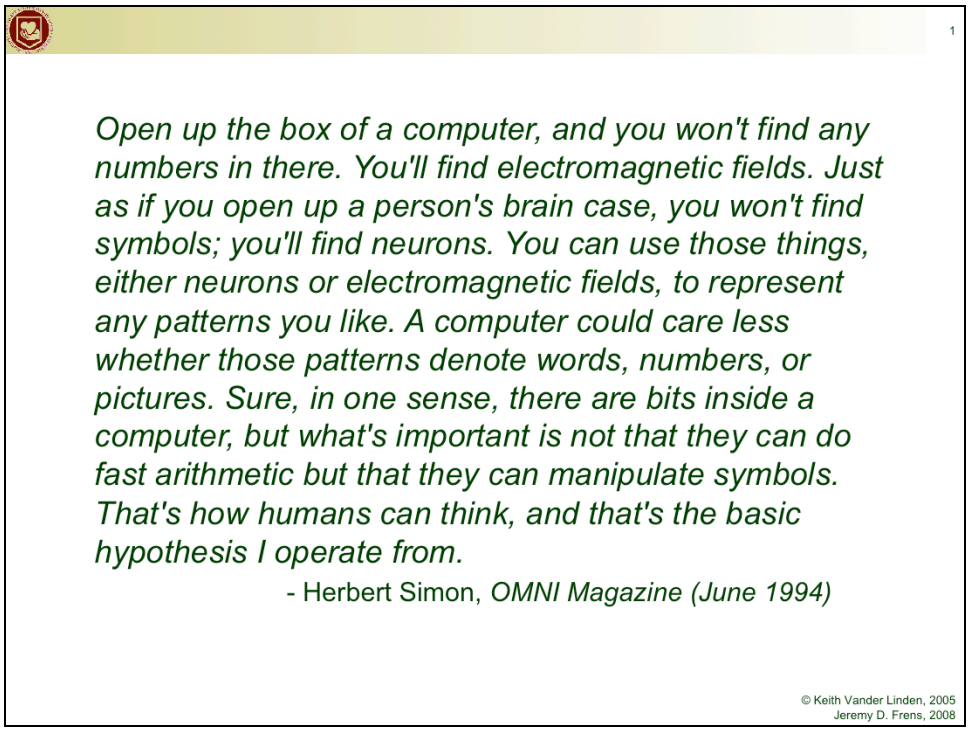
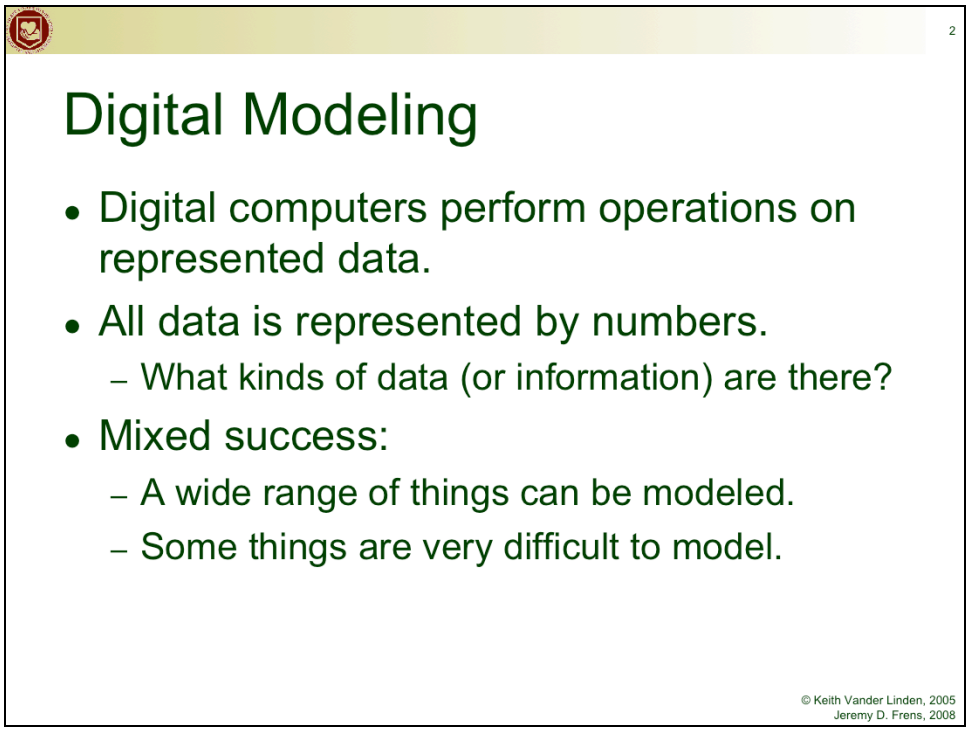


“omni”: of all things



kinds of data that need to modeled: numbers, letters, colors, pictures, sounds, videos, etc.

more to model: chemical reactions, disease spread, airflows, population growth/decline, etc etc etc . . .

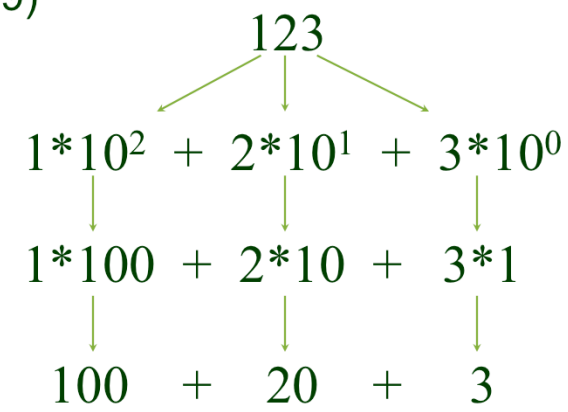
Binary

- Binary is a *base-2* numbering system.
- A *bit* is a "*b*inary dig*it*":
  - 0 (or "off")
  - 1 (or "on")
- Binary is just as powerful as decimal -- no more or less.

Decimal	Binary
0	0
1	1
2	10
3	11
4	100
5	101
6	110
7	111
8	1000
9	1001
10	1010
11	1011
12	1100
13	1101
14	1110
15	1111

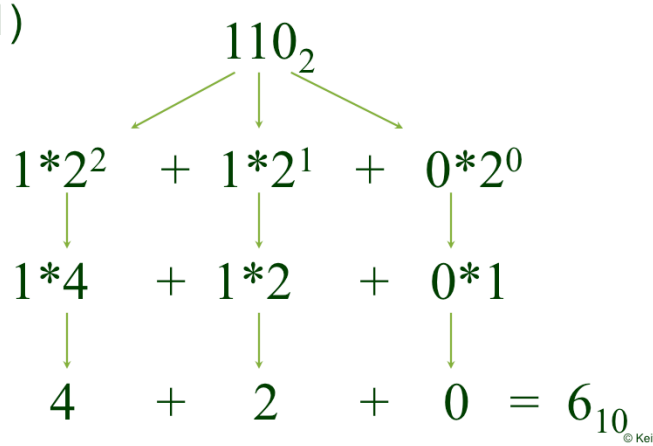
Decimal Numbers

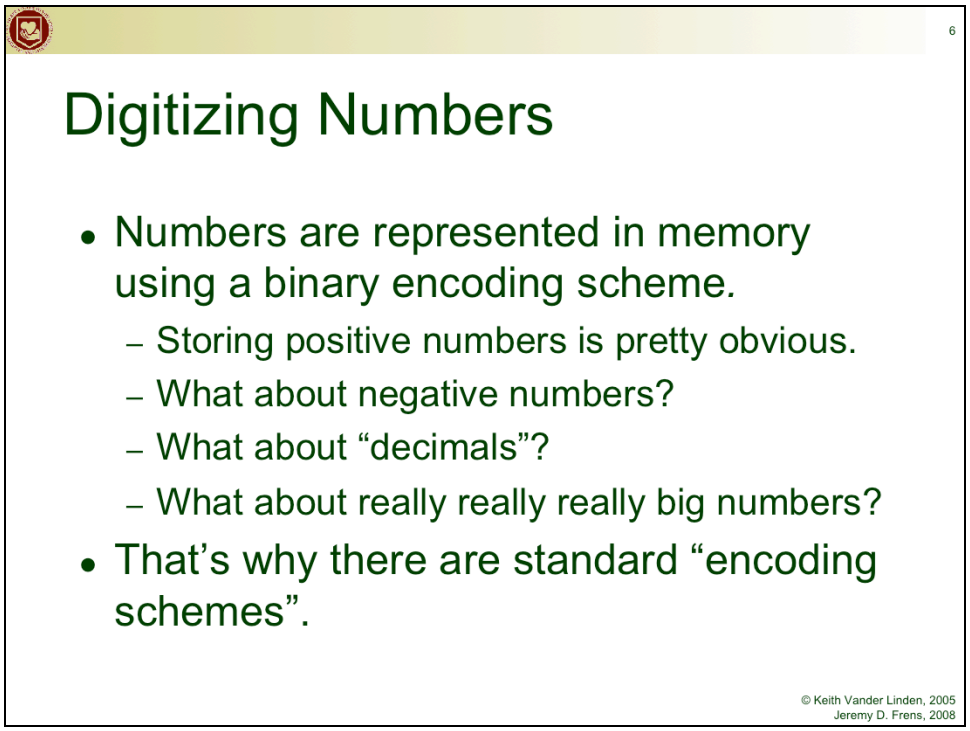
Decimal numbers are base-10 (using digits 0-9)



Binary Numbers

Binary numbers are base-2 (using digits 0 & 1)

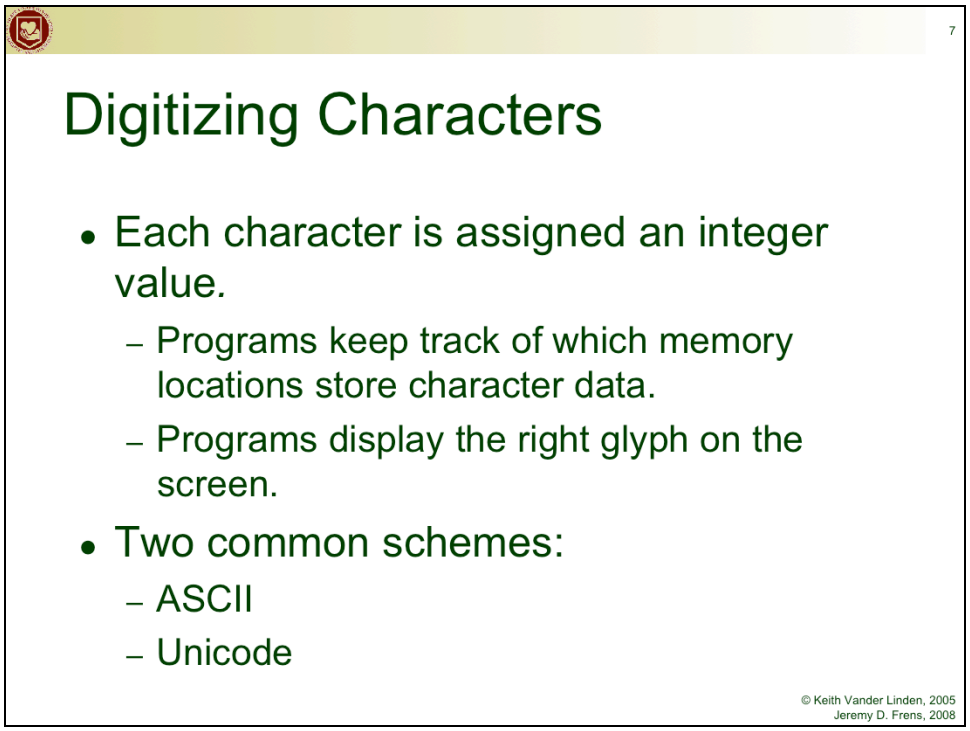




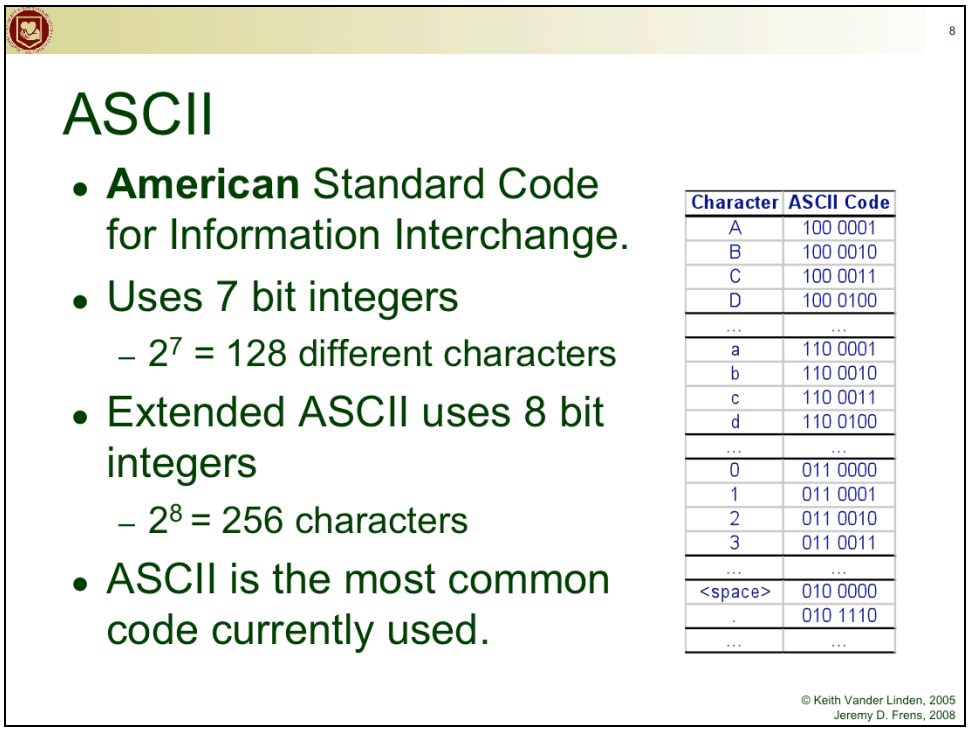
It’s all binary underneath!

2’s compliment for integers

Floating point for real numbers



It's all binary underneath!



Note the “numerals” have their own ASCII code which means that “1” is different from 1. Confusing!

Unicode

- Uses *8 - 32 bit integers*
  - over a million characters defined.
- Unicode supports a number of different character types.
  - Cyrillic
  - Ancient Coptic
  - All charts

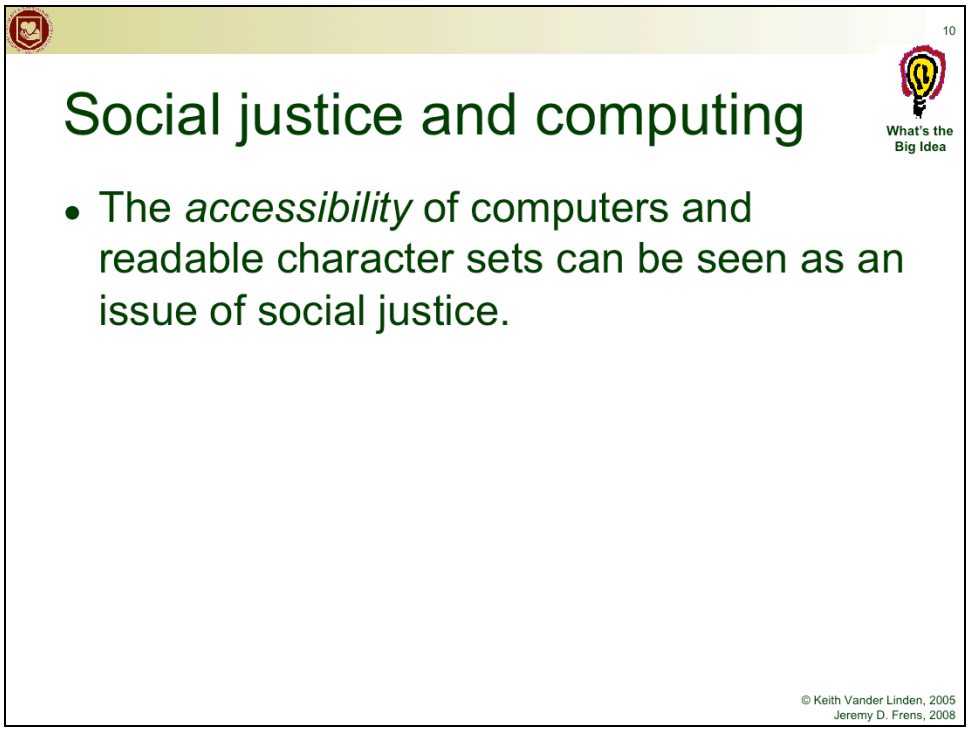
This is just a small part of the full unicode support for chinese characters.

Unicode is becoming more and more common.

unicode.org/charts

http://www.unicode.org/charts/PDF/U0400.pdf

http://www.unicode.org/charts/PDF/U2C80.pdf



Related to the digital divide material included in the computer anatomy lectures.

Given the English-centric nature of the web, one might more accurately call it the ***Western*-wide web.**

Digital divide – the WWW is hard to access in:

- the developing world
- the non-western world
- underpriviledged social classes
- the disabled community

What could we do to help bridge this divide?

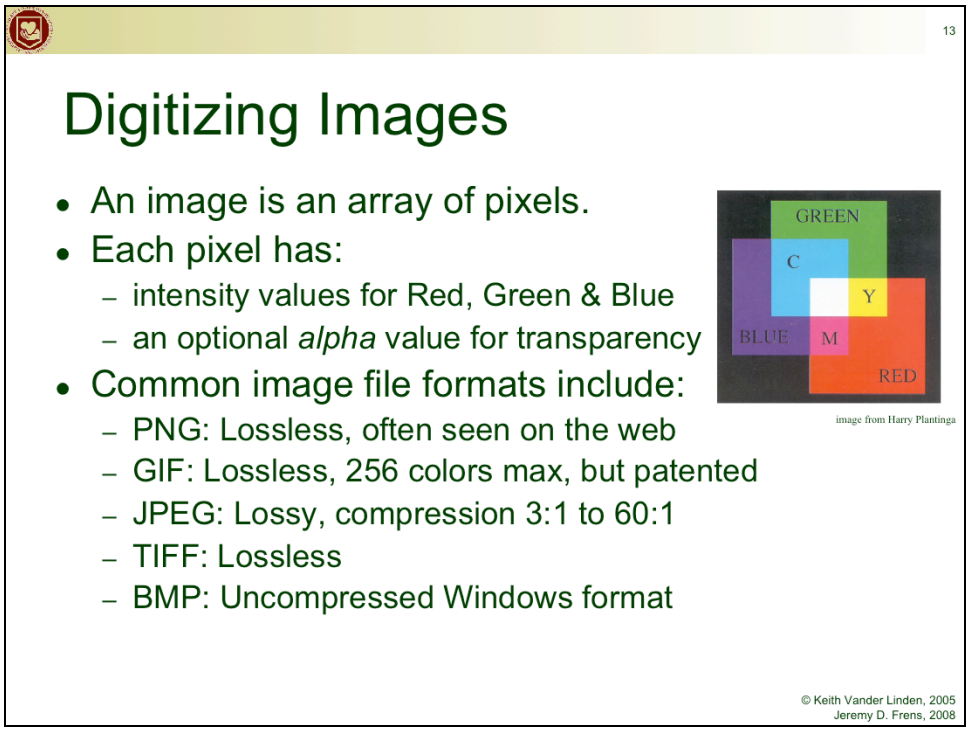
- Unicode
- internationalized domain name resolution
- better translation tools
- better international/disabled design and testing

Digitizing Big Data

- Using numbers and characters, we can digitize and model bigger things:
  - Documents
  - Accountant's ledger
  - City maps
  - Human behaviors
  - Calendars
  - Images, audio, video
- Programs make sense of the data.

Digitizing Multimedia Data

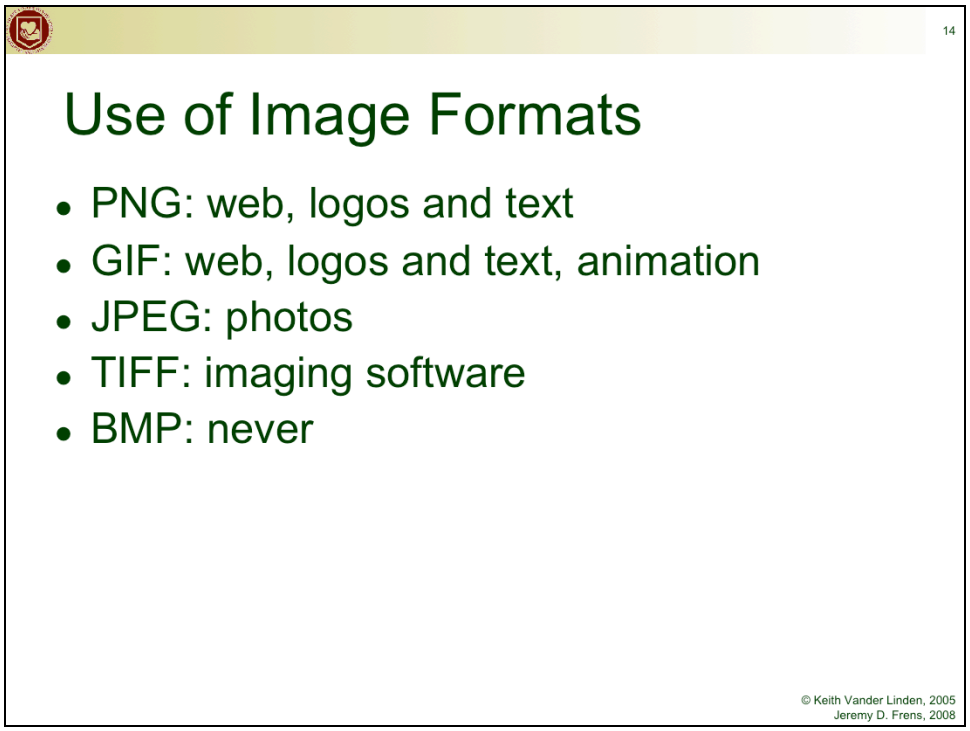
- Multimedia data is usually
  - ***HUGE***
  - And highly patterned
- *Compress* the data by taking advantage of the patterns to take up less space.
  - Lossless compression doesn't lose any information.
  - Lossy compression loses some information for better compression.



Check out this article on digitally edited photos:

Can Photos Be Trusted? - Popular Science

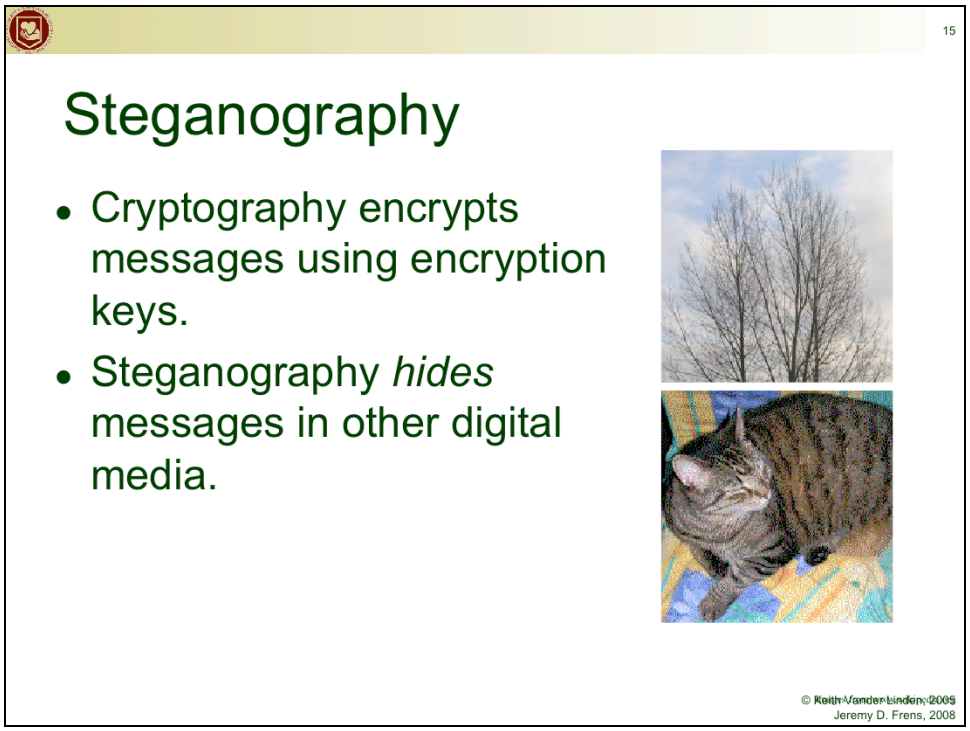
http://www.popsci.com/popsci/technology/generaltechnology/d6002684e4646010vgnvcm1000004eecbccdrcrd.html



Check out this article on digitally edited photos:

Can Photos Be Trusted? - Popular Science

http://www.popsci.com/popsci/technology/generaltechnology/d6002684e4646010vgnvcm1000004eecbccdrcrd.html



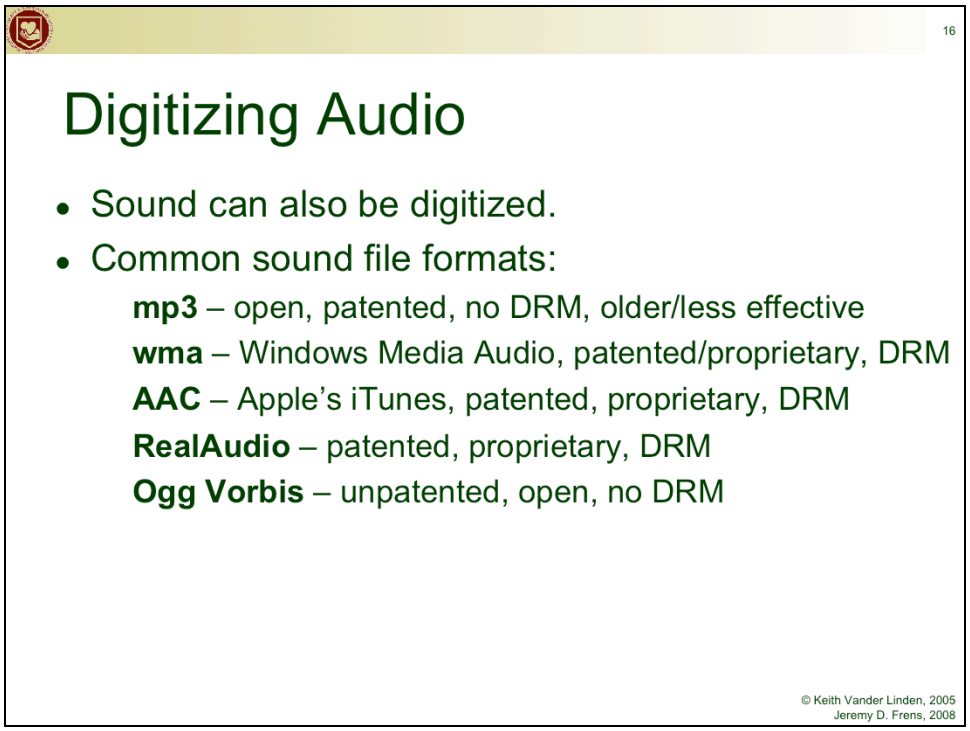
The bottom picture was hidden in the last 2 bits of the pixel codes of the top picture and recovered. People looking at the original would never have known that the cat was there, except, perhaps, by noticing that the relatively large size of the image file is not congruent with the relatively poor resolution of image.

http://en.wikipedia.org/wiki/Steganography

http://www.calvin.edu/~lave/s-tools/

This works for GIF but not for JPG (because of the way JPG codes the colors for compression).

http://www.stegoarchive.com/



DRM – digital rights management (see wikipedia)

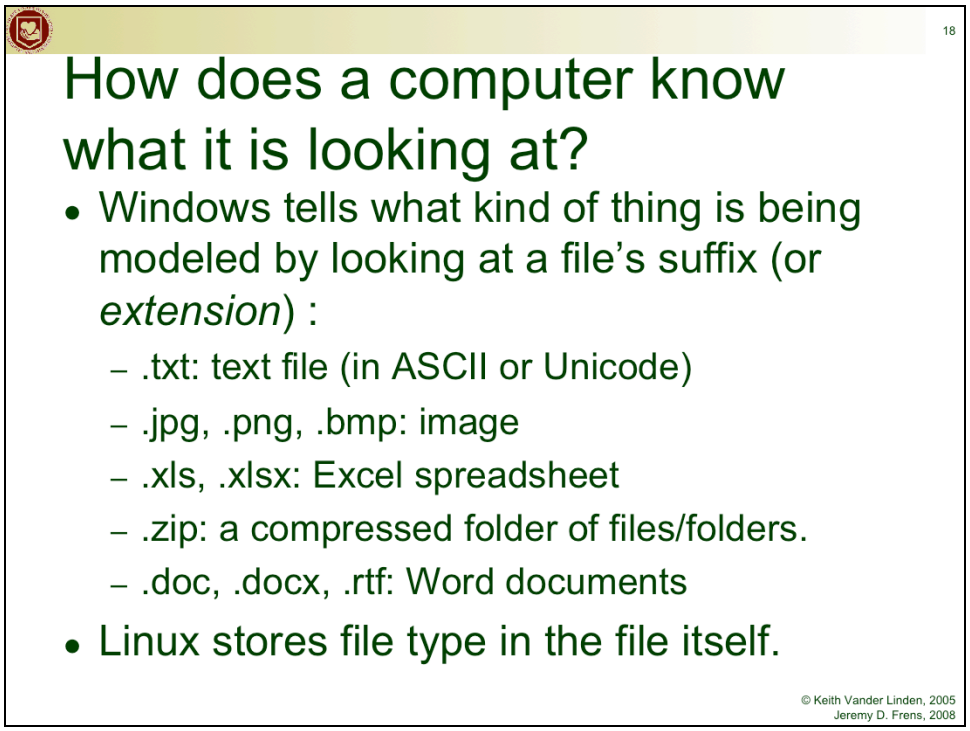
Lossy vs. lossless

Fights over "standards"

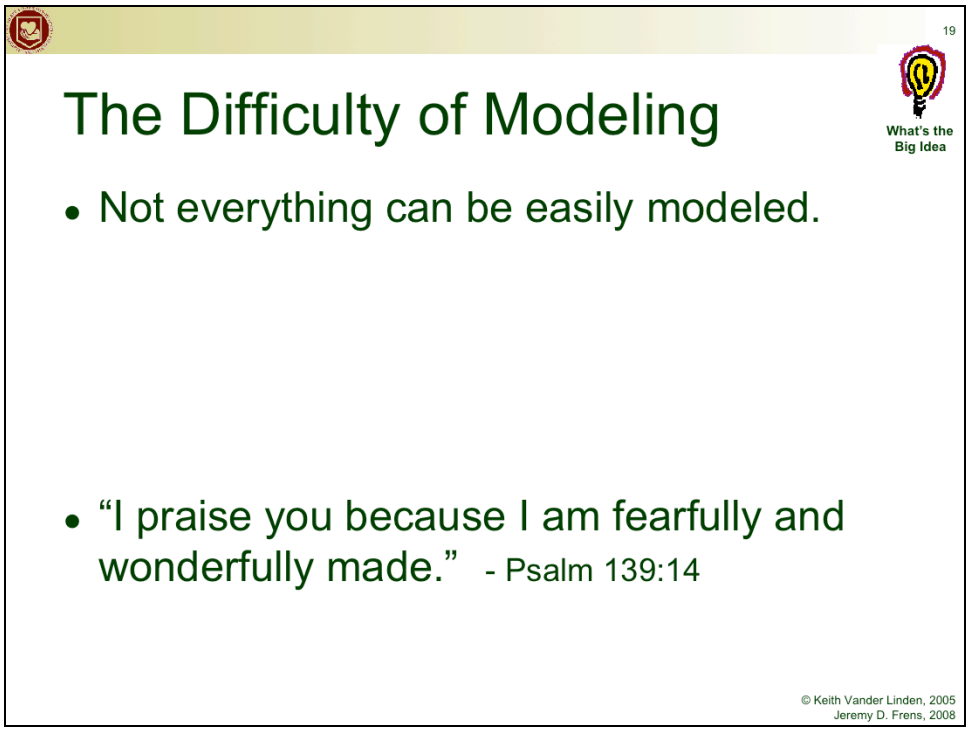
DRM is an ever-more-important issue

Digitizing Video

- Common movie file formats:
  - **mpeg** – Open (but patented) standard
  - **avi** – Windows Media Player
  - **DV** – As used in digital camcorders
  - **divx** – very high compression ratios



If you change the suffix of a file name, Windows thinks it is a different kind of file and will try to open it with a different program.



The weather (just too much stuff) (although this is getting much better every year)

The human genome (just too much stuff we don't understand)

Human intelligence (AI – to the sussman anomaly example here) - Easy things are hard, hard things are easy. E.g., Being human is harder than it looks. "One year in AI is enough to make one believe in God" – Alan Perlis.