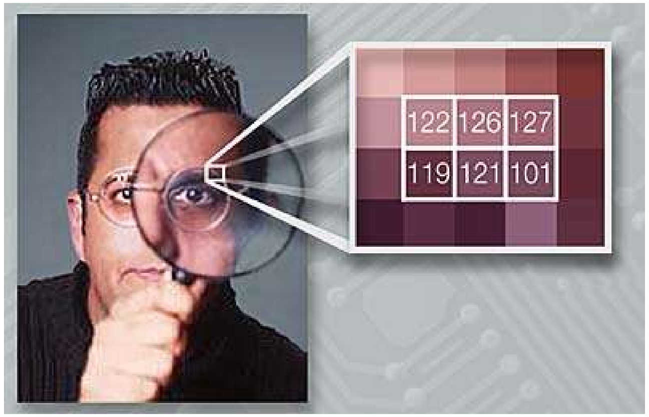
Steganography: The Art of Hiding Data In Plan Sight



A Capstone Project Presentation
By Jonathan Gould

Image #1

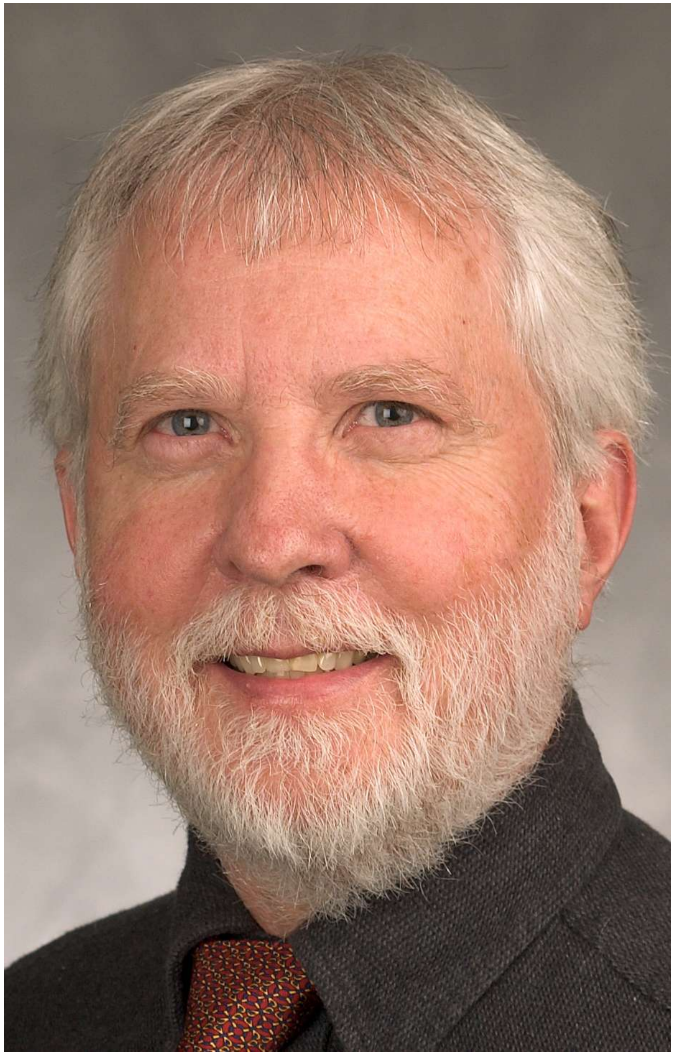


Image #2

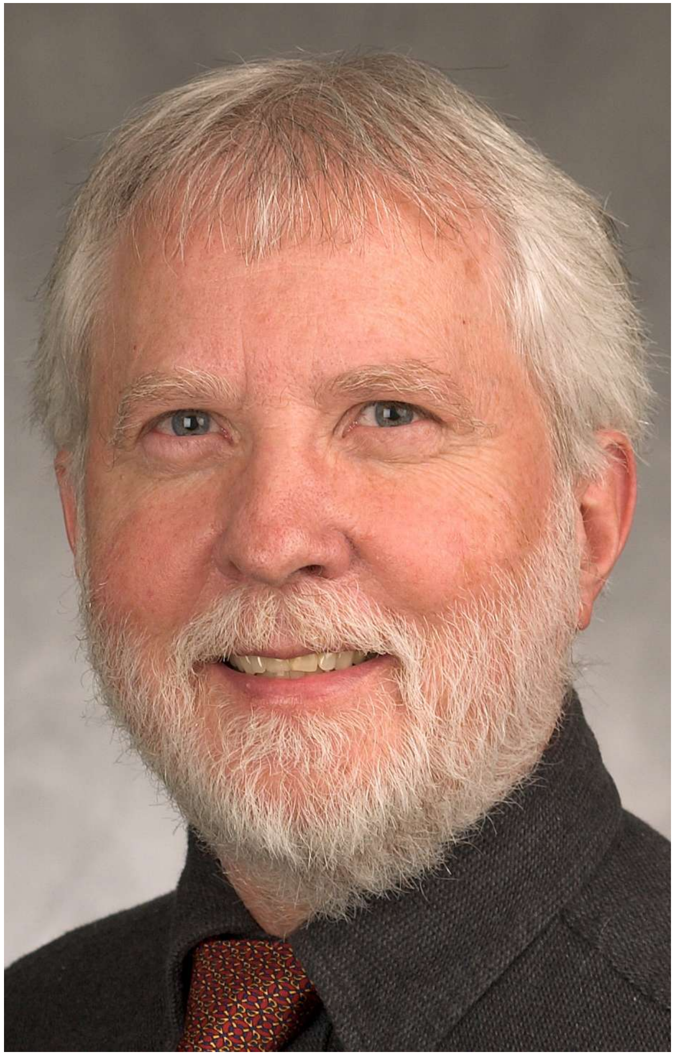


Image #3

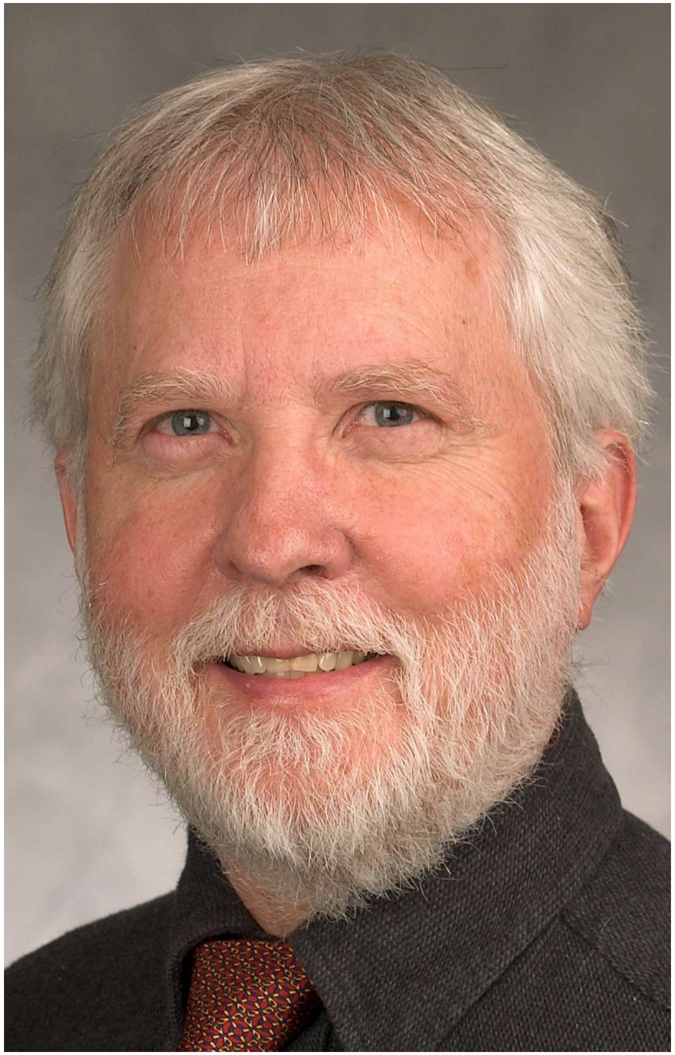


Image #4

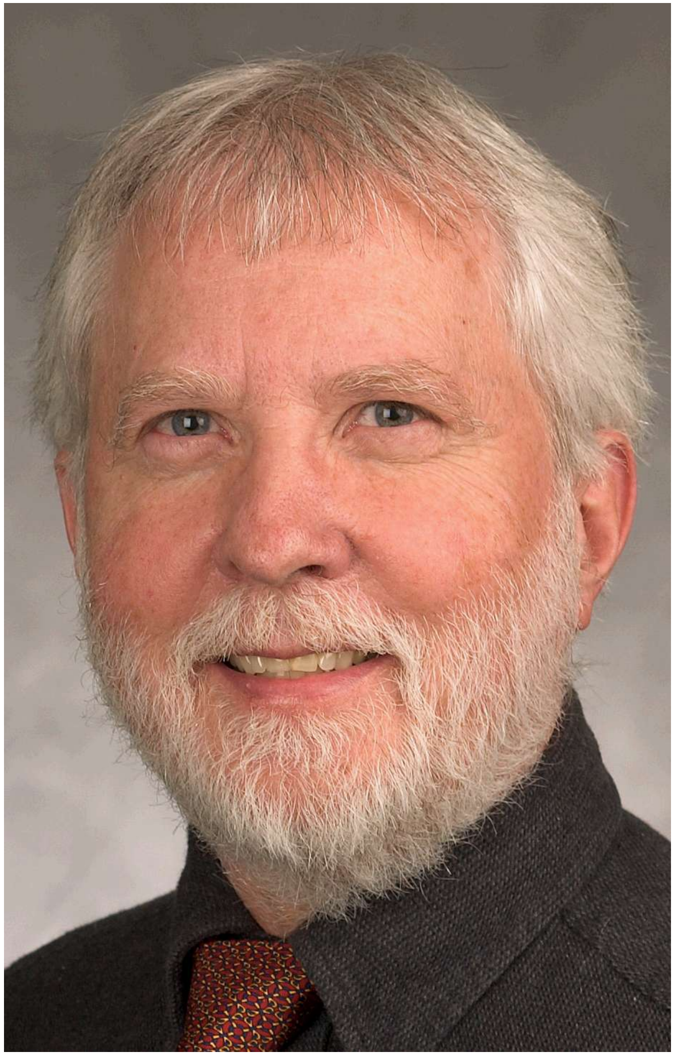


Image #5

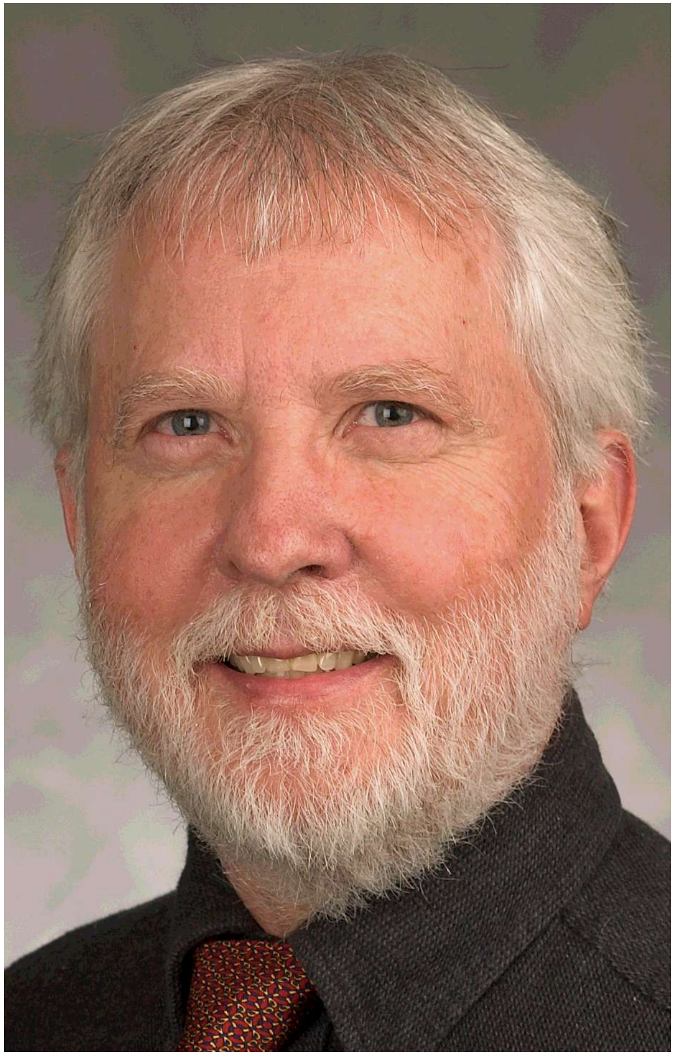


Image #6

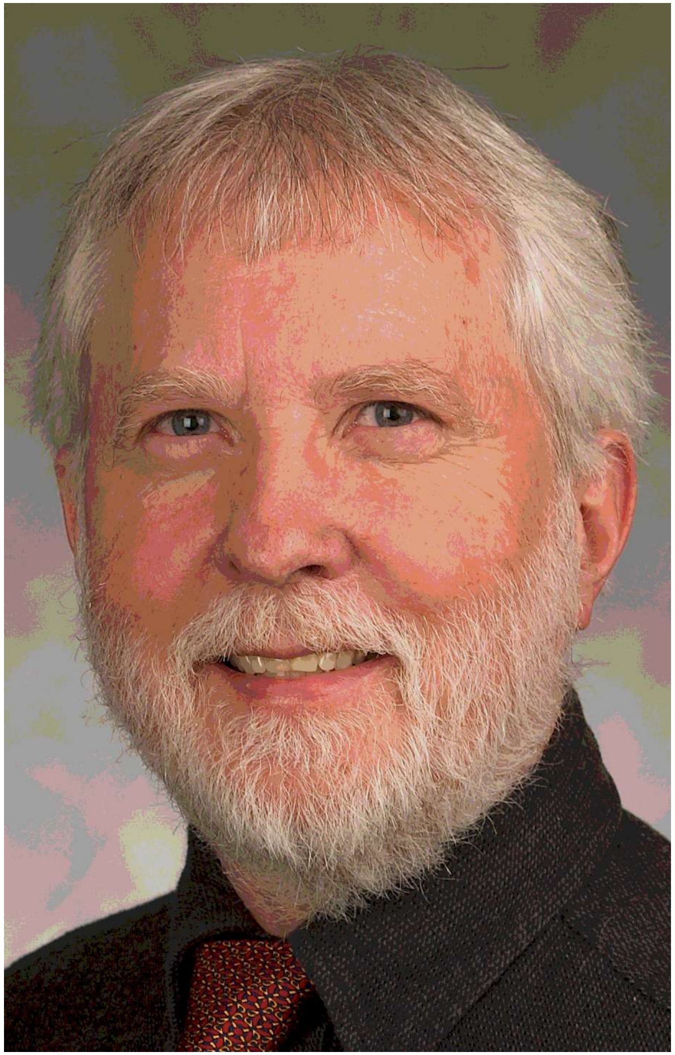


Image #7

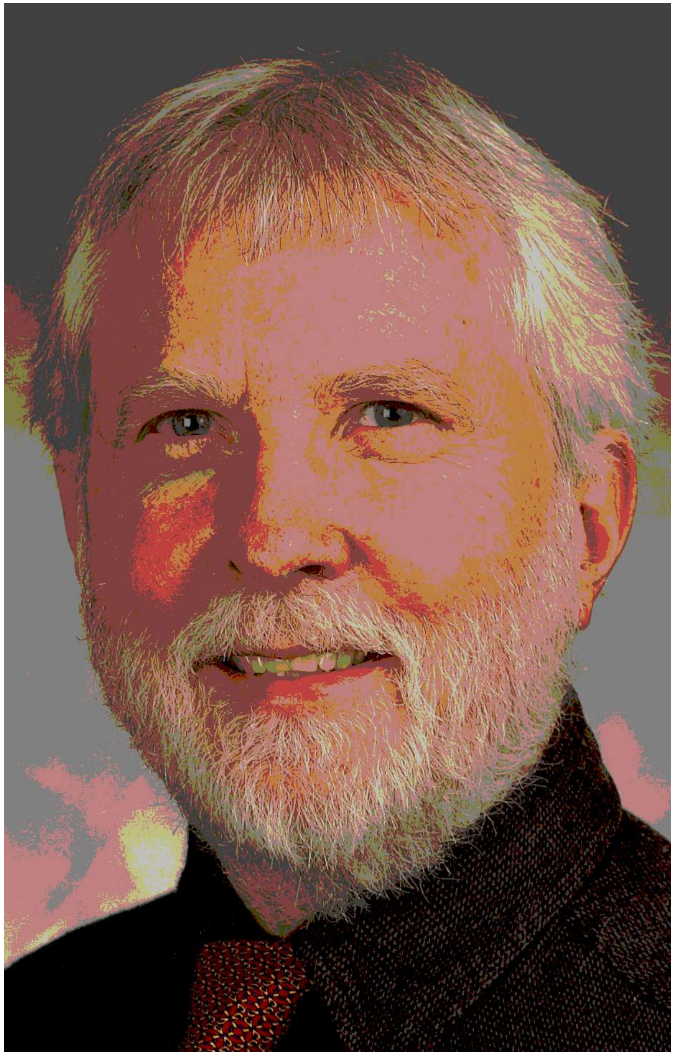


Image #8

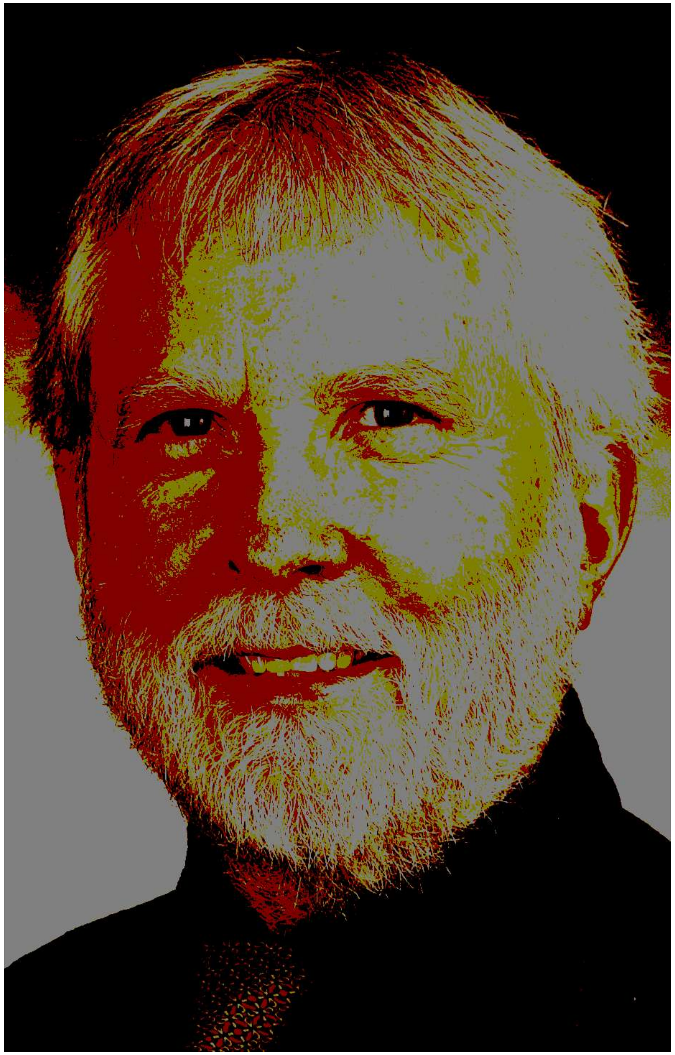


Image #9

[redacted]

How image file formats like .png and .jpg store data

- Every pixel's data is stored in three bytes:
  - How much red (0-255)
  - How much green (0-255)
  - How much blue (0-255)
- Combined, they form 16 million (2^{24}) colors.
- Many of these colors are very similar!

Two different but similar colors

Red = 33
Green = 72
Blue = 105

0	0	1	0	0	0	0	1
0	1	0	0	1	0	0	0
0	1	1	0	1	0	0	1

Red = 32
Green = 73
Blue = 104

0	0	1	0	0	0	0	0
0	1	0	0	1	0	0	1
0	1	1	0	1	0	0	0

Taking Advantage of This

- So we can drop a few of the least signficant bits (LSBs) without significantly damaging image quality.
- What if we used those bits to hold data? We could use image files to store:
  - Text
  - Other images
  - Audio
  - Anything!

My Software

- I wrote programs to do the following:
  - Remove 1-8 LSB from an image and replace them with zeros, ones, or a random combination of the two.
  - Hide either text or an image within another image file, using:
    - one, two, or three LSB and
    - one of five pixels patterns

My Software (cont.)

- Extract from an image data (image or text) hidden by the previous program.
- Determine whether a given image file is likely to contain text hidden in one of the ways discussed on the previous slide.

Sending Text In Secret

- So if two parties had this software:
  - One could take private text and encode it in a picture of a cat.
  - Upload it to an image-sharing website along with forty-nine other cat pictures.
  - The second party runs all fifty through the detection program, finds the right one, and extracts from that image the message.

Brief Demonstration

The Tools I Needed/Learned

- How to write reusable code.
- How to write flexible code – code that could be used after the addition of new features.
- How to write/refactor code to be easily read by others and by myself.

In terms of steganography, my biggest lesson was ..

Advantage: Hacker

- Any manner of obfuscation can work, so long as both parties know the pattern.
- There's always a way to further mask data.
- There is no good way to stop this technique other than brute-force.

The Risk of Misuse

- Conceivably, steganography can be used to:
  - Transmit stolen data such as blueprints of government installations
  - Leak the technology behind nuclear/chemical/biological weapons
  - Disseminate child pornography
  - Facilitate any number of illegal activities
- So it's bad, right?

Not Always

- The same tools used to defeat “good” governments can be used to subvert tyrannical ones.
- Steganography can aid those fighting for government transparency (WikiLeaks) or for open access (the late Aaron Swartz).
- … so it's good then?

… But thinking makes it so

- Steganography: a tool like any other.
- Our role is to understand it in theory, to recognize its potency in practice, and to serve as educators for those who do not.

Questions?

Thank you all
for your time
and attention!