners of a website to send alerts to people who have subscribed to the feed whenever the content of the website changes and to 'push' that new content to subscribers' computers. Subscribers often use RSS aggregators (called 'readers' or 'feeders') to gather together the content of the different websites they have subscribed to.

More sophisticated personalized filters are those that filter data based not on conscious decisions that you have made but on your *past behaviour*. Such filters are essentially prediction engines which examine your past actions on the network (such as what links you have clicked on or what kinds of things you have 'liked' on your social networking site) and

try to guess the kind of data (or products) that you are likely to find most interesting and useful in the future. These actions are often recorded and sent back to the server using small programs, which have been installed onto your computer, called 'cookies'. The search personalization features that are part of Google's set of search algorithms that we mentioned in our case study constitute an example of such a personalized filter. Another good example of a site that personalizes search based on users' past behaviour is Amazon.com, which returns results and 'makes recommendations' based on the items you have previously browsed and bought when logged into the site.

While the filtering capacity of algorithms has the enormous advantage of protecting us from unwanted data and making our searches more efficient, there are disadvantages as well. Sometimes filters filter out potentially useful information along with 'spam'. No matter how sophisticated an algorithm is, it can never be absolutely sure about the data you need. Results returned based on criteria such as popularity and relevance are not necessarily the best. Sometimes data that is less popular or falls outside the scope of the things you are usually exposed to is much more useful. It is important, therefore, to be aware of the ways the various algorithms we use filter data and to critically evaluate the criteria they use (see Chapter 7).

Apart from the filtering procedures performed by technological, social and personalized algorithms, there are also the procedures you follow in your own mind when you set about looking for relevant data. All of us bring to the task of information creation certain **mental algorithms**. In other words, whether we are conscious of it or not, we always execute a series of steps when we go about searching for useful data, and just as understanding how Google chooses and filters data for us can help us to use it better, becoming more conscious of the algorithms our own minds execute will make us better at finding useful data and creating useful information. These mental algorithms guide us to perform different procedures based on things like the *purpose* of our search, help us to determine appropriate *search terms*, and assist as we *interact with* and *evaluate* our results.

### Recovery or discovery?

One of the most important things to determine before searching for data is what your purpose is for doing so. In some cases, we are attempting to *recover data* which will help us to solve a particular problem. At other times, however, we are not trying to find a specific piece of data, but rather to *discover what kind of data is out there associated with a certain topic or issue*. Some tasks involve both recovery and discovery. Doing research is a good example. Most good research is a process of attempting to discover interesting questions, recovering data that can be used to answer those questions, and then discovering more questions. Most tasks, however, involve primarily recovery or discovery, and if you get lost discovering when all you really need to do is find the answer to the question you have, you end up wasting a lot of time. At the same time, if you are trying to find specific answers before you have sufficiently discovered the best questions to ask, you will also be led astray.

### The right word

Searching for data is fundamentally a *linguistic activity*, and for it to go smoothly requires that the words and phrases that you choose as search terms are accurate in terms of mean-

ing, **syntax** and, of course, spelling. Meaning has to do with whether or not the search terms you are using semantically match the kind of data that you want to find. Sometimes you will use words that you believe *describe* the kind of data you are looking for, and sometimes you will simply choose words that you think are in some way *associated* with that data. In choosing search terms, it is important to be flexible and try new terms when the old ones have not delivered satisfying results. Often the search results themselves can help us to refine our searches, suggesting more specific terms or related terms we had not thought of.

Syntax is about the way you combine words together and the order in which you combine them. The syntax that most search engines use is called Boolean logic. Boolean logic does not pay attention to the order of words unless you tell the search engine to. And so, typing 'Lady Gaga' is the same as typing 'Gaga Lady'. If you wish for these words to be found in a particular order as, for example, part of a phrase, you need to surround them with double quotation marks ("). Although some search engines purport to work on the syntax of natural language, most, like Google, do not, and so asking a question like 'Where was Lady Gaga born?' is no more effective than typing a series of words like 'Lady Gaga birthplace'. The most important thing is to be brief, avoiding extra words that might confuse the search engine. Google actually screens out common words like 'the', 'when' and 'where', and so if you want these terms to be included you need to precede them with a '+' - if you want to search for pages that contain either one term or another, you can separate the two terms with 'or'.

### The pragmatics of search

**Pragmatics** is a branch of linguistics which studies how we do things with words' and how conversations can be interpreted as a series of *actions and reactions*. To approach your search pragmatically means to first realize that, usually, you are not just trying to create information for the sake of information, but that you are trying to do something like buy a birthday gift or complete a school assignment. Sometimes people get so wrapped up with collecting data that they forget to ask whether or not the data they are collecting is really going to help them to solve their problem. To avoid this, you can 'think backwards': first determine what you need to accomplish, then determine what kind of information you need, and then think about the kind of data you need to create that information.

Another important aspect of search is that the process of searching for data, whether we are using technical algorithms, social algorithms, personalized algorithms, or relying on our own mental algorithms, is always a kind of *conversation*. When your search engine or social network or whatever tool you are using delivers its results, that is not the end of the conversation, but rather an invitation for you to respond, either by 'interrogating' the results that you have gotten or by further refining your search strategy.

### Evaluating data

We will have much more to say about how to 'interrogate' and evaluate data you find on the internet in Chapter 7. For now it will suffice to return to the notion of relationships with which we began this chapter. Usually the best way to judge data you find on the web is to examine the relationships they have to other pieces of data and to different kinds of people who have either produced, used or recommended them. In looking for relationships with people, for

example, you might ask who produced the data and what kinds of affiliations they have (e.g., a university, a religious organization, etc.). When it comes to publically available data, it is important to remember that all data has an agenda (See Chapter 7), which is usually to convince you of a certain idea or position or to make you do something (like make a purchase). In looking for relationships between this data and other data, you can consider how the data fits in with the data you already have, the relevance of the data to what you have to do, when it was produced and when you need to use it, and the other sources of data linked to it.

### Activity 2.2: Working with algorithms

In this activity you will reflect on and work to refine the various 'algorithms' or sets of procedures for gathering data that help you complete everyday search tasks. Plan a search strategy appropriate for the following scenarios:

1. You want to locate an old friend you went to primary school with.
2. You are writing an essay on whether or not 'collective intelligence' is really a form of intelligence.
3. You want to know how PageRank works.
4. You are going to the beach and you want to know how long before going out in the sun you should apply sunscreen.
5. You want to buy a portable hard disk.
6. You want to find out if a certain professor would be a good supervisor for your planned PhD studies.

For each scenario, write down the procedures that you would go through to locate and retrieve the appropriate data. Provide a list of steps or a 'flow chart', noting when the procedure involves a technological algorithm (such as one provided by a search engine), a social algorithm (such as recommendations from friends in your social network), a personalized algorithm (such as a set of choices you have made in the past), or a mental algorithm (a set of mental procedures you perform yourself such as brainstorming appropriate search terms). As you are doing this, consider the following questions:

1. Does this task primarily require a process of recovering data or discovering data?
2. What kinds of tools (e.g., search engines, social filters) are most appropriate for the job?
3. What do you know about the algorithms that these tools use that can help you to interpret the results that you get?
4. What kinds of questions will you ask about the data you find in order to evaluate its accuracy and relevance to your task?

Most good algorithms have built into them alternative strategies, so make sure to include in your plan alternate routes to take if your initial attempts do not yield the kind of data that allow you to create the information you need.

### CONCLUSION

In this chapter we have attempted to address the problem of 'information overload' often associated with digital media by making a distinction between data and information. Data is something that we either find or is 'pushed' to us by other people, and information is something that we create by forming useful relationships with this data and among different pieces of data. Forming these relationships always involves using various kinds of tools such as organizational systems and algorithms, and the affordances and constraints of these different tools affects the kind of information we are able to create and what we are able to do with it. In the next chapter we will further explain how we navigate through and create relationships with online data as we consider the new ways of reading and writing that digital media make possible.

### USEFUL RESOURCES

#### Print

- Battelle, J. (2005). *The search: How Google and its rivals rewrote the rules of business and transformed our culture*. Boston: Nicolas Brealey Publishing.
- Berlinski, D. (2000). *The advent of the algorithm: The idea that rules the world*. New York: Houghton Mifflin Harcourt.
- Davis, C. (2005). *Finding and knowing: Psychology, information and computers*. London: Routledge.
- Jones, W. (2007). *Keeping found things found: The study and practice of personal information management*. Waltham, MA: Morgan Kaufmann.

#### Web

- Library thing  
<http://www.librarything.com/>
- Google guide  
<http://www.googleguide.com/>
- Delicious  
<http://www.delicious.com/>
- Librarians' internet index (LII)  
<http://www.ipl.org/>
- Google zeitgeist  
<http://www.google.com/intl/en/press/zeitgeist/index.html>



## CHAPTER 7

## Critical Literacy

So far we have focused on various affordances and constraints of digital media and explored how these affect the ways people do things, make meanings, manage their relationships with others, construct their identities and think. In this chapter we will take a more critical stance towards digital technologies, attempting to discover how these affordances and constraints embed particular **ideologies** and the agendas of particular people or groups. The word 'critical' as we are using it here doesn't mean that you need to find fault with digital media or the practices people engage in with it. What we really mean by a critical stance is a *conscious stance* – a stance that puts you in the position to 'interrogate' the ideologies and agendas promoted in the texts that you encounter via digital media and by digital media themselves.

## TECHNOLOGY AND IDEOLOGY

Ideologies are systems of ideas, practices and social relationships that govern what is considered right or wrong, good or bad, and normal or abnormal. These systems are extremely powerful. They can determine which people in a society are included and which are marginalized, who has power and who doesn't, and how wealth and other resources are distributed. The problem with ideologies is not that they are necessarily bad, but that most people are unconscious of them, taking the particular ideology they subscribe to as the 'truth' rather than simply as one of many possible 'versions of reality'. As a result, it is difficult for them to question the assumptions of their own ideology.

One place where we can see ideology at work is in the way people talk about digital media. As we mentioned at the beginning of this book, many people are upset about digital technologies and the new kinds of literacy practices associated with them. They are afraid these technologies are affecting our minds and our relationships in negative ways and point as evidence to things like cyber-bullying, internet 'addiction', the erosion of personal privacy, the supposed deterioration of our language (see Chapter 5), and the apparent shortening of our attention spans (see Chapter 6).

As we said in the first chapter, the introduction of new media has always made people uncomfortable, but their worries often turn out to be unfounded. At the same time, it is important not to dismiss concerns about the effects of technology out of hand. Many of them are based on legitimate concerns that are supported by scientific evidence. The trick is to separate the reasonable arguments from those based on emotion and exaggeration.

It is also important to remember that sceptics of digital technologies are not the only ones that trade in emotion and exaggeration. Those who uncritically extol the supposedly positive effects of digital media are equally guilty.

The writer Evgeny Morozov, author of *The Net Delusion* (2011), for example, talks about the way the Western media portrayed the 'dramatic' role of Twitter and other social networking sites (see Chapter 10) in the Iranian protests of June 2009. The problem with this portrayal, claims Morozov, is that it was probably wildly exaggerated. Most of the 'tweets' about the protests, he points out, came from outside of the country rather than inside of it and there is no strong evidence that digital social networking (as opposed to old-fashioned word of mouth) played a significant role in the organization of the protests. Finally, he reminds us that, whether or not digital media played a role, the revolution was *not* successful – the government of President Mahmoud Ahmadinejad not only retained power, but itself harnessed the power of the internet and social networking sites to round up critics and encourage citizens to inform on dissidents.

The main problem with both digital sceptics and digital utopians is that they often focus on the technology itself without paying enough attention to the social contexts in which it is used and the intentions of those who use it. This position is known as technological determinism, a belief that technology controls people's thoughts and behaviour. The opposite of this position is one which argues that technology is ideologically neutral and that people can do whatever they want with it. This position is equally problematic. The position we have been taking in this book is that the truth lies somewhere between these two extremes, and the first step to taking a critical stance towards digital media is to understand both the potential for technology to control us and our potential to exercise control over technology.

## MEDIATION

In the first chapter we introduced the concept of *mediation* as the foundation of our approach to digital literacies. All human actions, we pointed out, are *mediated* through tools, either technological tools, like telephones and computers, or symbolic tools like languages and other semiotic systems. The crux of the concept of *mediation* is that we cannot interact with the world without doing it through some kind of *medium*, and the media that we use play an important role in determining how we perceive the world and the actions we can take. And so part of *mediation* has to do with how we are to some degree 'controlled' by the tools that are available to us to take action.

There are at least four ways that media can exert control over us. The first is through what we have been calling *affordances* and *constraints*. Different tools make some actions more possible and other actions less possible. They reveal certain aspects of reality and conceal others. They amplify participation in certain kinds of social practices and social groups, and constrain participation in others. In the last five chapters we have given examples of the affordances and constraints of various tools from hypertext to digital video. In the chapters in this section we have sometimes emphasized the affordances over the constraints. This is a natural reaction to any new technology – people tend to be more focused on the new things it enables them to do rather than the limitations it might impose. Taking a critical stance towards media means being just as conscious of what they *do not* allow us to do as what they *do* allow us to do.

## CRITICAL LITERACY

The second way media exert control over us is through the social conventions that grow up around their use. The way particular tools get used is not just a matter of what we can do with them, but also of the ways people have used them in the past. All tools carry the history of their past use. After people have used a particular tool in certain ways to perform a particular practice for a period of time, the conventions or 'social rules' that have grown up around the tool and the practice become 'solid'. We call the process by which social practices and conventions come to 'solidify' around various technologies the **technologization of practice** (Jones 2002). The technologization of practice often leads to situations in which, directly or indirectly, the way certain things are done comes to be controlled by the dominant technologies that we use to do them.

The third way media exert control over us is through who has access to them. The distribution of tools, both technological and symbolic, in any society is always unequal. What this means is that the kinds of actions that media make possible are always only available to certain people. In other words, the use of mediational means is always tied up with economic and political systems that govern the way access to them is distributed. As a result, the ways media end up being used usually support or perpetuate these political and economic systems.

Finally, media exert control over us through how easy or difficult they are for us to use. All tools require that people learn how to use them. Some require a great deal of time and energy to learn how to use, while others are much simpler. Media that are extremely easy to use can exert control over us by encouraging us to regard them and the actions they make possible as 'natural'. Media that are more demanding exert control by requiring us to make a 'commitment' to them. The more time and energy you devote to mastering a particular tool, the more likely you are to regard the kinds of actions that that tool makes possible as valuable.

The longer we use some technologies, the harder it becomes for us to get along without them or to switch to competing technologies, either because we have organized so many of our activities around the affordances of these technologies or because so many of our friends use them that abandoning them might result in social isolation. The phenomenon by which technologies make it progressively difficult for us to separate ourselves from them is called **lock-in**. One good example of lock-in can be seen in online video game servers which ensure that owners of a particular system can only play with other owners of that system: PlayStation gamers, for instance, can only play with other PlayStation gamers and only Xbox gamers can play together. When you buy a PlayStation, you become 'locked-in' to playing with other people who have this platform, then after a while the costs of switching to an Xbox become too high, not just because of all the games you have paid for, but also because you have developed relationships with other people using your system.

Fortunately, the way media exert control over us is only half the story of mediation. Mediation also involves human agency. While media exert considerable control over what we can do with them, we also exert control over how we use media. There are at least four ways that we can exert control over the media that we use.

The first has to do with our ability to choose (or **appropriate**) the media we use. Even though we may feel we are locked in to particular media choices, there are often alternative choices available. We might choose a free open source software (see below) over a proprietary one, or an academic database instead of Google. The second way we can control

designers — that is, by **adapting** them to fit our own purposes. Facebook, for example, can be used to share personal information with our friends (and with Facebook's advertisers), or it can be adapted to do other things like mobilize people to protest Facebook's privacy settings and terms of use.

The third way we can control media is by actually changing or **modifying** ('modding') them in sometimes small but significant ways to make them more suitable for our purposes, a practice which, as we will see in Chapter 9, is common among computer gamers. Finally, the fourth way we can control media is by **mixing** two or more tools together such that the constraints of one tool are cancelled out by the affordances of another, opening up actions which neither of the tools alone could have made possible, as people do when they create remixes and mashups.

We will refer to these four ways we can control media — appropriating, adapting, modifying and mixing — as **hacking**. Often when people think of 'hackers' they think of cybercriminals who break into other people's computers and alter or steal data, but our use of the term harkens back to the 'hacker culture' of the early days of the internet (The Mentor, 1986). In this sense, the word 'hacker' refers to expert computer programmers, who get satisfaction from exploring the limits of programmable systems. For our purposes, what we mean by hacking is simply taking control of media and appropriating, adapting, modifying or mixing it in creative ways rather than simply going along with the agendas of those who manufactured or control the use or distribution of the media we are using.

The process of mediation, then, is not just a matter of media controlling people or people controlling media. It is a matter of the *tension* between what technology wants us to do and what we want to do with it, between the limitations it imposes on us and our ability to get around these limitations by 'hacking' it.

Media become *ideological* when they become resistant to hacking, that is, when the affordances and constraints they embody are presented as 'just the way things are' rather than as 'workable' and adaptable. They do this in two ways. One way is by becoming **transparent**, encouraging us to forget that they are standing between us and reality (Bolter and Grusin, 1999). Transparent media encourage us to regard the kinds of actions that they make possible as 'natural' or desirable and the kinds of actions that they constrain as unnatural or undesirable. Technologies tend to become more transparent to us the more we use them. A person writing with a pen, for example, is apt to forget that the pen is actually coming between him or her and the words that are appearing on the page. They think *they* are writing and forget the pen altogether. Similarly, a person chatting with a friend online may momentarily forget the mediational means and feel like he or she is communicating 'directly' with the other person.

There are, of course, a lot of apparent advantages to media transparency. The more transparent a mediational means is, the more comfortable and convenient it is for us to use. But at the same time, the more invisible its built-in limitations are to us, the more difficult it becomes for us to question those limitations. Transparent tools are resistant to hacking because they make us forget that they are even there. Many marketers of media technologies extol the value of media transparency. When Apple first released its iPad, for example, it advertised it with the assertion: 'Technology is at its best when it feels completely natural, almost like there's no technology at all'. From the point of view of critical literacy, too much media transparency is inherently dangerous because it makes us less conscious that what we are experiencing is mediated.

The second way technology becomes resistant to hacking is by becoming **opaque**. When something becomes opaque, we are unable to understand it fully; unable to figure out how it works. Despite the increased access to information made possible by digital media, very few people know much about how the technologies they use on a day-to-day basis actually work. In his book *Program or Be Programmed* (2010), the author Douglas Rushkoff points out the irony that fewer children nowadays learn computer programming than did thirty years ago. At the same time, many technologies themselves seem designed to prevent users from learning how to re-program them.

A system that does not allow anybody apart from the technology manufacturer or those with whom it has agreements to alter it in any way and usually requires that users purchase a particular device to run particular software is called a **closed system**. Apple Corporation's devices and software are examples of closed system technologies. Users cannot run their software on any devices other than those manufactured by Apple and they cannot re-program these devices. Microsoft's Windows and Google's Android are examples of more **open systems** which can be run on different devices, though these systems are still licensed by the manufacturers. An example of a truly open system technology is Linux, an **open source** operating system whose code is publically available to anyone who wants to change it.

The more closed a system is, the more opaque the technologies within it become. Technologies are opaque when we can't open them up and see what's inside and when the code for the software or the algorithm for the search engine is kept secret from us. Of course there are often good reasons for media opacity having to do with intellectual property, the ability of technology companies to earn money, and the protection of other people who are also using the technology. Sometimes people who hack media have less than honourable intentions (which is one reason why the whole idea of hacking has gotten a bad name). Systems that are completely open make it difficult to maintain quality and are more vulnerable to abusive practices. Systems that are totally closed, on the other hand, lock people in to certain ways of behaving and thinking and discourage innovation.

It is important to remember that what we mean by transparency is not that we can see 'inside' of something, but that we 'see through it' and forget that it's there, and what we mean by opacity is not that something becomes obtrusive or obvious, but that it becomes resistant to adaptation and modification.

And so, opacity and transparency, the way we are using the terms here, are *not* opposites. In fact, the ironic thing about digital media is that transparency and opacity often go hand in hand. The more transparent a technology becomes, the less likely it is for us to understand how it works. Transparency makes the effects of media less apparent to us and discourages us from critically examining them, and opacity makes it more difficult for us to critically examine them, even if we want to.

### Activity 7.1: Your experience with technology

#### A. TRANSPARENCY AND OPACITY

List the media that you use on a daily basis (including both hardware like mobile phones and software interfaces like Facebook) and rate them according to their transparency (how 'natural' they feel to use) (10 = very transparent, 1 = not at all transparent) and their opacity (how difficult it is to understand and alter their 'inner

workings') (10 = very opaque, 1 = not at all opaque). Talk about the reasons why you rated these media the way you did and discuss the degree to which the transparency and opacity of media affect how much control you have over what you can do with them.

#### B. LOOKING FOR LIMITATIONS

As we have said, whereas the affordances of technological tools are often obvious to us (usually because manufacturers of those tools spend a lot of money communicating those affordances to us), sometimes the ways these technologies limit our abilities to take certain actions or engage in certain social practices is not so obvious. Choose one of the media you considered above and list the kinds of social actions, social practices, social relationships and social identities that this tool makes it *more difficult* for you to participate in.

#### C. ARE YOU A HACKER?

Think of some instance where you have selectively appropriated, adapted, modified or mixed technologies in order to get around their limitations or do something new with them. Describe how you did this and how it altered your relationship with these technologies.

## IDEOLOGY AND THE TECHNOLOGY OF LANGUAGE

When it comes to developing critical literacy with digital technology, a good place to start is by practising it first with the analogue technology of language. Like digital media, the medium of language also tends towards transparency and opacity. We are rarely aware of language as a mediator of meaning. Instead we tend to think that we are communicating with people directly. Thus, certain ways of constructing reality with language come to be taken for granted by us, considered 'natural' or 'normal'. Furthermore, most of us are just as unaware of how language actually works to create meaning and promote ideology as we are of the workings of our personal computer or smartphone.

Anytime a person uses a semiotic system like language to make meaning they always have an *agenda*. We produce texts in order to get things done – whether that means achieving some kind of material gain, fulfilling an obligation to someone, or making someone do something or believe something. We judge how successful our texts are by how well they help us to realize our agendas. The first question to ask whenever we encounter a text is what the agenda of the person or people who produced this text is.

This is not always an easy question to answer because people are often not completely straightforward about their true intentions when they produce texts, and sometimes they even try to hide them. Sometimes it's even difficult to find out who produced a text in the first place. The text may come from an institution or group (like a company or a government) rather than an individual, or the source of the text may be a complete mystery to us. However, we can never really critically assess a text unless we have some idea about 1) who produced the text and 2) what their agenda might be.

The second important thing about language is that it is never 'objective'. That is to say, the 'reality' that texts depict is never 'true' reality – it is always a matter of selection, a particular 'version' of reality. We depict reality by using various words that represent *things* and *actions*. Language is basically a 'filter' (see Chapter 2) – a system of inclusion and exclusion: whenever we choose to include or emphasize one thing in our version of reality, we are also choosing to exclude or downplay something else. When we refer to the World Wide Web as 'an information superhighway', for example, we emphasize the role of the web in the exchange of information and commerce and deemphasize the role the web plays in social interaction or entertainment.

The versions of reality that we construct with our language are not just a matter of the words we choose, but also how we choose to connect up these words to form **propositions** – statements about the world and how it works. The function of language in forming propositions is called the **ideational function** of language because it allows us to express our ideas about reality. We might, for example, attempt to construct propositions by asserting various kinds of 'logical relationships' like those discussed in Chapter 3 (e.g. part/whole, cause/effect) between things and/or actions. The problem is that these 'logical connections' are not always entirely logical. For example, in an article entitled 'The Internet Helps Build Democracies' published in *Newsweek* on April 30, 2010, the authors wrote:

The revolts in Ukraine, Kyrgyzstan, Lebanon, Burma, Xinjiang, and Iran could never have happened without the web.

The main problem with this sentence is that, although it seems to establish a cause and effect relationship, this relationship is neither logical nor scientifically verifiable. It is impossible to either prove or disprove. Since the web *did* exist and *did* apparently play some role in these revolts, we have no way of knowing if they would have occurred without the web. Another problem with this statement is that it confounds *association* with *causation*: just because the people involved in these revolts were using the internet does not mean that their internet use caused the revolts or even that it contributed to the revolts in any significant way.

Just because somebody constructs a proposition doesn't mean we have to accept it as the truth. There are, however, a number of linguistic tricks that people use to make their propositions more 'natural'-seeming (transparent) and more difficult to question (opaque). One of the most powerful is called **presupposition**. Presupposition is when, rather than asserting something directly, writers or speakers treat certain propositions as 'given'. When Steve Jobs, for example, said of the iPad, 'It is the best browsing experience you've ever had,' he was making a couple of assumptions: first, that the 'browsing experience' you have had in the past is somehow 'defective' or 'inferior' compared to the one you will have with his product (an assumption he made without much information about your particular past experiences with browsing); and second, that what makes a browsing experience positive is the *technology* that you use rather than the *content* that you access. This last proposition, though debatable, is very difficult to question since it is hidden deeply inside the bigger proposition about your browsing experience.

Texts promote ideologies not just through how they represent the world, but also through how they represent relationships between people, what linguists call the **interpersonal function** of language. They do this in many different ways, one way being through *advertising* certain genres or styles of writing or speaking. Different kinds of texts create

different kinds of relationships between their producers and the people who read them. Some kinds of texts make it easy for the reader to disagree with or debate the writer, whereas others make it almost impossible. A legal notice, for example, puts the reader in a different kind of position *vis-à-vis* the writer and the text than an online advice column.

One of the points we made in Chapter 3 was that web 2.0 has introduced new ways for people to 'talk back' to writers by commenting on the text, adding to the text, or even changing the text. Different tools, however, differ dramatically in the amount of power they give to readers. Blogs, for example, might allow readers to comment on what they have read, but don't have the capacity of other tools like wikis (see Chapter 11) to allow readers to alter the actual text that they have read.

Furthermore, writers can create a relationship with readers through the kind of language they use. Texts written in an informal style often create a sense of intimacy, whereas texts written in a formal style create a sense of distance. In the example above, when Steve Jobs said that the iPad is 'the best browsing experience you've ever had', he created a sense of closeness with the reader by addressing him or her directly and by using informal language ('you've').

Finally, authors of texts create relationships with their readers by adopting certain kinds of identities, projecting certain kinds of identities on to the reader, and portraying what they are doing in a certain way. For example, you might be more receptive to an opinion about a new technology product if it appears in the context of a news story or a customer review than if it appears in an advertisement. This is why technology companies spend so much money on public relations campaigns designed to get journalists and customers to express positive opinions about their products in the context of news stories and reviews.

## Activity 7.2: Analyzing texts

Analyze the following two texts using the questions below as guidelines. (The URLs provided are for both the original locations of the texts and their locations in the internet archive: <http://waybackmachine.org>)

### Text 1

'iPad: A magical revolutionary product at an unbelievable price', by Apple Corp (<http://www.apple.com/ipad/>)  
(Archived at <http://web.archive.org/web/20101211014634/http://www.apple.com/ipad/>)

### Text 2

'Why I won't buy an iPad and think that you shouldn't either' by Cory Doctorow (<http://boingboing.net/2010/04/02/why-i-wont-buy-an-ipad-and-think-you-shouldnt-either.html>)  
(Archived at <http://web.archive.org/web/20101206055212/http://boingboing.net/2010/04/02/why-i-wont-buy-an-ipad-and-think-you-shouldnt-either.html>)

1. What is the source of the text? Who produced it? Is the author an individual or an institution? What is the 'agenda' of the people who produced the text? What

opportunities do you have to comment on the text or communicate with the producer?

2. How does the text construct a version of reality? What *things* and *actions* are included and excluded? How are these things and actions characterized?
3. What kinds of *propositions* does the text contain? Are these propositions logical or verifiable?
4. How does the text use presupposition to make meanings seem more 'natural'?
5. What kinds of social relationships are constructed by the text through things like genre and style? What identities do the authors construct for themselves and their readers?

## IDEOLOGY AND DIGITAL TECHNOLOGIES

Just as texts promote certain versions of reality and certain social relationships and identities as 'natural' or desirable, so do media, and they often do so using similar kinds of techniques as those we have pointed out in texts.

Like texts, media are never 'objective'. They always promote a particular version of reality through the choices they require users to make, choices which, in the case of digital media, are often expressed in terms of drop-down menus or 'preference' settings. While the word 'preference' gives the illusion that the media can be truly personalized, this is rarely the case. The choices we are presented with always instantiate systems of inclusion and exclusion, which limit the kinds of actions we can take and the kinds of meanings we can make. Perhaps the main difference between digital technologies and earlier analogue technologies lies in the nature of the choices they make available to us. While digital technologies seem to make more choices available to us, sometimes dizzying in number, these choices are always expressed as discrete alternatives. We are not, as we are with many analogue technologies, able to invent our own choices by creating fine gradations of meaning which we can control ourselves. Even when it seems we can move along a continuum of, say, brightness or volume, the underlying mechanism is always governed by a predefined set of alternatives.

The semiotician and jazz pianist Theo van Leeuwen (2012) pointed this out in a comparison of what it's like to play a digital piano as opposed to an analogue one. He writes:

On a modern digital piano I cannot do what Billie Holiday did with her voice, or what I can even do, to some extent, on an acoustic piano. I cannot make the sound rougher or tenser as I play. Instead the designers of the instrument have provided me with a choice from a wide range of piano sounds, some tenser, some mellower, some higher, some lower, some smoother, some rougher (in a synthetic kind of way) and so on. All I can do myself is phrasing.

Furthermore, these fixed systems of inclusion and exclusion can become 'locked-in' in ways that they affect the development of later technologies. Jaron Lanier (2010), author of *You Are Not a Gadget*, for example, relates the story of how the MIDI format, which was developed in the 1980s to digitally represent the sounds of a keyboard, has become the

foundational technology for all digital music, including the mp3 files most people listen to today. The problem with this, he argues, is that, while it is good at representing the sounds of a keyboard, it is less accurate when it comes to representing the sounds of other instruments like the violin or the human voice. Unfortunately, this format has become such an integral part of digital music systems that it is too difficult to change it. 'Software' writes Lanier (2010: 1), 'expresses ideas about everything from the nature of a musical note to the nature of personhood ... (it) is also subject to an exceptionally rigid process of "lock-in." Therefore, ideas (in the present era, when human affairs are increasingly software driven) have become more subject to lock-in than in previous eras.'

Not only do technologies impose upon reality systems of inclusion and exclusion based on the limited number of choices they make available, but, like texts, they sometimes present these choices as 'givens' (presuppositions) – that is, they make certain choices seem 'normal' and others seem unusual. Most software programs, for example, present users with a series of 'default settings', which inevitably influence their perceptions of normalcy and deviance. No matter how much the palette of choices is expanded, giving the impression of freedom, default settings always steer users towards a certain set of normative behaviours and normative meanings.

Most programs go even further, building in things like 'wizards', 'templates', and 'themes' which instruct users not just about design choices but also about the kinds of social contexts in which the technology is meant to be used and the kinds of social actions it is meant to be used to perform. Apple's iMovie, for example, offers themes like 'Photo Album', 'Comic Book', 'Scrapbook' and 'Film Strip', each with different kinds of title pages and transitions. Such prefabricated templates are not just there to make it easier to use the software but to create the contexts in which users are expected to make movies. iMovie asks its users to use it to 'have fun' and 'share memories' with friends and relatives. It doesn't invite them to make probing documentaries or film noir features.

Another way systems of inclusion and exclusion can limit our perception of the world is through the various ways technology 'filters' the data that is available to us. Google's Page Rank algorithm, for example, promotes the proposition that the value of a piece of data is best determined by how popular it is. Perhaps the most insidious ideological filters are those that shape our view of reality based on the view of reality that we already have. Examples of such filters are Google's personalized search (see Chapter 2) and the EdgeRank algorithm in Facebook, which pushes status updates and news about the friends you most frequently interact with to the top of your News Feed. The danger of such filters is that we end up essentially 'brainwashing ourselves' by limiting what we are exposed to based on our past preferences. In his book, *The Filter Bubble: What the Internet Is Hiding From You* (2011), Eli Pariser, describes the problems associated with personalized filters in the following way:

With Google personalized for everyone, the query 'stem cells' might produce diametrically opposed results for scientists who support stem cell research and activists who oppose it. 'Proof of climate change' might turn up different results for an environmental activist and an oil company executive. In polls a huge majority of us assume that search engines are unbiased. But that may be because they are increasingly biased to share our own views. (2)

Finally, like texts, technologies always create particular relationships between users and producers of the technologies and between users and other users. A good example of this

is Microsoft PowerPoint, which has become a fixture in many classrooms and lecture halls. While there are obvious advantages to using PowerPoint to present information, there are also biases built into the software in terms of the kinds of meanings and the kinds of social relationships that it promotes. Not only does the program encourage users to express ideas (and conceive of problems) in terms of bulleted hierarchical lists, excluding other ways of thinking, it also promotes a certain kind of power relationship between teachers and students. Although it is possible for teachers and students to have discussions in the context of a PowerPoint presentation, the way the technology is designed to be used discourages interaction. Furthermore, because the person manipulating the slideshow has access to more modes of communication than the person watching, it is more difficult for audience members to question or challenge the presenter. Embedded within these affordances and constraints is a certain *ideology* of education, one which conceives of education as a matter of presenting authoritative information in the form of discrete, 'digestible' bits, rather than a matter of encouraging inquiry and debate.

One problem with the way technologies promote certain kinds of social relationships is that the nature and purpose of these relationships are not always obvious and straightforward. Perhaps the best example of this is Facebook. The default relationship with Facebook 'friends' is to disclose as much information as possible to them. Although the interface allows these 'default settings' to be modified, the technology itself and the applications associated with it encourage greater and greater disclosure from users, which, given Facebook's business model of selling user information to advertisers in aggregated form, is good for the company.

While Facebook is in the business of encouraging users to make their relationships with other users as visible as possible, the relationship Facebook itself has with its users is rather ambiguous. On the one hand, they are providing a service, making their users their customers, but on the other hand, they are selling information about and access to their users to advertisers, making their users their *product*.

Learning how to use technological tools, then, involves not just mastering the range of choices they present, but also being indoctrinated into the social practices that have come to be *technologized* around these tools. The range of actions these tools make available not only determine how people behave and communicate with each other, but they also end up promoting particular versions of reality and making some kinds of social relationships more possible and others less possible.

### Case study 7: Protecting children online

A good example of how different technological tools can promote different ideologies can be seen in the design of massively multiplayer online games (MMOGs) for young children. MMOGs are digital games that you can play with a large number of other players on the internet.

Typically, the designers of such games as well as the parents of children who play them share the belief that children require protection in online settings from unwanted online interactions, such as those associated with cyber-bullies or online predators. In order to provide a safe environment, game designers provide tools that structure

interaction in particular ways, and thereby promote particular beliefs about how to keep children safe.

In order to illustrate these points, we will consider three browser-based, cartoon-style MMOGs designed for children: *Poptropica* (by Pearson Education), *Moshi Monsters* (by Mind Candy) and *Club Penguin* (by Disney). All of these games are aimed at young children and have information for parents prominently linked from their home pages. Online safety for players is explicitly addressed in all three games and is clearly perceived as an important value by the game designers. What is interesting is the different ways that designers promote this value through the affordances and constraints of the communication tools they make available in their games, and how these affordances and constraints create different 'versions' of what does or does not constitute 'safety'.

The first game, *Poptropica*, is an adventure game consisting of a number of islands which players can visit. Each island is a different virtual world with its own puzzle-like adventure for players to solve. Players can also visit 'common' or 'multiverse' rooms and play a range of mini-games against other players. In these common rooms synchronous interaction with other players is possible, using a 'pre-scripted chat' system that provides players with a limited number of questions and answers. For example: 'Are you good at soccer?'; 'I am ok at it'.

Because the chat is pre-scripted in this way, the range of topics that can be explored by players is limited. This constraint on chat means that 'unsafe' conversations are technically impossible. As the developers note in the FAQ: 'There is purposely no free chat in *Poptropica* in order to make our virtual world as safe as possible for kids' (<http://www.poptropica.com/Poptropica-FAQ.html>). The particular choice of design in *Poptropica* reflects the belief that an effective way to provide protection for children is to greatly restrict the number of possible actions that they can take and topics they can discuss.

The second game, *Moshi Monsters*, is a social game in which players adopt and look after a pet monster. They interact with the monster by feeding it, playing with it and customizing the room that it lives in. They can also play mini-games to earn points so that they can visit shops in the virtual world and buy items for their monsters. In this game, players can interact with other players by adding each other to their buddy list. Once they have buddies, they receive 'news' about their buddies' actions and achievements, in the form of a non-interactive list of automatically generated status updates. In addition, they can visit their buddies' rooms and leave a note on their buddies' pinboards, a form of asynchronous messaging. Finally, they can connect with other players by posting comments to a community forum on the site.

Unlike *Poptropica*, the communication tools in *Moshi Monsters* are not pre-scripted, allowing users to compose freely. However, as asynchronous tools, the pinboard and forum systems are also constrained – the proximity of such asynchronous tools is

lower than that of synchronous chat, and so it is possible that this reduced proximity discourages undesirable behaviour. Nevertheless, additional monitoring is necessary to promote a respectful online culture. This takes the form of a filtering system for pinboard posts and self-policing tools throughout the site, which allow players to flag content for a moderator to review (see <http://www.moshimonsters.com/parents>). The choice of communication tools here reflects a subtly different ideology about online safety for children, one which aims to promote 'safety' by monitoring interaction rather than restricting it.

The third game, *Club Penguin*, is another social game. Players are penguins in an icy virtual world where they can interact with other players and play a range of mini-games. Out of the three games, *Club Penguin* is the one that makes synchronous interaction with others the easiest. It is possible to interact through synchronous chat with any other player anywhere in the virtual world. Chat can be either pre-scripted or free, but the default selection is free synchronous chat and changing this setting requires a separate parent account.

The chief way safety is promoted in this more open environment is through a set of 'rules' that are posted on the homepage of the website: 1. Respect other penguins; 2. Never reveal your personal information; 3. No inappropriate talk; 4. No cheating. Players are encouraged to report other players who violate these rules. In addition, the site employs moderators who monitor activity and chat and investigate complaints about particular players. In this model, safety is seen as a matter of people agreeing to follow a common set of rules. Responsibility is given to the children themselves (and their parents) to both refrain from inappropriate behaviour and to monitor the behaviour of others and report any transgressions to the 'authorities'.

All three of the games described here are clearly committed to the value of online safety for children and take steps to promote community norms that are consistent with this value. However, the approaches taken reflect different ideologies about the level of control that is necessary in order to achieve the goal of safety. Compared to the pre-scripted chat tools of *Poptropica*, the communication tools adopted in the design of both *Moshi Monsters* and *Club Penguin* reflect the belief that children should have more freedom of expression when networking with their friends. They also reflect beliefs about the role of parents and of children themselves in helping to maintain a safe environment.

Although these are just children's games, what children are exposed to when they play them are actually potent lessons in politics and citizenship. Each game indoctrinates them into a different model of what it means to be a citizen, from the restrictive censorship of *Poptropica* to the 'self-regulation' and 'rule of law' promoted by *Club Penguin*. Although it is difficult to say how much the experiences of playing these games might affect their later civic lives and political choices, it is important to note that even something as seemingly innocent as an online game about penguins cannot avoid reflecting and reproducing ideology.

**Activity 7.3: The ideology of media**

Choose one of the digital tools listed below or some other tool that you are familiar with and analyse it for 'biases' using the questions provided for guidance.

- Facebook (or some other social networking site)
- Google (or some other search engine)
- Microsoft Word (or some other word processing program)
- iPhoto (or some other photo management program)

**A. IDEATIONAL FUNCTIONS**

What kind of 'version of reality' does the tool construct through the choices it makes available to users? What are the 'default settings' and how do they reflect the way users are expected to use the tool? How easy is it to change these default settings? What kinds of things are users *not* able to do with the tool?

**B. INTERPERSONAL FUNCTIONS**

What kind of relationships does the tool construct between the manufacturers and the users or between different users. Does the tool encourage users to treat other people in a certain kind of way? How clear and visible are these relationships? How easy are they to change?

**CONCLUSION**

If there is one lesson you are meant to take from this chapter it is that critical literacy means trying as much as possible to learn how things work, whether we are talking about the way language works to influence our opinions and view of the world or the way software works to influence our behaviour and our relationships with other people. We cannot avoid using various mediational means to take action in the world. However, the more we know about how these media work, the better we can become at 'hacking' them through selective appropriation, adaptation, modification and mixing to fit our own purposes and promote our own agendas rather than the agendas of politicians, journalists, engineers and corporations. Of course, this is not always easy as producers of texts, software programs, digital devices and other mediational means increasingly work to make their tools more and more transparent and more and more opaque, and therefore more and more resistant to hacking.

In this first section of the book, we have outlined some of the affordances and constraints digital media have introduced into our lives, and how these affordances and constraints affect how we can perform various social actions, how we can make meaning, how we can relate to others and construct our own identities, and even how we can think. To really understand the effects of digital media on our behaviour, our social relationships and our thoughts, and to understand our potential to creatively appropriate, adapt and mix these media to fit our own purposes, we must examine them in the context of actual *literacy practices* embedded in real life situations. These literacy practices will be the subject of the second part of this book.

## USEFUL RESOURCES

## Print

- Lanier, J. (2010). *You are not a gadget*. New York: Alfred A. Knopf.
- Morozov, E. (2011). *The net delusion: The dark side of internet freedom*. New York: Public Affairs.
- Gee, J. P. (2010). *An introduction to discourse analysis*, 3rd edition. London: Routledge.
- Pariser, E. (2011). *The filter bubble: What the internet is hiding from you*. New York: The Penguin Press.
- Rushkoff, D. (2010). *Program or be programmed: Ten commands for a digital age*. New York: Or Books.
- Vaidhyanathan S. (2011). *The Googleization of everything (And why we should worry)*. Berkeley: University of California Press.

## Web

- The Mentor, The hacker manifesto  
<http://www.mithral.com/~beberg/manifesto.html>
- List of logical fallacies  
[http://web.cn.edu/kwheeler/fallacies\\_list.html](http://web.cn.edu/kwheeler/fallacies_list.html)
- Jay Lemke, Towards critical multimedia literacy: Technology, research, politics  
<http://www.personal.umich.edu/~jaylemke/papers/reinking2/htm>

## Video

- Net neutrality  
<http://www.youtube.com/watch?v=19jH0n0EW8U>
- TED, Eli Pariser, Beware online 'filter bubbles'  
[http://www.ted.com/talks/eli\\_pariser\\_beware\\_online\\_filter\\_bubbles.html](http://www.ted.com/talks/eli_pariser_beware_online_filter_bubbles.html)
- Iaron Lanier, You are not a gadget  
<http://youtu.be/T5JZFxr1lY>



Information overload: The movie  
<http://www.youtube.com/watch?v=MuwJevVFJF20>

Do you suffer from information overload syndrome?  
<http://www.youtube.com/watch?v=zhoRkxypAwE&feature=related>

Google, How Google makes improvements to its search algorithm  
<http://youtube.com/watch?v=J5RZOU6wK4Q>

CHAPTER 3

# Hyperreading and Hyperwriting

*Hyper Literacy*

With the development of digital media, the activities of reading and writing, which used to be mediated by pen and paper, typewriters and printing presses, are now increasingly mediated by digital tools like web browsers that we access through our computers and other digital devices. This shift from print-based media to digital media has been accompanied by the development of new literacy practices, shaped by the affordances and constraints of digital tools.

Most printed books are designed to be read in a linear way that limits the reading path that we can take. This is not true of all books, of course. Encyclopaedias and dictionaries, for example, are designed to allow us to move from place to place, but we are still restricted by the classification systems of the authors. Printed books also separate us from the people who wrote them and put us in the position of passive recipients of information. In digital media, these design limitations are overcome through three main affordances that are not available in print-based media, namely:

- **Hypertext, Interactivity; and**
- **Multimedia.**

Let's take multimedia first. Multimedia, by which we mean content that is conveyed as images, audio, video, as well as text, allows us to represent information and make meaning using a wider range of modes (for example visual, aural) than is possible in print. With print technologies it was more difficult to include visual elements in a document and impossible to include aural elements and moving images. Even the relatively straightforward job of inserting a picture on a page posed technical challenges in the early days of print. In contrast, digital technologies have made these operations childishly easy, with the result that we now encounter multimedia everywhere. We consider these changes in detail in the next chapter on **multimodality**.

Another affordance available in digital media is hypertext. In essence, hypertext is electronic text which is **hyperlinked** to other electronic text. Hypertext is so fundamental to the architecture of the internet that we tend to take it for granted. However, it has had a profound effect on the way writers can structure and organize information, and the way readers can navigate their way through it. Unlike the pages of a book, which unfold in a linear sequence, hypertext can be organized in a variety of different ways.

Finally, digital media provide writers and their readers with the ability to interact with texts in ways that were previously difficult or impossible. For example, if you are reading a book and find an error, then there is no easy way for you to 'write back' to the author and