International Journal of Communication 6 (2012), 907–919 1932–8036/20120907

Information Capacity
IJoC

One in a Million: Information vs. Attention

MICHAEL LESK
Rutgers University

Aristotle could write that we ascribe "universal education to one who in his own individual person is thus critical in all or nearly all branches of knowledge, and not to one who has a like ability merely in some special subject." Today nobody can know about everything. The flow of information so far exceeds what anyone can observe, learn, or appreciate that we must look at methods of compression, summarization, and filtering. These methods must achieve reductions of a million to one to cope with what is now routine. Fortunately, technology is making this possible.

The amount of information flowing around the world today is much greater than the ability of people to pay attention to it. The Global Information Industry Center says U.S. consumers are getting 34 GB/day (Bohn & Short, this Special Section), which for a typical lifetime of perhaps 25,000 days would be around 1 petabyte total. Some years ago, Tom Landauer estimated the size of human memory to be somewhere in the neighborhood of a gigabit, and less than a gigabyte (Landauer, 1986). That means that only 1 byte in 1 million of what you receive can possibly be remembered. So even the present, let alone the future, is not about the information we have; it's about our ability to search for, condense, and extract the tiny piece of that information that we might ever use (see also Neuman, Park, & Panek, this Special Section).

Looking ahead, what will happen to all the information people never see? Some of it will go into automatic systems. For example, weather forecasting with computer models can use many more observations than any person doing a forecast by hand could ever have studied. It's somewhat like transportation: In the United States every year there are about 4.7 trillion passenger-miles of people moving around (U.S. Department of the Interior, 2000), which at an average weight of 185 pounds is 423 billion ton-miles. In terms of freight, that is less than the Union Pacific Railroad moves all by itself. But the end purpose of all information and all transportation has to be to serve people; the railroads do not haul coal around the country for the benefit of the coal. Similarly, we send some information to machines instead of people, but often they then deliver something—like a weather forecast—to people. As the amount of information grows, our efficiency in processing it and summarizing it has to grow as well. It is

Michael Lesk: lesk@acm.org
Date submitted: 2011–07–11

no longer about just piling up disk drives; it is about utility, whether directly for people or indirectly through systems. We should ask not "how much," but "why?"

By contrast, when I started the "how much information" series, it was mostly done by looking at how much information devices could store. There were a number of obvious problems with such metrics, including duplication, detail, and destination.

Duplication, Detail, Destination

Duplication is a problem if we wish to understand how much unique information there is in the world, rather than how many copies there are of some of it. Most film snapshots and handwritten personal letters were unique, but the largest contributor to static information for a long time has been digital storage, making information easy to duplicate. Computer disks contain many redundant copies of things like operating system binary code.

Detail is the question of how precise any piece of information needs to be. If a photographer upgrades from a 1 megapixel image to a 10 megapixel camera, the images are not necessarily worth much more. Yes, there will be things visible in the higher-resolution image that cannot be distinguished in the lower-resolution one, but for most purposes, the extra detail is not going to be vital. Similarly, the change from NTSC video to HD video adds a lot of bits to a movie, but not much useful information: "A kiss is still a kiss."

Destination is the key question: Does anybody look at the information, or is it only going into some digital dumpster? For example, as Bohn and Short (this Special Section) point out, CCTV cameras now scan many stores and streets, but only rarely does anybody make use of any of the images. The largest source of digital information today is the output from digital sensors, spewing enormous quantities of data that must go immediately to data reduction systems. The next astronomy instrument, the "large synoptic survey telescope," will be generating several petabytes per day (LSST, 2001). We have gone, in information collection, from "just in time" (measure something you need to know) over to "just in case" (collect anything that might be wanted in the future). As mentioned above, in the long run, larger and larger fractions of our information will be going at least temporarily through automatic systems.

Given these and other problems with estimates of static storage, the Bohn study of consumer information looked at how much time people spent with what information source. As a result, they identified television and video games as the primary information providers. These are, however, forms of information that are highly duplicated; *Madden NFL 11* sold more than 2.5 million copies (Bilbao, 2011). Are all those users getting different information? As time goes on, such blockbusters are becoming less important. Chris Anderson, in his work on the "Long Tail" theory (Anderson, 2006), pointed out that we have many more books and TV shows since our distribution mechanisms are now able to successfully cater to individual tastes. So the amount of unique information flowing to people grows not only because we have more bandwidth, but because we can route a greater variety of information along our channels. What happens when that information arrives?

If we look at how much information somebody can absorb, we get widely variant numbers depending on the medium. Text, for example, can be read at perhaps 300 words/minute, at which rate it would take 11,000 hours to read a gigabyte. Since a lifespan is about 250,000 hours awake and active, a person who read and did nothing else might get through 20 GB, or about 20,000 books. But that is far-fetched: The home of a traditional scholar might have something like 5,000 books.

Writing is slower. Even typing is slower: A skilled typist would not do more than 80 words per minute, which full-time would be about 50 MB per year. If you have to generate the words to be written or typed, that takes longer. To my surprise, the prolific English novelist Anthony Trollope actually wrote 11,817,517 words (Newlin, 2005), which would be about 50 MB for a lifetime. But by modern standards of disk size, that is still chickenfeed.

However, the information in the texts is at least close to minimal. Compressing ordinary ASCII text is typically going to yield a factor of 4. By contrast, the information in pictures and sounds has a great deal more redundancy. As a comparison, if one reads text aloud, speaking 120 words per minute, that's about 10 characters per second of text; those 10 bytes in 128 kpbs MP3 format would correspond to 16 KB. So for the same content, the sound is more than 1,000 times bulkier. Admittedly, there is affect as well as words; some of the words are pronounced louder or slower, and that nuance has to be imagined in the printed form. But that doesn't sound like 1,000 times more useful information.

Listening to sounds, perhaps MP3 at 128 kbps, would be 17 hours per GB. But does one really need that much bandwidth? Voice communication can be done at 2.4 kbps, or about 1/50th the bit rate. The finer points of music are going to be lost at that rate, but the composition would be recognizable. Storing audio signals up to 96 kHz or higher, as recommended by the Audio Engineering Society, is of little use to humans (Link, 1999). The market for "videos for cats" (who can hear 60 kHz) is still pretty small. As Flanders and Swann sang some decades ago, "The ear can't hear as high as that. Still I ought to please any passing bat" (Flanders & Swann, 1960).

An extremely prolific classical composer—think Vivaldi—might have written 1,000 works. These are likely to add up to less than 20 GB even of MP3. As scores, they would be even more compact. The International Music Score Library Project (IMSLP, 2011) has about 100,000 scores, using 18 GB to store all the scanned pages; if they were encoded into MIDI, they would be smaller yet. For example, to look at Bach's Brandenburg Concerto No. 5, an example of a MIDI version is 90 KB, a set of scanned pages is 500 KB, and an MP3 recording is 38 MB, or some 75 times as large as MIDI.

Images, similarly, may have millions or billions of pixels, but what is significant, and what we absorb, is much less. Landauer estimated a typical image as using less than 14 bits of memory. This makes more sense if you think of the amount of processing people do when they see something. Years ago, I recall somebody at Bell Labs asking an information theorist, as part of planning for video telephony in the early 1970s, how many bits it took to represent a human face and mood. The theorist thought for a moment and replied, "Sixty bits." The engineer, flabbergasted, asked how that could be. The theorist said, "Well, there are fewer than 10 billion people in the world—that's 30 bits—and they have fewer than 10 billion recognizable facial expressions, that's another 30 bits." This sounded silly only because Leon

Harmon had not yet become famous for his pixelated image of Abraham Lincoln, using only a grid of 14x19 squares with 4 bits of greyscale on each square (Harmon & Julesz, 1973).

We can recognize these images, of course, because we know what they are. We don't see the pixelated Lincoln as 1,000 bits; we just think "Lincoln." The total data rate that might be perceived by our eyes is perhaps 10^{13} bits/second (Furness, 2003), but nobody can process this much. Instead, you identify interesting things in your field of view, and pay attention to them. You read words, for example, and convert them into thoughts. Falcons have three times your visual acuity, and pay no attention to printed words; but they are better than you are at recognizing mice and voles.

Is Less More?

To think of artworks, compare Piet Mondrian with Jackson Pollock. Below are thumbnails of works by Mondrian and Pollock. Both are about the same physical size: Mondrian, 32 x 20 inches; Pollock, 23 x 31 inches. When I run compression algorithms on the images, however, the Mondrian compresses to half the size of the Pollock. Does that make it less important or valuable? Certainly not in dollar terms. This particular Pollock sold at Sotheby's in 2010 for $8.7 million, while the Mondrian sold at Christie's in 2009 for $27 million. (See Figures 1 and 2.)

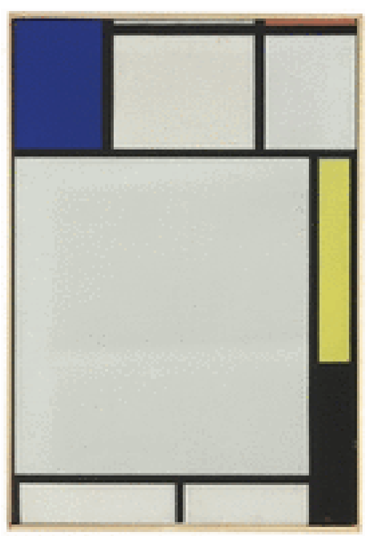

Figure 1. Piet Mondrian (1922). Composition avec bleu, rouge, jaune et noir.

Figure 2. Jackson Pollock (1948). Number 12a, 1948, Yellow, Grey, Black.

So our goal ought not to be simply increasing the number of bits flying around, but the value of those bits. Extra resolution in images, after a point, has minimal value. Thinking of Mies van der Rohe's dictum, "Less is more," reminds us that a simpler and more rapidly analyzed image may communicate more information with more impact.

How much do we reduce information when we take it in? That is not clear, and it depends on your familiarity with the subject. As an example, chess masters are much better than novices at memorizing a position from a chess game, when shown the board briefly. But they are no better if the pieces are not in a game position, but simply scattered at random on a chessboard. Some time ago, Marvin Minsky tried to estimate how many high-level "chunks" of information an expert maintained. He proposed, based on experiments with chess masters, that an expert maintained about 50,000 units of information about a subject. Minsky suggested that a mental activity where about 50,000 chunks are needed for mastery is challenging but reasonable (such as chess). An activity where many fewer chunks will exhaust the tasks (such as tic-tac-toe) is either a children's game or boring. An activity where many, many more chunks would be needed (perhaps chess on a 256 x 256 board) would simply be frustrating.

In terms of useful information, we have gigabytes of raw perception turned into megabytes of recognizable things, turned into kilobytes of something analyzed and useful. As the traditional hierarchy goes, *information* becomes *wisdom* which becomes *knowledge* (Eliot, 1934, n.p.). Just increasing the number of bits flying around does not help people; at best you will select what you need; at worst you will get confused. Accidents and mistakes are often attributed to information overload, including the Three Mile Island nuclear accident, some medical mistakes, and airplane crashes. What we can learn is more important than what we can see. There is little point in improving the resolution of book scanning beyond what is needed to capture the text, or in raising the quality of sound conversion beyond what can be heard.

Even for information flowing to computers, there will be a point of overload. As we add sensors to a weather-prediction system, we get better models and thus better predictions. We also, of course, increase the amount of data, which either slows the computation (meaning that we cannot generate predictions rapidly enough) or makes it more expensive as we buy faster computers. So far, increases in computer speeds have meant that we can afford to buy bigger machines. Since the sensors and networks are also getting cheaper, there may come a time when we have to choose between better predictions and predictions available more rapidly or more cheaply. Given the insensitivity of most people to slight differences in predicted weather (other than the snow/rain boundary), we might decide at that time to choose less information even when machines are the recipients.

When people are the recipients, we face an enormous mismatch between our ability to generate information and human capabilities. Donald Lindberg, director of the National Library of Medicine, said, "If I read and memorized two medical journal articles every night, by the end of the year I'd be 400 years behind" (deBronkart, 2010, entry for June 26, para. 6). A common metaphor is "Trying to interpret the torrent will be like drinking from a fire hose" (McCellan, 1998, p. 39).

However, that is not new. We have always lived in a world of stimuli, only some of which we think about. For example, the amount of visual information in the world around us is also huge, as mentioned earlier. The acuity of human vision is under 1 arc minute, and Clark (2009) estimates that the eye is equivalent to a 324-megapixel camera; see also Furness (2003). The dynamic range of each pixel is 10,000 to 1 (that's 16 bits), and we also have color vision, at lower resolution, so let's only double the grey-scale value, and the eye can see flickering images at perhaps 20 per second (this is why silent movies at 16 fps are inadequate but modern movies at 24 fps are OK). The total bitrate into the eye would thus be over 25 GB/second. How much does the eye reduce this as it sends signals to the brain? Koch et al) (2009) suggest that about 10 million bits per second, or about 1 MB/s, move from the retina to the brain. That is a reduction of 25,000.

Of course, other animals must do the same thing. Hawks and other raptors have several times our visual acuity and, if you believe they have less brain capacity than humans, their eyes must be doing an even better job of simplifying the information flowing to the brain.

How does sound compare? Human hearing has a spatial resolution of about 1 degree horizontally and 4 degrees vertically (Grantham, 1995), which would divide space (assuming facing one direction) into 8,000 "sound pixels," with each one involving an amplitude range of over 100 dB and with a frequency range up to 20 kHz. The audio industry, long ago, decided that sound is adequately transmitted at 1.5 MB/s (CD sound). Again, this must be greatly reduced as it reaches your brain, but the task starts out with much less information. Again, some animals do better. The domestic cat can hear sounds about 1/10th as faint and three times as high in frequency; it can also localize sounds more accurately, to an acuity of 1/6 of a degree, suggesting 100 times as much audio information (Tollin, Populin, Moore, Ruhland, & Yin, 2005). The hearing abilities of barn owls may be even better.

So we are accustomed, from evolution, to the idea that there may be lots of information, only some of which needs to be retained. In fact, we are presumably adapted to looking only at an appropriate amount of information. I once mentioned to a neuroscientist a suggestion made by the late Jim Gray, namely, that the designers of digital prosthetics should not just aim to give those with vision problems average eyesight, but should try to give all us the sight of eagles. His reaction was that if your eye tried to send your brain the amount of information that an eagle's eye does, you would probably go crazy.

What is new is that we now can save vastly more information, and try to sort through it automatically. Our goal needs to be not just to save information or to route it around, but to select what matters and present some version of that. This is a very practical problem. On a small scale, for example, a typical medical record is now more than 200 pages. Obviously, no doctor can read through every page for every patient. How can we intelligently present a doctor with the relevant and timely information for a particular patient? This only gets worse as electronic medical records, including imaging and sensor data, increase. Beth Israel Deaconess Medical Center at Harvard creates 20 terabytes of data per year (95% of that is image data), averaging 80 MB per patient (Halamka, 2011). This will only grow over time, with the increasing resolution of imaging devices, the advent of 3-D imaging on a large scale, and additional sensors in and out of the hospital.

Filtering, Summarization, Visualization and Parellelism

How can we reduce the amount of data presented to people, from the huge amount that we can store? There are several basic approaches, including selection, summarization, and visualization; usually more than one is used at a time. The most recent addition to this portfolio is collaboration, such as crowdsourcing, where multiple people are employed to look at a lot of data in parallel. Given the amount of data, however, multiple approaches are needed:

- *Filtering* or *selection* to pick out what is worth reporting.
- *Summarization* to give a sample value, or an average value.
- *Visualization* to display a sequence or table.
- *Parallelism*, whether human or machine, to process multiple data streams.

As an extreme example of automated processing, consider the search for the Higgs boson—"the God particle." The sensors in the accelerator are sending out 1,000 GB/sec, or a terabyte per second. That would be 30 exabytes per year if it were saved. Instead, the instrumentation immediately selects events, which are delivered at a rate of about 75 GB/second, and then those are filtered to 5 GB/second. That is then reduced to about 0.1 GB/second, or about 2 PB/yr (Djorgovski, 2007). In the end, the physicists are saving 1/15,000 of the original data, suggesting that the entire data-processing facility at CERN is comparable to your eye at simplifying and extracting useful information, but not any better.

What else can we do? One approach may be the generation of textual descriptions of data. Two early examples are stock market reports and weather forecasts. Automatic synthesis method for weather forecasts was developed by Richard Kittredge (Kittredge, Polguère, & Goldberg, 1986) when he generated linguistic versions of Canadian weather forecasts automatically in his RAREAS system. His system could create forecasts in both English and French. Similarly, Kukich (1983) demonstrated how to take the stock tables and turn them into a paragraph. More recently, Northwestern University has publicized a system that generates sports reports from text (Carr, 2009). All of these systems take tabular data and generate a natural language text that describes something about the data, typically following known patterns of journalism. This is more than form-letter generation, since the software must decide what is interesting in the data, and then create sentences that express those thoughts. Thus the process involves both filtering and language synthesis.

For example, Kukich's methods involved a language model based on left-to-right synthesis. The program would decide to say something about the overall market averages, and emit a beginning such as "The Dow Jones closed 50 points higher today." It would then look for something to say that could be added, such as "on heavy volume," and then still try to continue, perhaps with "led by General Electric rising to a new yearly high." Eventually it would find itself unable to lengthen the sentence in any useful way. This process could not create a sentence with an introductory adverbial or prepositional phrase; nobody ever noticed.

What was the compression factor of the stock market reports? Combining NYSE, Euronext, and NASDAQ, there are more than 10,000 traded securities, and the NYSE reports some 3 billion trades per day. If we assume each is perhaps 10 bytes (price, time, quantity), that would be 30 GB per day. An Associated Press daily news story on the markets is about 3,000 bytes, or a "compression" of 10 million to one. An alternative, of course, is visualization: standard plots of price, volume, and so on during a time period. And indeed, the daily chart that goes with a market article is also about 3,000 bytes (compressed).

Visualization has received more attention than language synthesis, although the rise of mobile devices may encourage researchers to build applications for motorists, who may be able to listen but not see. Visualization, however, can be very effective at replacing even extremely large data sets with easily grasped pictures. Perhaps the best known popular exponent of visualization techniques is Edward Tufte, whose three books (*The Visual Display of Quantitative Information, Envisioning Information,* and *Visual Explanations*) are classics on the design of display. Some spectacular examples of data visualization can be found at www.improving-visualisation.org and other sites.

Consider, for example, the "helioviewer" site–www.helioviewer.org–which delivers images of the sun. When it started in the 1990s, astronomers were getting about 250 MB/day of solar images from the SOHO (ESA/NASA Solar and Heliospheric Observatory) satellite. Now, with the SDO (NASA Solar Dynamics Observatory) satellite, there are 1.5 TB/day to be examined (Mueller et al., 2010). The helioviewer interface makes it easy for any Web user to look at individual pictures, sequences of images, or comparisons across time or regions of the sun. Again, petabytes of data are selected, compressed, and delivered as simple images that represent at most a few dozen kilobytes, perhaps a million-to-one selection. There are multiple other applications aimed at visualizing astronomy on a large scale (Hassan & Fluke, 2011). Similarly, Google Earth stored about 70 TB when it was new (Google, 2006) and is widely rumored to be over a petabyte now. Again, this is all boiled down to a single screen image.

Both the sun and earth viewers, however, are storing and displaying imagery. In terms of compression and visualization, there are also ways of displaying data that did not start with spatial imagery. For example, de Oliveira and Levkowitz's (2003) survey techniques, including iconographic, geometric, pixel-based, and hierarchical. The choice may depend on the data, the user, the technology available, or the ultimate purpose of the work.

Shneiderman (2003) pointed out that computer visualizations can provide abilities that are beyond those possible in the real world, such as showing objects hidden behind other objects or enabling an instantaneous shift to a different physical viewpoint. Or, as in the Furnas (1986) fisheye viewer, by changing scale within a diagram, portions of an image can be shown in greater detail while other portions can be shown in less detail.

Parallelism, the idea of processing lots of data in multiple streams, is, of course, an effective way of reducing the flow of data into any one receiver. Although the idea used to mean multiple computer processors running simultaneously, recently the idea of processing lots of data by using many people at once has become popular under that *crowdsourcing* name. Perhaps the prototypical project is "Galaxy Zoo," which uses amateurs to classify galaxies (Raddick et al., 2010). As of 2009, some 200,000 people had manually classified 100 million galaxies; no small group of astronomers or assistants could possibly have done this. The project starts with images and identifies each one with a few bytes of tag. Again, extremely great compression is achieved, albeit in this case at the expense of a large number of human hours.

Similarly, many of the "collaborative filtering" methods use the judgments of many individuals to select books or movies that someone else may wish to pay attention to. There may be several hundred thousand books published each year, but Amazon has an effective system of collecting user judgments and actions to suggest books for any particular visitor to its website.

Conclusion, More or Less

In summary, we have a variety of methods for large-scale data selection, compression, and presentation. We are better off than our predecessors, who also had libraries no single researcher could read, but had no search engines to help them find the particular items they wanted. They also had large amounts of data; for example, the herbarium at the Royal Botanic Gardens in London dates to 1853 and contains 7 million specimens. They had little in the way of tools to help them with that data; again, we are much better off today.

Research continues on building the technology that will help us filter and select information. Computer scientists have worked for decades on visualization, summarization, and filtering. More recently, collaborative filtering has done amazingly well at suggesting things you, personally, would like to read or see. The PageRank algorithm, which propelled Google into the leading place among search engines, is an extremely effective way to use the hyperlink structure of the Web to provide relevant rather than random choices among the documents that contain your search terms.

Moving still further forward, researchers have imagined systems that would maintain a user model and place new information in that model, telling you what you need or want to know, given your history (Brusilovsky & Maybury, 2002). In the long run, we care little about the quantity of information reaching a person, but rather about its quality and impact. After the message has been received, what is it that a recipient now knows or has enjoyed?

Reducing all incoming information to what people retain from it is oversimplified. I would not claim that *Huckleberry Finn* and *Don Quixote* are the same just because both can be abbreviated to "two guys go on a trip." But bits pouring into your brain are heavily filtered and indexed. There is no point in being overwhelmed with information that you will ignore. To live, I need a small amount of water, measured in quarts, every day. Providing me with an acre-foot a day would not help; it would literally drown me. What will matter in the future is not how much information is being pumped into our houses, but how much of it we can use.

References

Anderson, C. (2006). *The long tail: Why the future of business is selling less of more.* New York: Hyperion.

Aristotle. (350 BCE). *Poetics.* Book I, Part I. (W. Ogle, Trans.). Retrieved from http://classics.mit.edu/

Bilbao, R. (2011, July 15). EA Tiburon brings gaming success to Orlando. *Orlando Business Journal*.

Brusilovsky, P., & Maybury, M., (2002). From adaptive hypermedia to the adaptive web. *Communications of the ACM*, *45*(5), 30–33.

Carr, D. (2009, October 29). The robots are coming! Oh, they're here. *The New York Times.*

Clark, R. N. (2009), *Notes on the resolution and other details of the human eye*. Retrieved from http://clarkvision.com/imagedetail/eye-resolution.html

deBronkart, D. (2010). E-patient Dave blog, June 26 entry. Retrieved from http://patientdave.blogspot.com/2010_06_01_archive.html

Djorgovski, S. G. (2007, July). The big picture: Information technology revolution, and science in the 21st century. Retrieved from http://astro.caltech.edu/~george/bravo/Djorgovski-BRAVO-Lec4.pdf

Eliot, T. S. (1934). *The Rock*. London: Faber & Faber.

Flanders, M., & Swann, D. (1960). A song of reproduction. In their album *At the drop of a hat*. London: Parlophone.

Furnas, G. W. (1986, April). Generalized fisheye views. *SIGCHI Bulletin, 17*(4), 16–23.

Furness, T. (2003). Visual pixel flow rate. Retrieved from http://www.hitl.washington.edu/people/tfurness/courses/inde543/LECTUR E_HANDOUTS-03/FURNESS/visualpixelflowrate.pdf

Google. (2006). Retrieved from http://googlesystem.blogspot.com/2006/09/how-much-data-does-google-store.html

Grantham, D. W. (1995). *Spatial hearing and related phenomena*. New York: Academic Press.

Halamka, J. D. (2011, April 6). Storage costs of electronic health records surprisingly low. *MedCity News*. Retrieved from http://www.medcitynews.com/2011/04/storage-costs-of-electronic-health-records-surprisingly-low

Harmon, L. D., & Julesz, B. (1973, June). Masking in visual recognition: Effects of two-dimensional filtered noise. *Science 15, 180*(4091), 1194–1197.

Hassan, A., & Fluke, C. (2011). Scientific visualization in astronomy: Towards the petascale astronomy era. *Publications of the Astronomical Society of Australia, 28*, 150–170.

IMSLP Petrucci Music Library. (2011). Retrieved from http://imslp.org/wiki and, in particular, http://imslp.org/wiki/IMSLP:Site_support for the size of the library.

Kittredge, R., Polguère, A., & Goldberg, E. (1986). Synthesizing weather forecasts from formatted data. *Proceedings of the 11th conference on Computational Linguistics* (COLING '86). Association for Computational Linguistics (pp. 563–565).

Koch, K., McLean, J., Segev, R., Freed, M. A., Berry, M. J., Balasubramanian, V. et al. (2006, July 25). How much the eye tells the brain. C*urrent Biology, 16*, 1428–1434.

Kukich, K. (1983). Knowledge-based report generation: A technique for automatically generating natural language reports from databases. *Proceedings of the 6th annual international ACM SIGIR conference on Research and Development in Information Retrieval* (SIGIR '83) (pp. 246–250).

Landauer, T. K. (1986). How much do people remember? Some estimates of the quantity of learned information in long-term memory. *Cognitive Science, 10*, 477–493.

Link, M. (1999, May). *Digital audio at 96 kHz sampling frequency: Pros and cons of a new audio technique*. (Paper 4985). Audio Engineering Society.

LSST. (2001). Retrieved from http://www.lsst.org/lsst/public/tour_software

McLellan, F. (1998). 'Like hunger, like thirst': Patients, journals, and the Internet. *Lancet*, Supp. 2, 39–43.

Mueller, D., Hughitt, V., Langenberg, M., Ireland, J., Pagel, S., Schmidt, L. et al. (2010). The Helioviewer Project: Browsing, visualizing and accessing petabytes of solar data. American Astronomical Society, AAS Meeting #216, #402.23; *Bulletin of the American Astronomical Society*, *41*, 876. Also: No Author. (2010). *ESA offers a new way of looking at the sun*. Retrieved from http://sci.esa.int/science-e/www/object/index.cfm?fobjectid=48123

Neuman, W. R., Park, Y. J., & EPanek, J. E. (2012). Tracking the Flow of Information into the Home: An Empirical Assessment of the Digital Revolution in the U.S. from 1960–2005. *International Journal of Communication, 6,* (This Special Section).

Newlin, G. (2005). *Everything and everyone in Trollope*. Armonk, NY: M. E. Sharpe.

de Oliveira, M. C. F., & Levkowitz, H. (2003, July–September). From visual data exploration to visual data mining: A survey. *IEEE Transactions on Visualization and Computer Graphics. 9*(3), 378–394.

Raddick, M. J., Szalay, A. S., Vandenberg, J., Bracey, G., Gay, P. L., Lintott, C. et al. (2010). Galaxy Zoo: Exploring the motivations of citizen science volunteers. *Astronomy Education Review*, *9*(1), item 010103.

Shneiderman, B. (2003, November/December). Why not make interfaces better than 3D reality? *IEEE Computer Graphics and Applications,* 12–15.

Tollin, D. J., Populin, L. C., Moore, J. M., Ruhland, J. L., & Yin, T. C. T. (2005, March). Sound-localization performance in the cat: The effect of restraining the head. *Journal of Neurophysiology, 93*(3), 1223–1234. Also: The sound-localization ability of cats. *Journal of Neurophysiology*, *94*(5), 3653–3655.

Tufte, E. R. (1997). *Visual explanations.* Cheshire, CT: Graphics Press.

Tufte, E. R. (2001). *Envisioning information.* Cheshire, CT: Graphics Press.

Tufte, E. R. (2001). *The visual display of quantitative information.* Cheshire, CT: Graphics Press.

U.S. Department of the Interior. (2000). Transportation chapter. In *The national atlas of the United States*. Retrieved from http://nationalatlas.gov/transportation.html