# Commando Forensics: What Dongle?

Cory Altheide

# Who are you?

(I'm kind of a big deal)

- Forensics & IR for ~10 years
  - National Nuclear Security Agency
  - Google
  - IBM
  - Mandiant
  - Google... again.
- Long-time proponent of open source forensics
  - Wrote some papers
  - Worked on a couple books
  - Technical Committee for DFRWS

# What are you doing here?

- **Survey of open source forensics tools**
  - "Beyond the file system"
  - Extracting and exploiting higher-order artifacts
  - Extracting data outside of the file system structure
- **Not a detailed how-to**
- **Intentionally omitting or glossing over:**
  - File system
  - Memory forensics
  - Timelining

# Why should I care?

(about non-file system artifacts)

- More information == happy responders
- “Truthier” information in metadata
  - More effective timelining
  - More effective attribution
  - Better understanding of system activities
- You can't always get file system information
- You should use the whole buffalo

# Why should I care?

(about open source tools)

- Open tools make you smarter
  - (through necessity)
- Open tools can be reviewed
  - (but are you actually going to do so?)
- Open tools can be modified
  - (but are you capable of doing so?)
- Did I mention that I'm kind of a big deal?

# What Artifacts Are We Talking About?

## Carving

- Scalpel
- Foremost
- PhotoRec
- Ftimes
- Post-carve processing

## Documents & Images

- Office Documents
  - OLE Binary
  - Open XML
- PDF
- JPEG
- PNG

# What Artifacts Are We Talking About?

## Internet Artifacts

- Browser History
  - IE
  - Firefox
  - Safari
  - Chrome
- Mail
  - MBOX/Maildir
  - Outlook

## Windows Artifacts

- Event Logs
- Registry
- Prefetch
- LNK files
- Recyle Bin Records

# Carving

# Carving

aka “making lemonade”

- Hunting for headers (and footers) in unallocated space
  - Gross simplification (sorry guys)
- Various “smart carving” mechanisms:
  - Knowledge of the file system
  - Knowledge of the file(s) being carved

PUMPKIN MASTERS
Deluxe
Pumpkin Carving Kit
PATTERN BOOK
12 Patterns • 1 4-in-1 Tool • 2 Saws • 1 Drill • 1 Super Poker • 2 Candles

# Foremost & Scalpel

- Scalpel forked from Foremost 0.69
- Both updated since then
- Both define headers/footers
- Both can restrict to user defined block boundaries
  - Useful if you know the cluster size
- Neither appears to have a major competitive advantage over the other
  - I am happy to be argued with at this point

# PhotoRec

- Much more “automagic” than scalpel/foremost
  - Autodetects block size
  - Extracts based on knowledge of the file type being extracted
  - Rolls knowledge about extracted blocks into subsequent extraction attempts
- A great tool to try if human resources are streched thin

# FTimes

- “system baselining and evidence collection tool”
  - Incredibly robust analysis and carving utilities
  - Incredibly steep learning curve
  - Can be used to automate 80% of the work – applying examiner intelligence to the other 20% can yield big results
- ftimes-crv2raw.pl - main carving script
- Xmagic - ftimes eXtended magic file
- Ftimes is capable of much more
  - Outside of the scope of this talk

# Post-carving

- **Problem: The stuff you carve out might suck**
  - Fragmentation may mean you have a partially clobbered file
  - Truncation/overwriting may mean you have a snipped file
  - Cleaning it up is not impossible
- **Frag find (part of NPS bloom package @ afflib.org)**
  - **Used to identify portions of a known file in unallocated space**
- **ssdeep**
  - Fuzzy hashing to determine similarity between two files

# What *don't* we have?

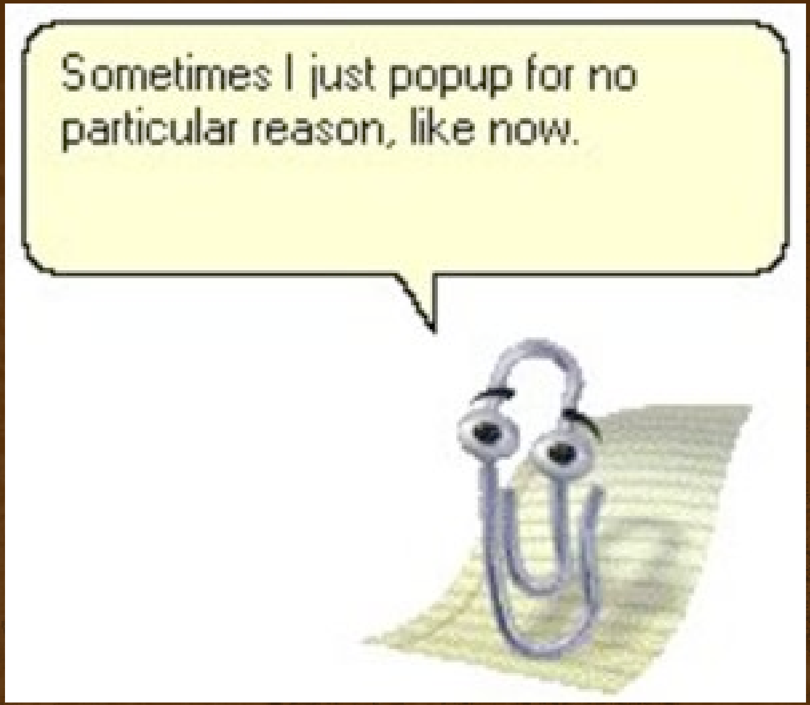
- Viable open-source “smart” carving
- Current useable zero-storage carving
- Parallelized carving
- ???

# Documents & Images

# Documents & Images

(thousands of words worth of pictures. also, thousands of words)

- **Documents**
  - Containers full of text (or numbers)
- **Images**
  - **Containers full of image data**
- **PDFs**
  - **Containers full of both!**
  - **(and remote code execution)**



# Microsoft Office Binary

- OLE Structured Storage
  - A miniature dedicated FAT-esque file system
  - Various streams contain useful metadata
  - Various streams have independent time stamps.
- Libforensics:
  - olecat: dumps an OLE stream
  - olels: list entries (streams) in a compound file
    - -m: display in mactime format
  - olestat: compound file statistics (like 'fsstat')
  - wmg.py: extracts MS Word specific metadata

# OpenXML

- Office 2007+/Open Office
- Just a ZIP file with some XML and other junk inside!
  - Simson has a paper that describes this “other junk” well.
  - Zipped items have individual time stamps
  - XML documents may have application-specific metadata & time stamps as elements
  - Images may contain metadata from their point of origin (camera, editing application, geolocation)
- unzip -l for zip-contained time info
- sgrep and/or xmlstarlet for XML parsing/grepping

# JPEG

- **EXIF info**
  - Standard metadata format for JPEG images
  - Valuable info like:
    - original camera used to create image
    - “True” date/time of capture
    - Geolocation
- **jhead, exif, and exiv2 all present this information**

# PDF

- Abuse of PDFs could be an entire talk on its own
  - Why yes I would love a document that executes javascript and flash in the context of my local machine
  - Also please make this format mandatory for doing anything thanks.
- Didier Stevens' pdfid.py and pdf-parser.py
  - pdfid.py - initial triage
    - Scans PDF for specific strings, assigns count, makes judgement on general “shadiness.”
  - pdf-parser.py - deep dive
    - Command-line parser for PDFs - examine and extract PDF elements

# Random Metadata Extraction

- Hachoir-metadata
  - Part of the hachoir suite of tools
  - Can trivially extract metadata from 33 different file types
    - Interesting types include:
      - JPEG
      - PNG
      - Torrents
      - EXEs

# What *don't* we have?

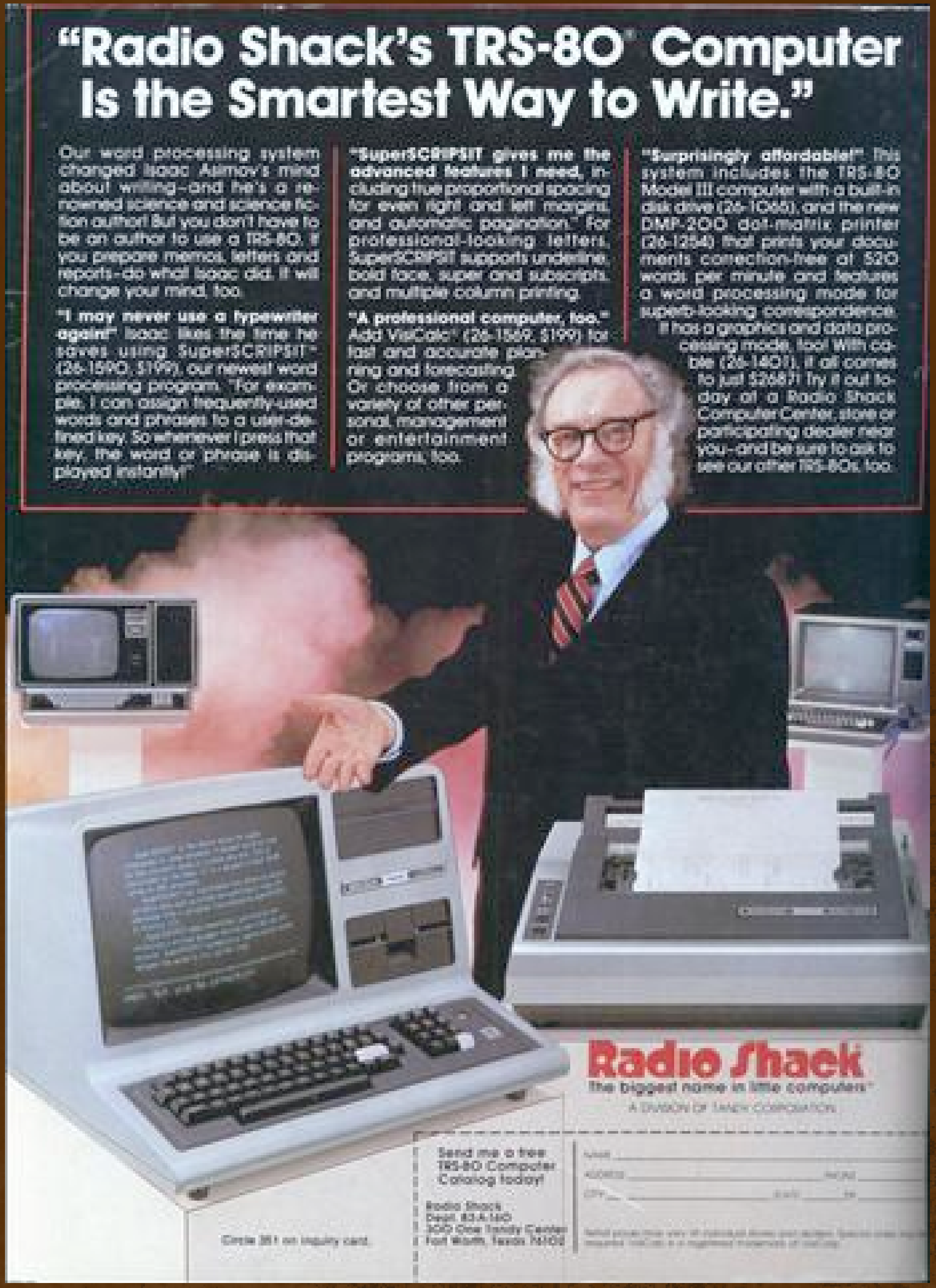
- Open source parser for Excel binary (BIFF) format
- Integration of metadata from files with file system information
- ???

# Internet Artifacts

# Internet Artifacts

(stuff that's left from the tubes)

- Web "history"
  - URLs accessed
  - Searches performed
  - Downloads
  - etc.
- Email
  - "local" mail storage only
  - Only covering mbox/maildir & Outlook



"Radio Shack's TRS-80® Computer Is the Smartest Way to Write."

Our word processing system changed Isaac Asimov's mind about writing—and he's a renowned science and science fiction author! But you don't have to be an author to use a TRS-80. If you prepare memos, letters and reports—do what Isaac did. It will change your mind, too.

**"I may never use a typewriter again!"** Isaac likes the time he saves using SuperSCRIPSIT™ (26-1590, $199), our newest word processing program. "For example, I can assign frequently-used words and phrases to a user-defined key. So whenever I press that key, the word or phrase is displayed instantly!"

**"SuperSCRIPSIT gives me the advanced features I need,** including true proportional spacing for even right and left margins, and automatic pagination." For professional-looking letters, SuperSCRIPSIT supports underline, bold face, super and subscripts, and multiple column printing.

**"A professional computer, too."** Add VisiCalc® (26-1569, $199) for fast and accurate planning and forecasting. Or choose from a variety of other personal, management or entertainment programs, too.

**"Surprisingly affordable!"** This system includes the TRS-80 Model III computer with a built-in disk drive (26-1066), and the new DMP-200 dot-matrix printer (26-1254) that prints your documents correction-free at 520 words per minute and features a word processing mode for superb-looking correspondence. It has a graphics and data processing mode, too! With cable (26-1401), it all comes to just $2687! Try it out today at a Radio Shack Computer Center, store or participating dealer near you—and be sure to ask to see our other TRS-80s, too.

Radio Shack
The biggest name in little computers™
A DIVISION OF TANDY CORPORATION

Send me a free TRS-80 Computer Catalog today!

Radio Shack
Dept. 83-A-160
300 One Tandy Center
Fort Worth, Texas 76102

NAME ____
ADDRESS ____ PHONE ____
CITY ____ STATE ____ ZIP ____

[illegible]

Circle 351 on inquiry card.

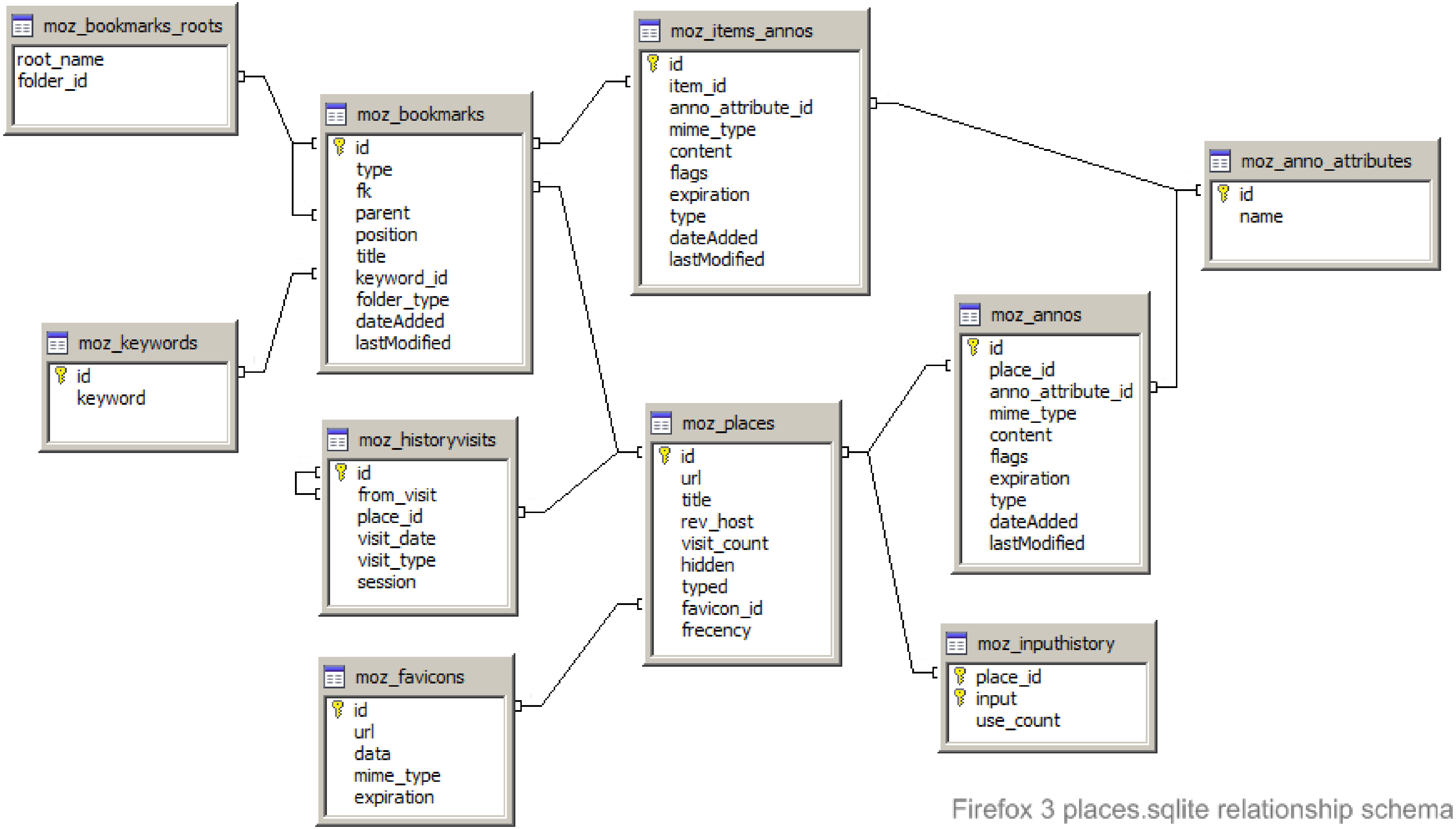
# Internet Explorer

- **Old and busted: pasco. New hotness: libmsiecf**
- **libmsiecf -**
  - msiecfinfo: index.dat metadata
  - msiecfexport: dumping index.dat records

# Firefox 3+

- Stores history info in SQLite database(s)
  - Schema mapped out and available visually at firefoxforensics.com
- Who needs forensics tools?
  - Pick your favorite SQLite3 interface, graphical or other
  - Start browsing or issuing SQL queries

# Firefox 3+



Firefox 3 places.sqlite relationship schema
www.firefoxforensics.com

# Safari

- Stores history in binary plist file
  - Can be opened on a Mac with Plist Editor
  - Can be parsed directly with JAFAT Safari Forensics Tools
- Cache is stored in a SQLite database
  - Process directly using SQLite tools, simple schema

# Google Chrome

(just the best browser ever, you're welcome)

- **Surprisingly enough, a SQLite database**
  - Schema defined... in the source.
  - Summarized nicely on the SANS blog: http://blogs.sans.org/computer-forensics/2010/01/21/google-chrome-forensics/
- **Moral of the story:**
  - Learn you some SQL

# Email

- “Unix” formats:
  - MBOX
  - Maildir
- Outlook PST/OST

# “Unix” formats

- MBOX & Maildir
  - “Beyond the file system”
  - Extracting and exploiting higher-order artifacts
  - Extracting data outside of the file system structure
- Grepmail for quick & dirty
- Mairix for more robust/larger inquires

# Outlook

- OST/PST/PAB are all stored in “Personal Folder File” (PFF) format
- libpff -
  - pffinfo – displays metadata about the PFF file
  - pffexport – dumps items (including deleted/recovered) from the PFF file.
- libnk2 - “nickfile” parser - Autocomplete file
  - nk2info – nickfile metadata
  - nk2export – dumps nickfile entries

# What *don't* we have?

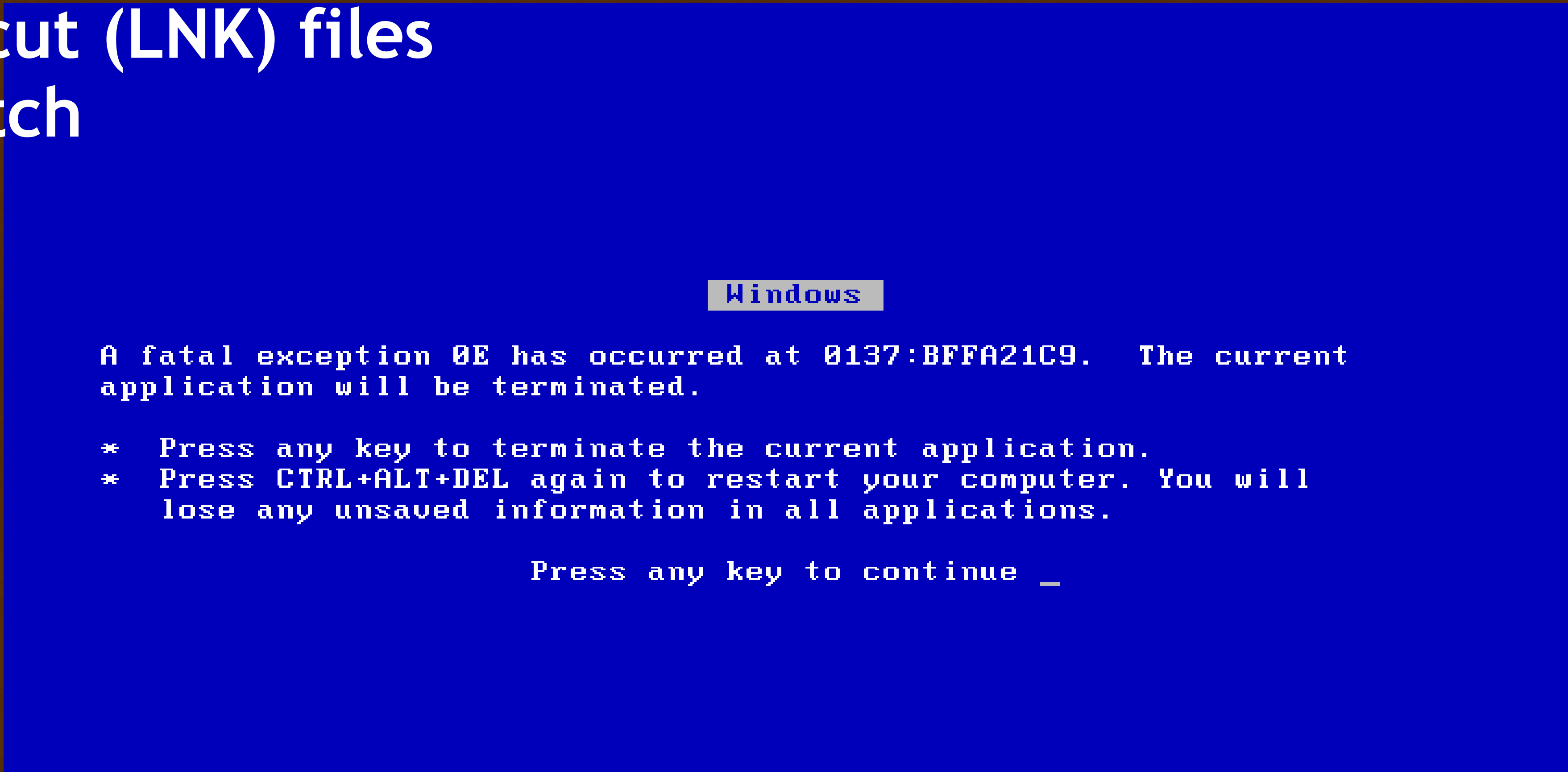
- Unified “automated” Internet History analysis
- Other mail format processing (Lotus, etc).
- … Open Source e-Discovery suite? …
- ???

# Windows-specific Artifacts

# Windows Specific Artifacts

*“Where we're going, we won't need Windows”*

- Registry
- Event Logs
- Recycle Bin Info
- Shortcut (LNK) files
- Prefetch



```
                         Windows

A fatal exception 0E has occurred at 0137:BFFA21C9.  The current
application will be terminated.

*  Press any key to terminate the current application.
*  Press CTRL+ALT+DEL again to restart your computer. You will
   lose any unsaved information in all applications.

                      Press any key to continue _
```

# Registry

- Regripper & Friends
  - The only actively developed and updated registry parser around
  - Sane plugin architecture
  - Robust set of included plugins
- rr/rip - standard regripper (GUI v. CLI)
- regslack - identifies registry entries in registry slack space
- ripxp - runs regripper plugins against hives in restore points

# Event Logs

- evtrpt
  - Provides statistical information on events contained in provided event log
- evtparse
  - Dumps event logs to line-based log output
  - Current version outputs to TLN format

# Recyle Bin

- INFO2/$Recycle.Bin
  - INFO2 prior to Vista
  - $Recycle.Bin for Vista & 2008
  - Extracting data outside of the file system structure
- Libforensics -
  - info2ls - listing of INFO2 entries
  - info2stat - detailed listing of a single INFO2 entry
- rifiuti2 - operates like the classic "rifiuti", but supports INFO2 and Vista/2008 scheme.

# Shortcut (LNK) Files

- **LNK files**
  - LiNK files, aka shortcuts.
  - LNK files are particularly interesting artifacts with respect to local attribution
  - Can also help flesh out a timeline where access times are trampled
- **libforensics -**
  - lnkinfo – parses and presents full content of a single LNK file

# Prefetch

- **Prefetch** files are artifacts of the Windows "prefetcher"
  - Speeds up execution of files
  - Generates interesting data for us – including a span of executions, and number of executions in that span
  - Sometimes the only evidence that a piece of code ran on the system.
- **Prefetch Tool:** **http://code.google.com/p/prefetch-tool/**

# What *don't* we have?

- ???

# Miscellaneous Tools

# Miscellaneous Tools

- A grab bag of stuff I found interesting, but that didn't fit anywhere else.

# Poorcase

- **Perl script to map split DD image to mountable disk**
  - Using loopback devices and device mapper
  - If you are awesome you may remember doing this manually \m/
  - Now you don't have to
- **Note: You can also do this with affuse**

# AnalyzeMFT

- **Python script to parse and interpret a raw MFT**
  - File Name Attribute time stamps
  - MFT Record Numbers
  - Time stamps for unallocated MFT records
- **Current shortcomings:**
  - Doesn't recreate directory structure
  - Doesn't parse out data runs
- **Harlan if going to mention this too and he totally copied it from me. :(**

# Forensie

- “Google Wave Robot designed to perform very basic file forensic analysis”
    - MBR of a hard drive
    - FAT Boot Sector
    - FAT Dates/Times
    - Hex/Binary decoding
- Useful? Maybe not. Awesome? Yes.

AWESOME
It's Comforting to Know That no Matter What You do in Life
It Will Never Be as Awesome as This Picture.

Questions?
cory@google.com