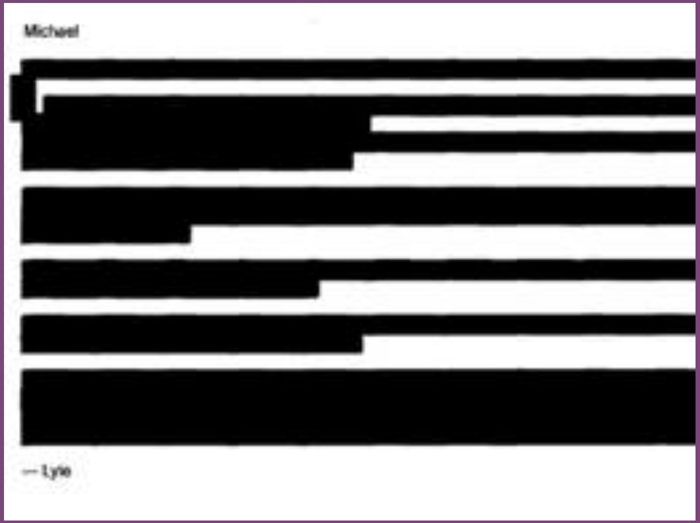
Original
Empower your people
with the IT tools they need
Learn More
Altered
Zwiększ możliwości pracowników swojej fi
za pomocą odpowiednich narzędzi informatycznych

CS121: Our Digital World



Michael

[redacted]

— Lyle

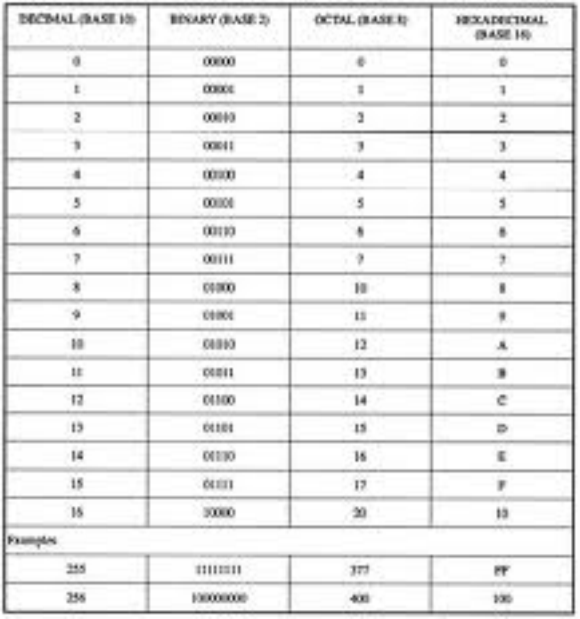
Redacted Documents

Activity: Badly Redacted Documents

- Download the following file:
  - http://zeus.cs.pacificu.edu/shereen/cs121sp12/redacted.pdf
- Can you find the secret message?

Electronically Redacted Document Examples

- TSA: Screening Management Standard Operating Procedure
  - Two examples: 2009 and 2010
- US Military
- Washington Post
- What went wrong?
  - WYSIWYG



DECIMAL (BASE 10)	BINARY (BASE 2)	OCTAL (BASE 8)	HEXADECIMAL (BASE 16)
0	00000	0	0
1	00001	1	1
2	00010	2	2
3	00011	3	3
4	00100	4	4
5	00101	5	5
6	00110	6	6
7	00111	7	7
8	01000	10	8
9	01001	11	9
10	01010	12	A
11	01011	13	B
12	01100	14	C
13	01101	15	D
14	01110	16	E
15	01111	17	F
16	10000	20	10
Examples			
255	11111111	377	FF
256	100000000	400	100

+

Number Systems

ASCII, binary, decimal, hex

+ ASCII

- In text files, each character is represented as an ASCII character
- ASCII Table: http://www.neurophys.wisc.edu/comp/docs/ascii/ascii-printable.html
- So, the letter ‘a’ is represented as:
  - Decimal: 097
  - Binary: 01100 0001
  - Hex: 061
- ‘A’ is represented as:
  - Decimal: 065
  - Binary: 0100 0001
  - Hex: 041

+ Decimal Numbers

- Start counting: 0, 1, 2, 3, 4, 5, 6, 7, 8, 9
- Oops! We've ran out of digits
- Add a second column: 10, 11, 12, …., 19, 20, 21, …, 98, 99
- Oops! We've ran out of digits
- Add a third column and so on…

Decimal Numbers

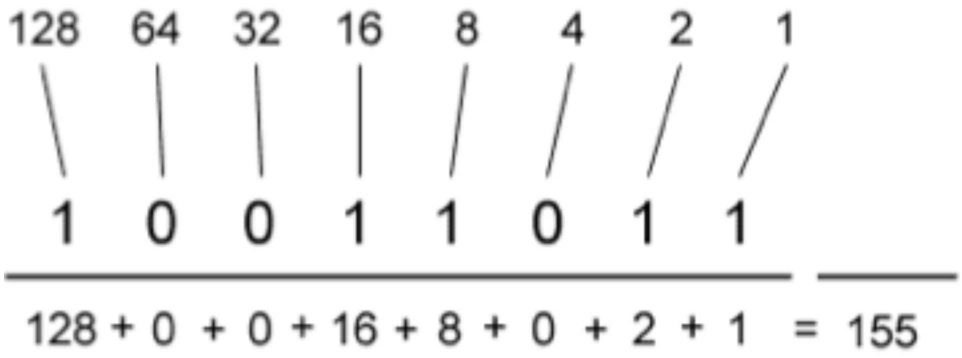
- Another way:
- 365 is:
  - 3 x 100 + 6 x 10 + 5 x 1
  - Or $3 \times 10^2 + 6 \times 10^1 + 5 \times 10^0$

Binary Numbers

- Start counting 0, 1
- Oops, we’ve run out of digits
- Add a column: 00, 01, 10, 11
- Oops, we’ve run out of digits
- Add another column: 000, 001, 010, 011, 100, 101, 110, 111
- And so on.

+ Binary Numbers

- Another way:
  - 1010 is:
  - $0 \text{ x } 2^3 + 1 \text{ x } 2^2 + 0 \text{ x } 2^1 + 1 \text{ x } 2^0$



128 64 32 16 8 4 2 1

1 0 0 1 1 0 1 1

128 + 0 + 0 + 16 + 8 + 0 + 2 + 1 = 155

+ Activity: Binary Numbers

- Convert these numbers from binary to decimal
  - 10
  - 111
  - 10101
  - 11110
- Represent the following decimals in binary
  - 1
  - 3
  - 6
  - 9

+

Track Changes & Metadata

Changes that can be Traced

- Track Changes
- Metadata
  - Which also can be forged

Track Changes Embarrassment

- UN document on the assassination of Lebanese Prime Minister Rafiq hariri
- Dodgy Dossier used by British government to justify their 2003 entry into the war on Iraq

Representation

- ASCII vs. raster
- What are the pros and cons of each?
- Example:
  - Book of Kells

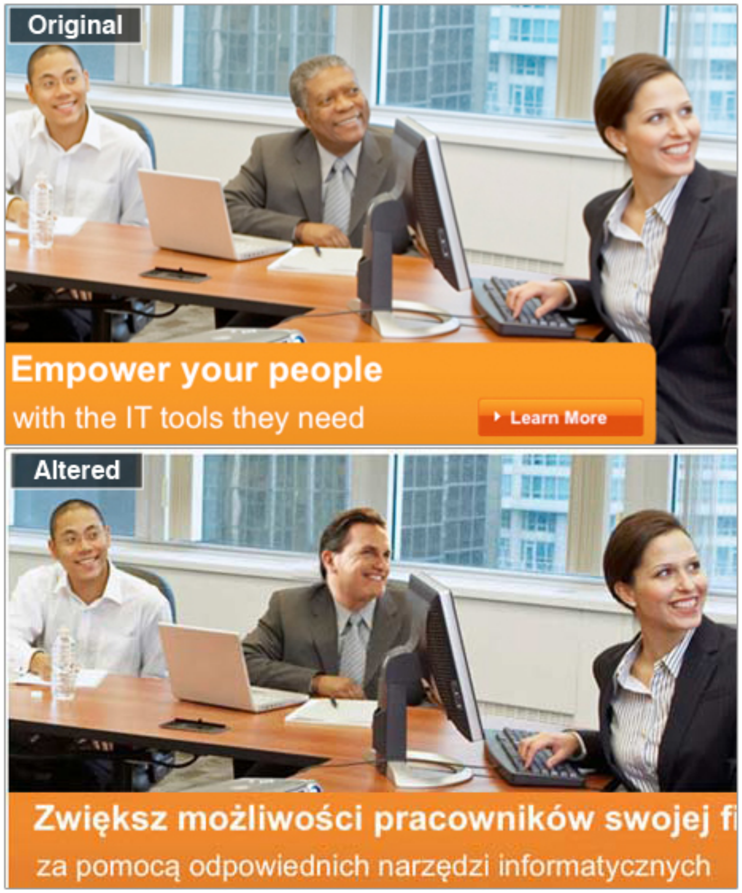
+
Compression

- Compression is necessary sometimes due to the size of the original image or file
- Two kinds of compression:
  - Lossless
  - Lossy
- Raster example: http://g2cs.org/media/interactives/image-compression/

Hiding Things

Photo Editing

- Microsoft edited a photo advertisement


Original
Empower your people
with the IT tools they need
Learn More
Altered
Zwiększ możliwości pracowników swojej fi
za pomocą odpowiednich narzędzi informatycznych

+
Steganography

- It is the art and science of writing a message in a way to where the only the recipient knows of its existence.
- The word **Steganography** is of Greek origin and means "covered, or hidden writing."
- Steganographic messages will generally appear as something else such as a picture or a text file.
- Provide security over insecure channel

+ Steganography

+

Watermarked pdf Example

+ Activity: Watermarked PDFs

- Download the pdf file to the desktop http://zeus.cs.pacificu.edu/shereen/cs121sp12/mad-men-denouement.pdf
- Open it and view it in Preview.
- Open Terminal and type the following:
  - cd Desktop
  - head -2 mad-men-denouement.pdf
- What did you notice?

Watermarked PDFs

- Download the python script (http://zeus.cs.pacificu.edu/shereen/cs121sp12/sign-pdf.py)
- Type the following in Terminal:
  - python sign-pdf.py mad-men-denouement.pdf JaneDoe
  - ls *.pdf
  - tail -2 JaneDoe.pdf
- Repeat the process above using a different name

Watermarked PDFs

- Okay, the files look identical, but can we tell the difference at the "bit" level? Yes. There is a way of boiling down a file (of any length) down to a 128 bit "signature". The technique is called MD5 (the "MD" stands for "Message Digest"). The digest goes by several names, such as "checksum" (since it's very often used to check the integrity of a file: is a copy correct?).

- Download the python file (http://zeus.cs.pacificu.edu/shereen/cs121sp12/md5.py) then type in the following in the terminal:
  - python md5.py *.pdf

- Notice that each file has been boiled down to a "signature" or "checksum". In this case, it's 128 bits long, and the bits are described by a string of *hexadecimal digits. This allows us to detect any tampering with the file, as well as being able to ascertain identity.*

Activity: Hard Disks

- How do they work?
- How can hard disks be truly sanitized?
- Conduct a little research and come up with three techniques:

  1.
  2.
  3.

+ References

- http://boardingarea.com/blogs/thewanderingaramean/2009/12/the-tsa-makes-another-stupid-move/
- http://www.shaunakelly.com/word/sharing/publicexamplesoftrackchanges.html
- http://news.bbc.co.uk/2/hi/8221896.stm
- http://www.khanacademy.org/video/binary-numbers?topic=core-pre-algebra
- http://www.brown.edu/cis/policy/datarmv.php