MSIE Cache File (index.dat) format specification

Analysis of the index.dat file format

By Joachim Metz <joachim.metz@gmail.com>

Summary

The index.dat filename is used by Microsoft Internet Explorer to cache different types of information. This specification is based on the work by [JONES03] and [CHAPPELL10] and was complimented by reverse engineering of the file format.

This document is intended as a working document for the MSIE Cache File (MSIECF) specification. Which should allow existing Open Source forensic tooling to be able to process this file type.

Document information

Author(s): Joachim Metz <joachim.metz@gmail.com>

Abstract: This document contains information about the MSIE Cache File format.

Classification: Public

Keywords: MSIE Cache File, index.dat, Client URLCache MMF

License

Version

Version	Author	Date	Comments
0.0.1	J.B. Metz	May 21, 2009 May 22, 2009 May 23, 2009 May 24, 2009	Initial version.
0.0.2	J.B. Metz	June 13, 2009	Additional information about date and time values.
0.0.3	J.B. Metz	June 18, 2009 June 20, 2009	Additional information about leak records.
0.0.4	J.B. Metz	June 27, 2009	Additional information about DOMStore cache file.
0.0.5	J.B. Metz	September 12, 2009	Additional information about file format.
0.0.6	J.B. Metz	October 13, 2009 October 18, 2009	Small corrections.
0.0.7	J.B. Metz	October 24, 2009	Additional information about file format.
0.0.8	J.B. Metz	August 2010	Email change.
0.0.9	J.B. Metz	September 2010	Additional information about MSIE 9 beta.
0.0.10	J.B. Metz	August 2012	Email and license update.
0.0.11	J.B. Metz	January 2013	Additional information about file format.
0.0.12	J.B. Metz	March 2013	Small changes.
0.0.13	J.B. Metz	March 2013	Additional findings regarding MSIE 4.
0.0.14	J.B. Metz	April 2013	Additional findings regarding the header data, hashing algorithm and groups.
0.0.15	J.B. Metz	June 2013	Additional information regarding Download URL record data with thanks to A. Case.

Table of Contents

1. Overview

The MSIECF (MSIE Cache File) is used by Microsoft Internet Explores to cache different types of information. There are MSIECF named index.dat for multiple uses, some are:

- the Temporary Internet Files;
- the History;
- the Cookies;
- the Feed Cache.

The MSIECF is also known as the WININET Cache.

These index.dat file are stored in the corresponding directories.

A MSIECF consist of the following distinguishable elements:

- file header
  - cache directory list
  - allocation bitmap
- records
  - hash record
  - URL record
  - redirected record
  - leak record

Characteristics	Description
Byte order	little-endian
Date and time values	in both UTC and local time
Character string	ASCII strings are stored in extended ASCII with a codepage. Unicode strings are stored in UTF-16 little-endian without the byte order mark (BOM).

1.1. Versions

There are multiple version of the MSIECF format.

MSIE version	MSIECF version	Filename	Remarks
MSIE 3 (32-bit)	Unknown	mm256.dat mm1024.dat	256 byte record size 1024 byte record size
MSIE 4	4.7	index.dat	128 byte block based record size
MSIE 5 to 9	5.2	index.dat	128 byte block based record size

This document mainly focuses on MSIECF versions 4.7 and 5.2.

1.2. File types and locations

The History (periodic) in the following sections indicates both the History (daily) and History (weekly) MSIECF file types.

1.2.1. MSIE 4 on Windows 98

Type of MSIECF file	Characteristics
Temporary Internet Files (Cache)	%WINDIR%\Temporary Internet Files\index.dat
History (global)	%WINDIR%\History\index.dat
History (periodic)	%WINDIR %\History\\MSHist01yyyymmddyyyymmdd\index.dat
Cookies	%WINDIR%\Cookies\index.dat

1.2.2. MSIE 5 on Windows 2000

Type of MSIECF file	Characteristics
Temporary Internet Files (Cache)	%USERPROFILE%\Local Settings\Temporary Internet Files\Content.IE5\index.dat
History (global)	%USERPROFILE%\Local Settings\History\History.IE5\index.dat
History (periodic)	%USERPROFILE%\Local Settings\History\History.IE5\MSHist01yyyymmddyyyymmdd\index.dat
Cookies	%USERPROFILE%\Cookies\index.dat
User data	%USERPROFILE%\Application Data\Microsoft\Internet Explorer\UserData\index.dat

1.2.3. MSIE 6 on Windows XP SP1, Windows 2003

Type of MSIECF file	Characteristics
Temporary Internet Files (Cache)	%USERPROFILE%\Local Settings\Temporary Internet Files\Content.IE5\index.dat
History (global)	%USERPROFILE%\Local Settings\History\History.IE5\index.dat
History (periodic)	%USERPROFILE%\Local Settings\History\History.IE5\MSHist01yyyymmddyyyymmdd\index.dat
Cookies	%USERPROFILE%\Cookies\index.dat
User data	%USERPROFILE%\UserData\index.dat

1.2.4. MSIE 7 on Windows XP SP2

Type of MSIECF file	Characteristics
Temporary Internet Files (Cache)	%USERPROFILE%\Local Settings\Temporary Internet Files\Content.IE5\index.dat
History (global)	%USERPROFILE%\Local

Type of MSIECF file	Characteristics
	Settings\History\History.IE5\index.dat
History (periodic)	%USERPROFILE%\Local Settings\History\History.IE5\MSHist01yyyymmddyyyymmdd\index.dat
Cookies	%USERPROFILE%\Cookies\index.dat
Feeds Cache	%USERPROFILE%\Local Settings\Application Data\Microsoft\Feeds Cache\index.dat
User data	%USERPROFILE%\UserData\index.dat

1.2.5. MSIE 7 on Vista

Type of MSIECF file	Characteristics
Temporary Internet Files (Cache)	%USERPROFILE%\AppData\Local\Microsoft\Windows\Temporary Internet Files\Content.IE5\index.dat %USERPROFILE%\AppData\Local\Microsoft\Windows\Temporary Internet Files\Low\Content.IE5\index.dat
History (global)	%USERPROFILE%\AppData\Local\Microsoft\Windows\History\History.IE5\index.dat %USERPROFILE%\AppData\Local\Microsoft\Windows\History\Low\History.IE5\index.dat
History (periodic)	%USERPROFILE%\AppData\Local\Microsoft\Windows\History\History.IE5\MSHist01yyyymmddyyyymmdd\index.dat %USERPROFILE%\AppData\Local\Microsoft\Windows\History\Low\History.IE5\MSHist01yyyymmddyyyymmdd\index.dat
Cookies	%USERPROFILE%\AppData\Roaming\Microsoft\Windows\Cookies\index.dat %USERPROFILE%\AppData\Roaming\Microsoft\Windows\Cookies\Low\index.dat %USERPROFILE%\AppData\Local\Temp\Low\Cookies\index.dat
User data	%USERPROFILE%\AppData\Roaming\Microsoft\Internet Explorer\UserData\index.dat %USERPROFILE%\AppData\Roaming\Microsoft\Internet Explorer\UserData\Low\index.dat
Feeds Cache	%USERPROFILE%\AppData\Local\Microsoft\Feeds Cache\index.dat

1.2.6. MSIE 8 on Windows XP SP3

Type of MSIECF file	**Characteristics**
Temporary Internet Files (Cache)	%USERPROFILE%\Local Settings\Temporary Internet Files\Content.IE5\index.dat
History (global)	%USERPROFILE%\Local Settings\History\History.IE5\index.dat
History (periodic)	%USERPROFILE%\Local Settings\History\History.IE5\MSHist01yyyymmddyyyymmdd\index.dat
Cookies	%USERPROFILE%\Cookies\index.dat
User data	%USERPROFILE%\UserData\index.dat
InPrivate Filtering	%USERPROFILE%\PrivacIE\index.dat
Compatibility Cache	
TLD Cache	%USERPROFILE%\IETldCache\index.dat
Feeds Cache	%USERPROFILE%\Local Settings\Application Data\Microsoft\Feeds Cache\index.dat
DOM store	%USERPROFILE%\Local Settings\Application Data\Microsoft\Internet Explorer\DOMStore\index.dat

1.2.7. MSIE 8 on Windows 2008

Type of MSIECF file	**Characteristics**
Temporary Internet Files (Cache)	%USERPROFILE %\AppData\Local\Microsoft\Windows\Temporary Internet Files\Content.IE5\index.dat %USERPROFILE %\AppData\Local\Microsoft\Windows\Temporary Internet Files\Low\Content.IE5\index.dat
History (global)	%USERPROFILE %\AppData\Local\Microsoft\Windows\History\History.IE5\index.dat %USERPROFILE %\AppData\Local\Microsoft\Windows\History\Low\History.IE5\index.dat
History (periodic)	%USERPROFILE %\AppData\Local\Microsoft\Windows\History\History.IE5\MSHist01yyyymmddyyyymmdd\index.dat %USERPROFILE %\AppData\Local\Microsoft\Windows\History\Low\History.IE5\MSHist01yyyymmddyyyymmdd\index.dat
Cookies	%USERPROFILE %\AppData\Roaming\Microsoft\Windows\Cookies\index.dat %USERPROFILE %\AppData\Roaming\Microsoft\Windows\Cookies\Low\index.

Type of MSIECF file	Characteristics
	dat %USERPROFILE%\AppData\Local\Temp\Low\Cookies\index.dat
User data	%USERPROFILE%\AppData\Roaming\Microsoft\Internet Explorer\UserData\index.dat %USERPROFILE%\AppData\Roaming\Microsoft\Internet Explorer\UserData\Low\index.dat
InPrivate Filtering	%USERPROFILE%\AppData\Roaming\Microsoft\Windows\PrivacIE\index.dat %USERPROFILE%\AppData\Roaming\Microsoft\Windows\PrivacIE\Low\index.dat
Compatibility Cache	%USERPROFILE%\AppData\Roaming\Microsoft\Windows\IECompatCache\index.dat %USERPROFILE%\AppData\Roaming\Microsoft\Windows\IECompatCache\Low\index.dat
TLD Cache	%USERPROFILE%\AppData\Roaming\Microsoft\Windows\IETldCache\index.dat %USERPROFILE%\AppData\Roaming\Microsoft\Windows\IETldCache\Low\index.dat
Feeds Cache	%USERPROFILE%\AppData\Local\Microsoft\Feeds Cache\index.dat
DOM store	%USERPROFILE%\AppData\Local\Microsoft\Internet Explorer\DOMStore\index.dat %USERPROFILE%\AppData\LocalLow\Microsoft\Internet Explorer\DOMStore\index.dat

1.2.8. MSIE 9 on Windows 7

Type of MSIECF file	Characteristics
Temporary Internet Files (Cache)	%USERPROFILE%\AppData\Local\Microsoft\Windows\Temporary Internet Files\Content.IE5\index.dat %USERPROFILE%\AppData\Local\Microsoft\Windows\Temporary Internet Files\Low\Content.IE5\index.dat
History (global)	%USERPROFILE%\AppData\Local\Microsoft\Windows\History\History.IE5\index.dat %USERPROFILE%\AppData\Local\Microsoft\Windows\History\Low\History.IE5\index.dat

Type of MSIECF file	Characteristics
History (periodic)	%USERPROFILE%\AppData\Local\Microsoft\Windows\History\History.IE5\MSHist01yyyymmddyyyymmdd\index.dat %USERPROFILE%\AppData\Local\Microsoft\Windows\History\Low\History.IE5\MSHist01yyyymmddyyyymmdd\index.dat
Cookies	%USERPROFILE%\AppData\Roaming\Microsoft\Windows\Cookies\index.dat %USERPROFILE%\AppData\Roaming\Microsoft\Windows\Cookies\Low\index.dat
User data	%USERPROFILE%\AppData\Roaming\Microsoft\Internet Explorer\UserData\index.dat %USERPROFILE%\AppData\Roaming\Microsoft\Internet Explorer\UserData\Low\index.dat
InPrivate Filtering	%USERPROFILE%\AppData\Roaming\Microsoft\Windows\PrivacIE\index.dat %USERPROFILE%\AppData\Roaming\Microsoft\Windows\PrivacIE\Low\index.dat
Compatibility Cache	%USERPROFILE%\AppData\Roaming\Microsoft\Windows\IECompatCache\index.dat %USERPROFILE%\AppData\Roaming\Microsoft\Windows\IECompatCache\Low\index.dat
TLD Cache	%USERPROFILE%\AppData\Roaming\Microsoft\Windows\IETldCache\index.dat %USERPROFILE%\AppData\Roaming\Microsoft\Windows\IETldCache\Low\index.dat
Feeds Cache	%USERPROFILE%\AppData\Local\Microsoft\Feeds Cache\index.dat
DOM store	%USERPROFILE%\AppData\Local\Microsoft\Internet Explorer\DOMStore\index.dat %USERPROFILE%\AppData\LocalLow\Microsoft\Internet Explorer\DOMStore\index.dat
Download history	%USERPROFILE%\AppData\Roaming\Microsoft\Windows\IEDownloadHistory\index.dat

1.3. Test version

The following version of programs were used to test the information within this document:

- MSIE 4, 5, 6, 7, 8, 9

2. File header

The file header is of 72 bytes of size and consists of:

offset	size	value	description
0	28	“Client\x20UrlCache\x20MMF\x20Ver\x20#.#\x00“	The signature and version string ASCII string with an end-of-string character The # characters contain the major and minor versions.
28	4		The file size
32	4		The first hash table record offset The file offset to the first part of the hash table This value always should be a multitude of 128 and greater equal 0x4000 or 0 if no hash table is available.
36	4		The total number of blocks
40	4		The number of allocated blocks
44	4	0	Unknown Empty value
48	4		The cache size (quota) limit of the container Contains the number of bytes
52	4	0	Unknown Empty value or 64-bit extension of previous value
56	4		The cache size of the container Contains the number of bytes
60	4	0	Unknown Empty value or 64-bit extension of previous value
64	4		The non-releasable cache size of the container. (The size of the cache container exempt from scavenging) Contains the number of bytes
68	4	0	Unknown Empty value or 64-bit extension of previous value

The container is the cache container e.g. the files in the corresponding cache directories.

Note: Can the signature and version string be set in the registry? Is it variable of length? Or is 28 bytes the maximum length?

2.1. The cache directory table

The file header is followed by the cache directory table.

The cache directory table is variable of size and consists of:

offset	size	value	description
72	4		Number of cache directory entries A maximum of 32 ?
76	...		Cache directory entries

The cache directory entry is 12 bytes of size and consists of:

offset	size	value	description
0	4		The number of cached files in the directory
4	8		Cache directory name ASCII string without an end-of_string character

Note that a cache directory can contain files besides the cached files.

2.2. Header data

The cache directory table is follow by the header data which is an array of 32 x 32-bit values.

offset	size	value	description
460 (0x1cc)	4		Unknown 0a 00 00 00
464	4		Unknown 02 00 00 00 04 00 00 00 9f 00 00 00
468	4		Unknown 04 00 00 00 c5 05 00 00 d2 00 00 00
472	8		Unknown Empty values
480	4		Unknown 08 40 00 00
484	...		
588	4		Unknown Empty values

2.3. The allocation bitmap

The allocation bitmap is situated at offset 592 (0x250). The allocation bitmap continues to offset 16384 (0x4000) but only the number of bytes necessary for the file size is used.

Every bit represents a block of 128 byte (0x80) starting at the (base) offset 0x4000. The bitmap is stored byte-wise where the LSB refers to the lowest offset, e.g.

```
base offset                   : 0x4000
first allocation bitmap byte  : 0xf0
unallocated range             : 0x4000 – 0x4200
allocated range               : 0x4200 - 0x4400
```

3. Hash table record

The hash table record consists of:
- the hash table header
- hash table entries

The hash table record is 4096 bytes of size.

The hash table header is 16 bytes of size and consist of:

offset	**size**	**value**	**description**
WININET FILEMAP_ENTRY			
0	4	“HASH”	The signature
4	4	32 (0x20)	The number of blocks in hash table The block size is 128 bytes 32 x 128 = 4096 This value includes the size of the hash table header The hash table entry data size is: 4096 – 12 = 4084
WININET LIST_FILEMAP_ENTRY			
8	4		Next hash table record offset The file offset to the next part of the hash table or 0 if this is the last part of the hash table
12	4		The sequence number 0 identifies the first hash table record

3.1. Hash table entry

The hash table entry (HASH_ITEM) is 8 bytes of size and consists of:

offset	**size**	**value**	**description**
0	4		Record hash
4	4		Record offset This value always should be a multitude

offset	size	value	description
			of 128 and greater equal 0x4000 If the record offset contains the same value as the record hash the value is unused

Hash table entries that contain the same record hash and offset seem to be unused.

3.2. Record hash

The record hash is 32-bits of size and consists of:

offset	size	value	description
0.0	5 bits		Record hash flags
0.5	1 bit		Unused
0.6	26 bits		Record hash value

The record hash can contain different values:

Value	Description
0x#######0	Valid URL record
0x00000001	Invalid URL record Some of the record offsets might be still valid
0x00000003	unknown record only seen as unused: with a record offset of 0x00000003
0x#######5	REDR record
0x0badf00d	Hash table entry is uninitialized
0xdeadbeef	Hash table entry is uninitialized (MSIE 8)

The uninitialized hash table entries can occur in all hash table records not only the last one. Often the hash tables are only filled to offset 0xe00 ?

The valid URL records in the hash table do not refer to all the allocated URL records in the file.

Maybe the record hash refers to a bucket of records not a single record.

3.3. Record hash flags

Value	Description
0x01	Entry is unused
0x02	Entry is locked
0x04	Entry is redirected (point to a REDR record)
0x08	Entry is part of group
0x10	Entry is part of a list of groups

3.4. *Hash algorithm*

TODO

4. URL record

The URL record represents a cached entry. [JONES03] refers to this record as the URL activity record.

4.1. URL record format version 4.7

The URL record version 4.7 (URL_FILEMAP_ENTRY) is variable of size and consists of:

offset	**size**	**value**	**description**
WININET FILEMAP_ENTRY			
0	4	“URL\x20”	The signature
4	4		The number of blocks in URL record The block size is 128 bytes
8	8		The secondary time value Filetime Contains 0 if not set See section: 4.5 Use of the filetime values
16	8		The primary time value Filetime Contains 0 if not set See section: 4.5 Use of the filetime values Can contain 0x7fffffff 0xffffffff
24	8		Expiration date and time Filetime Contains 0 if not set
32	4		Cached file size Contains the number of bytes
36	12		Unknown Empty values
48	4		Unknown Empty values
52	4		Unknown
56	4	104 (0x0068)	The location offset The value is relative to the start of the URL record or 0 if not set
60	1		Cache directory index A value of 0 represents the first cache directory
61	3		Unknown

offset	size	value	description
64	4		The filename offset The value is relative to the start of the URL record or 0 if not set
68	4		Cache entry flags
72	4		The data offset The value is relative to the start of the URL record or 0 if not set
76	4		The data size
80	4		Unknown Empty values
84	4		Last checked date and time (Last synchronization date and time) FAT date time Contains 0 if not set See section: 7 FAT date time
88	4		Number of hits
92	4		Unknown Empty values Use count used in memory?
96	4		Unknown Last cache synchronization date and time FAT date time Contains 0 if not set See section: 7 FAT date time
100	4		Unknown Empty values
URL record data variable of size			
(location offset) 104	...		The location ASCII string with an end-of-string character Special characters are URL encoded (4 byte aligned)
(filename offset)	...		The filename in cache directory ASCII string with an end-of-string character
(data offset)	(data size)		Data
...	...		Unknown Uninitialized

4.2. URL record format version 5.2

The URL record version 5.2 (IE5_URL_FILEMAP_ENTRY or IE6_URL_FILEMAP_ENTRY) is variable of size and consists of:

offset	**size**	**value**	**description**
WININET FILEMAP_ENTRY			
0	4	“URL\x20”	The signature
4	4		The number of blocks in URL record The block size is 128 bytes
8	8		The secondary time value Filetime Contains 0 if not set See section: 4.5 Use of the filetime values
16	8		The primary time value Filetime Contains 0 if not set See section: 4.5 Use of the filetime values
24	4		Expiration date and time FAT date time Contains 0 if not set See section: 7 FAT date time Can contain 0xffff 0xffff (seen in a Visited URL record with an about: URI) -1 => never
28	4		Unknown Empty values
32	4		Cached file size Contains the number of bytes
36	4		Upper part of a 64-bit cached file size Contains the number of bytes
40	4		Group or group list offset
44	4		The non-releasable time delta (exempt time delta) Contains the number of seconds Contains the number of seconds before the cached item may be released. The time delta is relative to the last access time. Often it contains the value 86400 (0x00015180) seconds or 24 hours.
48	4	96 (0x0060)	Unknown offset The value is relative to the start of the URL record
52	4	104 (0x0068)	The location offset The value is relative to the start of the URL record or 0 if not set
56	1		Cache directory index A value of 0 represents the first cache

offset	size	value	description
			directory 0xfe => special type (cookie/iecompat/iedownload)? 0xff => ? Note: A value of 0xFF could be a special flag. There is no associated file in the cache and the URL has the a *.cdf extension. Could be Channel Definition Files.
57	1		Unknown (synchronization count) 0x00 => 0x01 = > 0x02 => 0x03 =>
58	1		Format version 0x00 => IE5_URL_FILEMAP_ENTRY 0x10 => IE6_URL_FILEMAP_ENTRY
59	1		Copy of format version 0x00 => IE5_URL_FILEMAP_ENTRY 0x10 => IE6_URL_FILEMAP_ENTRY
60	4		The filename offset The value is relative to the start of the URL record or 0 if not set
64	4		Cache entry flags
68	4		The data offset The value is relative to the start of the URL record or 0 if not set
72	4		The data size
76	4		Unknown (file extension offset) Empty values
80	4		Last checked date and time (Last synchronization date and time) FAT date time Contains 0 if not set See section: 7 FAT date time
84	4		Number of hits (number of times the entry has been locked)
88	4		Unknown Empty values Use count used in memory? (level of lock nesting of the entry)
92	4		Unknown Last cache synchronization date and time

offset	size	value	description
			(entry creation time?) FAT date time Contains 0 if not set See section: 7 FAT date time
URL record data variable of size			
(unknown offset) 96	4		Unknown value (8 byte aligned)
100	4		Unknown Uninitialized
(location offset) 104	...		The location ASCII string with an end-of-string character Special characters are URL encoded (8 byte aligned)
(filename offset)	...		The filename in cache directory ASCII string with an end-of-string character
(data offset)	(data size)		Data
...	...		Unknown Uninitialized

4.3. Cache entry flags

The cache entry flags consist of the following values:

Value	Identifier	Description
0x00000001	NORMAL_CACHE_ENTRY	Normal cache entry; can be deleted to recover space for new entries.
0x00000002	STABLE_CACHE_ENTRY	
0x00000004	STICKY_CACHE_ENTRY	Sticky cache entry that is exempt from scavenging for the amount of time specified by release (exempt) delta. The default value set by the function CommitUrlCacheEntry is one day. Has extended flags (unknown value at offset 40): 0x00004008 ?
0x00000008	EDITED_CACHE_ENTRY	Cache entry file that has been edited externally. This cache entry type is exempt from scavenging. Set for iecompat: and TLD Cache entries
0x00000010	TRACK_OFFLINE_CACHE_ENTRY	Not currently implemented.
0x00000020	TRACK_ONLINE_CACHE_ENTRY	Not currently implemented.

Value	Identifier	Description
0x00000040		Is cached/cache-able? Not set if header contains “Pragma: no-cache” or other cache related headers are present
0x00001000		HTTP request method 0 => GET 1 => POST
0x00010000	SPARSE_CACHE_ENTRY	Partial response cache entry.
0x00020000	OCX_CACHE_ENTRY	OLE Control Extension (OCX) cache entry. OCX is a predecessor of ActiveX Set for PrivacIE:
0x00100000	COOKIE_CACHE_ENTRY	Cookie cache entry.
0x00200000	URLHISTORY_CACHE_ENTRY	Visited link cache entry.
0x00400000	PENDING_DELETE_CACHE_ENTRY	Cache entry is pending deletion.
0x10000000	INSTALLED_CACHE_ENTRY	Unknown
0x80000000	IDENTITY_CACHE_ENTRY	Unknown

4.4. URL record types

The URL record values have different meanings for different types of MSIECF files.

Type of MSIECF file	**Location:**
Temporary Internet Files (Cache)	<URI>
History (global)	Visited: <username>@<URI>
History (periodic)	:<date range>: <username>@<URI> Date range is formatted as: yyyymmddyyyymmdd What about Host: in the visited URI e.g. :2013011020130111: test@:Host: My Computer :2013010920130110: test@file:///C:/test.txt
Cookies	Cookie:<username>@<URI>
InPrivate Filtering	PrivacIE:<URI filter expression>
Compatibility Cache	iecompat:<filename>

Type of MSIECF file	Location:
TLD Cache	ietld:<filename>
Feeds Cache	feedplat:<URI>
User data	userdata:<username>@<protocol>@<URI>
DOM store	DOMStore:<URI>
Download history	iedownload:<GUID>

Note both History (global) and History (periodic) set URLHISTORY_CACHE_ENTRY but global sets STICKY_CACHE_ENTRY and periodic sets STICKY_CACHE_ENTRY.

According to [BUNTING] the History (global) URL record type contains:

```
<username>@<URL>
```

This has not been seen in MSIE 4 - 9 cache files.

4.4.1. Cache URL record data

The cache URL record contains a string with headers of the HTTP response.

```
flags: 0x00000001
HTTP/1.0 200 OK^M
Content-Type: image/gif^M
Pragma: no-cache^M
Content-Length: 43^M
^M
~U:username^M

flags: 0x00000005
HTTP/1.0 200 OK^M
ETag: "13e-411e677a07f80"^M
Content-Length: 318^M
Content-Type: image/x-icon^M
X-Cache: MISS from sq25.wikimedia.org^M
X-Cache-Lookup: HIT from sq25.wikimedia.org:3128^M
X-Cache: MISS from knsq26.knams.wikimedia.org^M
X-Cache-Lookup: HIT from knsq26.knams.wikimedia.org:3128^M
X-Cache: HIT from knsq3.knams.wikimedia.org^M
X-Cache-Lookup: HIT from knsq3.knams.wikimedia.org:80^M
^M
~U:username^M

flags: 0x00000041
HTTP/1.1 200 OK^M
Content-Length: 1445^M
Content-Type: image/gif^M
ETag: "096398e49cc81:bd5"^M
X-Powered-By: ASP.NET^M
^M
~U:username^M

flags: 0x00000045
HTTP/1.1 200 OK^M
```

```
Content-Length: 25214^M
Content-Type: image/x-icon^M
ETag: "931c9030e226c61:284"^M
X-UA-Compatible: IE=EmulateIE7^M
X-Powered-By: ASP.NET^M
^M
~U:username^M

flags: 0x00001001
HTTP/1.0 200 OK^M
P3P: CP="NOI DEVo TAIa OUR BUS"^M
X-Function: 101^M
Pragma: no-cache^M
Content-Type: application/x-javascript^M
Content-Length: 209^M
^M
~U:username^M

HTTP/1.0 200 OK^M
P3P: CP="NON NID PSAa PSDa OUR IND UNI COM NAV STA",policyref="/w3c/p3p.xml"^M
P3P: CP="NON NID PSAa PSDa OUR IND UNI COM NAV STA",policyref="/w3c/p3p.xml"^M
ETag: "c9e504-2b-428a378f"^M
Content-Length: 43^M
Content-Type: image/gif^M
^M
~U:username^M

flags: 0x00001041
HTTP/1.1 200 OK^M
Content-Type: text/html; charset=UTF-8^M
Transfer-Encoding: chunked^M
^M
~U:username^M

HTTP/1.1 200 OK^M
Content-Type: text/html; charset=utf-8^M
P3P: CP="ALL IND DSP COR ADM CONo CUR CUSo IVAo IVDo PSA PSD TAI TELo OUR SAMo
CNT COM INT NAV ONL PHY PRE PUR UNI"^M
X-Powered-By: ASP.NET^M
X-UA-Compatible: IE=EmulateIE7^M
X-AspNet-Version: 2.0.50727^M
Transfer-Encoding: chunked^M
^M
~U:username^M
```

4.4.2. Visited URL record data

The visited URL record contains information which user visited what URI.

The URL record location consists of the following string

```
Visited: <username>@<URI>
```

If set the URL record data contains multiple entries in the following format:

offset	size	value	description
0	2		The entry size

offset	size	value	description
2	1		The entry type
3	1		The value type
4	(entry size - 4)		Value data

The last entry is an empty entry consisting of 4 zero-bytes.

Entry type	Value type	Identifier	Description
0x02	0x00		Unknown
0x0e	0x1e		Unknown A GUID formatted as a string {000000-0000-0000-0000-00000000} with NUL character (5 trailing empty bytes)
0x10	0x1f		Page title with NUL characters (4 trailing empty bytes)
0x11	0x01		Filenames Special characters are URL encoded (4 trailing empty bytes)
0x14	0x03		Unknown (4 trailing empty bytes)
0x15	0x1e		HTTP URI of favicon with NUL character (4 trailing empty bytes)
0x16	0x1f		File URI with NUL characters Special characters are URL encoded (4 trailing empty bytes)
0x17	0x13		Unknown (4 trailing empty bytes)
0x18	0x40		Unknown Contains a filetime (4 trailing empty bytes)
0x1c	0x03		Unknown (4 trailing empty bytes)
0x1e	0x40		Unknown Contains a filetime

Entry type	Value type	Identifier	Description
			(4 trailing empty bytes)
0x20	0x03		Unknown (4 trailing empty bytes)

The value types are similar to the values used by the OLE variant types (VT) and MAPI data (property) types (PT):

Value type	Identifier	Description
0x00	VT_EMPTY	Empty
0x01		Multi value UTF-16 string?
0x03	VT_I4	Integer 32-bit signed
0x1e	VT_LPSTR	Extended ASCII string NUL terminated
0x1f	VT_LPWSTR	Unicode string (UTF-16 little-endian) NUL terminated

What about the first entry it has an empty value type but contains values.
Perhaps it's some kind of header? It is present in every validation data.
first 32-bit value contains:
00 00 00 00 00 00 00 00 00 00 00 00 (0 or 1 entries)
00 00 00 10 00 00 00 00 00 00 00 00 (0 or 1 entries)
00 00 00 10 00 00 00 00 01 00 00 00 (multiple entries)
00 00 00 10 00 00 00 00 03 00 00 00 (multiple entries)

4.4.3. InPrivate Filtering URL record data

data does not change between items

```
libmsiecf_url_read: data:
00000000: f1 ff 00 00 01 00 00 00  04 00 00 00 81 41 33 21   ........ .....A3!
00000010: 00 00 00 00
```

4.4.4. Compatibility URL record data

data does not change between items

```
libmsiecf_url_read: data:
00000000: 58 49 71 17 00 00 08 00  00 00 00 00               XIq..... ....
```

4.4.5. TLD URL record data

data does not change between items

```
libmsiecf_url_read: data:
```

```
00000000: 01 00 00 00 01 00 00 00  00 00 00 00                ........ ....
```

4.4.6. Download URL record data

If set the URL record data contains data in the following format:

offset	size	value	description
0	4	0x00000085	Unknown Value does not change, maybe version indicator
4	4		Download status 0x00000001 => in progress? 0x00000003 => paused 0x00000006 => interrupted 0x0000000b => completed
8	8		Unknown (Empty values)
16	4		Unknown sometimes 0
20	16		GUID Should match the GUID in the location string or 0 if not set e.g. in canceled download
36	8		Download start time Filetime
44	4		Unknown (Empty values)
48	8		Unknown
56	4		Unknown
60	4		Unknown
64	4		Unknown
68	4		Unknown
72	8		Total download size Value in bytes
80	8		If status is in progress Number of bytes downloaded?
88	8		Unknown
96	8		Unknown (Empty values)
104	8		Unknown Set to 1 if string array contains a company/organization name?
112	8		Unknown 0x06
120	4		Flags 0x01 => complete download (otherwise

offset	size	value	description
			partial) 0x02 => unknown (related to signing info) 0x04 => unknown (related to signing info) 0x08 => string array contains a company/organization name 0x10 => string array contains originating website URL 0x40 => unknown (related to signing info)
124	16		Unknown (Empty values)
140	4		Unknown Set to 1 if string array contains a company/organization name?
144	2		Unknown 2 => HTTP/HTTPS
146	2		Unknown 0x5000 HTTP
148	4		Unknown (hash or checksum?) not set for ftp
152	152		Unknown (Empty values)
304	8		Unknown 0 most of the time also seen 1
312	...		Array of strings

Does the Last cache synchronization date and time of the URL contain the download time or is this just a common side effect of original purpose of the date and time value?

```
iedownload:{7EAE5A0A-00F9-11E2-8E4F-705AB642E02F}

GUID related to GUID in URL location?
FILETIME (same as primary time? Sep 17, 2012 18:57:26.719662000)
downloaded file size
00000000: 85 00 00 00 0b 00 00 00  00 00 00 00 00 00 00 00   ........ ........
00000010: e9 fd 00 00 45 88 5d 33  f9 00 e2 11 8e 4f 70 5a   ....E.]3 .....OpZ
00000020: b6 42 e0 2f cc 7e 99 47  06 95 cd 01 00 00 00 00   .B./.~.G ........

00000030: 91 01 00 00 00 00 00 00  01 00 00 00 01 00 00 00   ........ ........
00000040: 00 00 00 00 01 00 00 00  28 cc 04 01 00 00 00 00   ........ (.......

00000050: 59 b5 00 00 00 00 00 00  01 00 00 00 00 00 00 00   Y....... ........
00000060: 00 00 00 00 00 00 00 00  01 00 00 00 00 00 00 00   ........ ........
00000070: 06 00 00 00 00 00 00 00  19 00 00 00 00 00 00 00   ........ ........
00000080: 00 00 00 00 00 00 00 00  00 00 00 00 01 00 00 00   ........ ........
00000090: 02 00 00 50 17 40 0f 1a                            ...P.@.. ........

00000090:                          00 00 00 00 00 00 00 00   ...P.@.. ........
000000a0: 00 00 00 00 00 00 00 00  00 00 00 00 00 00 00 00   ........ ........
```

```
000000b0: 00 00 00 00 00 00 00 00  00 00 00 00 00 00 00 00   ........ ........
000000c0: 00 00 00 00 00 00 00 00  00 00 00 00 00 00 00 00   ........ ........
000000d0: 00 00 00 00 00 00 00 00  00 00 00 00 00 00 00 00   ........ ........
000000e0: 00 00 00 00 00 00 00 00  00 00 00 00 00 00 00 00   ........ ........
000000f0: 00 00 00 00 00 00 00 00  00 00 00 00 00 00 00 00   ........ ........
00000100: 00 00 00 00 00 00 00 00  00 00 00 00 00 00 00 00   ........ ........
00000110: 00 00 00 00 00 00 00 00  00 00 00 00 00 00 00 00   ........ ........
00000120: 00 00 00 00 00 00 00 00  00 00 00 00 00 00 00 00   ........ ........
00000130: 00 00 00 00 00 00 00 00                            ........

Company/Organization name
00000130:                          4d 00 69 00 63 00 72 00            M.i.c.r.
00000140: 6f 00 73 00 6f 00 66 00  74 00 20 00 43 00 6f 00   o.s.o.f. t. .C.o.
00000150: 72 00 70 00 6f 00 72 00  61 00 74 00 69 00 6f 00   r.p.o.r. a.t.i.o.
00000160: 6e 00 00 00                                        n...

URL download originating webpage
00000160:             68 00 74 00  74 00 70 00 3a 00 2f 00       h.t. t.p.:./.
00000170: 2f 00 77 00 77 00 77 00  2e 00 6d 00 69 00 63 00   /.w.w.w. ..m.i.c.
00000180: 72 00 6f 00 73 00 6f 00  66 00 74 00 2e 00 63 00   r.o.s.o. f.t...c.
...
000001c0: 77 00 6e 00 6c 00 6f 00  61 00 64 00 2e 00 61 00   w.n.l.o. a.d...a.
000001d0: 73 00 70 00 78 00 00 00                            s.p.x...

URL download
000001d0:                          68 00 74 00 74 00 70 00            h.t.t.p.
000001e0: 3a 00 2f 00 2f 00 64 00  6f 00 77 00 6e 00 6c 00   :././.d. o.w.n.l.
...
000002c0: 2e 00 6d 00 73 00 75 00  00 00                     ..m.s.u. ..

Destination filename
000002c0:                                43 00 3a 00 5c 00              C.:.\.
..
00000340: 68 00 50 00 6b 00 67 00  2e 00 6d 00 73 00 75 00   h.P.k.g. ..m.s.u.
00000350: 00 00                                              ..
```

4.5. Use of the filetime values

The filetime values in the URL record have different meanings for different types of MSIECF files.

Type of MSIECF file	**Primary time value**	**Secondary time value**
Temporary Internet Files (Cache)	Client last accessed date and time in UTC	Server modification date and time in UTC
History (global)	Last visited date and time in UTC	Last visited date and time in UTC
History (weekly)	History creation date and time in UTC Contains the creation date and time of the MSIECF index.dat file	Last visited date and time in local timezone
History (daily)	Last visited date and time in UTC	Last visited date and time in local timezone
Cookies	Cookie last accessed date and time in UTC	Cookie modification date and time in UTC
InPrivate Filtering	UTC	emtpy
Compatibility Cache	UTC	emtpy

Type of MSIECF file	**Primary time value**	**Secondary time value**
TLD Cache	UTC	emtpy
Feeds Cache	UTC	emtpy
User data	UTC	Emtpy
DOM Store	UTC	Emtpy
Download history	Downloaded file creation time in UTC (download start time)	Emtpy

Note on Windows 7 the MA.B times of the downloaded file are the same as the downloaded file creation time.

5. Redirected record

The redirected record represents a redirected entry. [JONES03] refers to this record as the REDR activity record.

The redirected record is variable of size and consist of:

offset	**size**	**value**	**description**
WININET FILEMAP_ENTRY			
0	4	“REDR”	The signature
4	4		The number of blocks in redirected record The block size is 128 bytes
8	4		Target hash table entry offset The offset is relative from the start of the file (However it does not seem to be the corresponding URL record, perhaps run-time remnant data)
12	4		Target URL hash value
16	...		The location ASCII string with an end-of-string character

6. Leak record

The leak record represents a cached URL record that is not longer consider as a valid part of the cache but that was not removed. See [CHAPPELL10] for a more detailed description.

The leak record is variable of size and consist of:

offset	**size**	**value**	**description**
WININET FILEMAP_ENTRY			
0	4	“LEAK”	The signature

offset	size	value	description
4	4		The number of blocks in URL record The block size is 128 bytes
8	24		Unknown Uninitialized
32	4		Cached file size in bytes
36	8		Unknown Uninitialized
44	4		Next leak record offset The offset is relative from the start of the file
48	8		Unknown Uninitialized
56	1		Cache directory index A value of 0 represents the first cache directory
57	3		Unknown Uninitialized
60	4		The filename offset The value is relative to the start of the URL record or 0 if not set
64	24		Unknown Uninitialized
88	4		Unknown Empty values
92	4		Unknown Last cache synchronization date and time FAT date time Contains 0 if not set See section: 7 FAT date time
LEAK record data variable of size			
96	8		Unknown Uninitialized
(filename offset) 104	...		The filename in cache directory ASCII string with an end-of-string character (8 byte aligned)
...	...		Unknown Uninitialized

7. FAT date time

The FAT date time consists of 4 bytes:

offset	size	value	description
0	2		date
2	2		time

In little-endian the 16-bit date value corresponds to:

offset	size	value	description
Bit 0 (LSB)	5 bits		Day of the month
Bit 5	4 bits		Month 0x01 => January
Bit 9	7 bits		Year 0x00 => 1980

In little-endian the 16-bit time value corresponds to:

offset	size	value	description
Bit 0 (LSB)	5 bits		Seconds in 2 second intervals
Bit 5	6 bits		Minutes
Bit 11	5 bits		Hours

8. Notes

```
INTERNET_CACHE_ENTRY_INFO Structure defines 4 time stamps
  FILETIME LastModifiedTime;
  FILETIME ExpireTime;
  FILETIME LastAccessTime;
  FILETIME LastSyncTime;
```

```
typedef struct _INTERNET_CACHE_TIMESTAMPS {
  FILETIME ftExpires;
  FILETIME ftLastModified;
}INTERNET_CACHE_TIMESTAMPS, *LPINTERNET_CACHE_TIMESTAMPS;
```

8.1. Cache entry control flags

Value	Identifier	Description
0x00000004	CACHE_ENTRY_ATTRIBUTE_FC	Sets the cache entry type.
0x00000010	CACHE_ENTRY_HITRATE_FC	Sets the hit rate.

Value	Identifier	Description
0x00000040	CACHE_ENTRY_MODTIME_FC	Sets the last modified time.
0x00000080	CACHE_ENTRY_EXPTIME_FC	Sets the expire time.
0x00000100	CACHE_ENTRY_ACCTIME_FC	Sets the last access time.
0x00000200	CACHE_ENTRY_SYNCTIME_FC	Sets the last sync time.
0x00000400	CACHE_ENTRY_HEADERINFO_FC	Not currently implemented.
0x00000800	CACHE_ENTRY_EXEMPT_DELTA_FC	Sets the exempt delta.

8.2. Cookie state flags

The InternetCookieState enumeration defines the state of the cookie.
Syntax

```
typedef enum  {
  COOKIE_STATE_UNKNOWN     = 0x0,
  COOKIE_STATE_ACCEPT      = 0x1,
  COOKIE_STATE_PROMPT      = 0x2,
  COOKIE_STATE_LEASH       = 0x3,
  COOKIE_STATE_DOWNGRADE   = 0x4,
  COOKIE_STATE_REJECT      = 0x5,
  COOKIE_STATE_MAX         = COOKIE_STATE_REJECT
} InternetCookieState;
```

Constants

COOKIE_STATE_UNKNOWN

Reserved.

COOKIE_STATE_ACCEPT

The cookies are accepted.

COOKIE_STATE_PROMPT

The user is prompted to accept or deny the cookie.

COOKIE_STATE_LEASH

Cookies are accepted only in the first-party context.

COOKIE_STATE_DOWNGRADE

Cookies are accepted and become session cookies.

COOKIE_STATE_REJECT

The cookies are rejected.

COOKIE_STATE_MAX

Same as COOKIE_STATE_REJECT.

8.3. Cache entry flags

res://

- location
- filename
- flags: 0x00000001

http://

- location
- filename
- data (string)
- flags: 0x00000001
- flags: 0x00000005 (Has unknown value at offset 40)
- flags: 0x00000041
- flags: 0x00000045 (Has unknown value at offset 40)
- flags: 0x00001001
- flags: 0x00001041

Cookie:

- location
- filename
- flags: 0x00100001

Visited:

- location
- data (binary)
- flags: 0x00200000
- flags: 0x00200001

PrivacIE:

- location
- data (binary)
- flags: 0x00020004 (Has unknown value at offset 40)

iecompat:

- location
- data (binary)
- flags: 0x00000009

ietld:

- location
- data (binary)
- flags: 0x00000009

feedplat:

- location

- filename
- data (string)
- flags: 0x00000001

userdata:
- location
- filename
- flags: 0x00000001

iedownload:
- location
- filename
- data (binary)
- flags: 0x00000009

8.4. Header data

offset	**size**	**value**	**description**
460	4		number of changes to any of many WININET settings (CACHE_HEADER_DATA_CURRENT_SETTINGS_VERSION)
464	4		number of changes to container list for same registry set (CACHE_HEADER_DATA_CONLIST_CHANGE_COUNT)
468	4		number of changes to Cookies container (CACHE_HEADER_DATA_COOKIE_CHANGE_COUNT)
472	4		window handle for cache notifications (CACHE_HEADER_DATA_NOTIFICATION_HWND)
476	4		window message for cache notifications (CACHE_HEADER_DATA_NOTIFICATION_MESG)
480	4		file offset of first GROUP_ENTRY, else zero (CACHE_HEADER_DATA_ROOTGROUP_OFFSET)
484	4		low 32 bits for generation of most recently allocated GROUPID, else zero (CACHE_HEADER_DATA_GID_LOW)
488	4		high 32 bits for generation of most recently allocated GROUPID, else zero (CACHE_HEADER_DATA_GID_HIGH

offset	size	value	description
			)
492	4		
496	4		
500	4		
504	4		
508	4		
512	4		
516	4		CACHE_HEADER_DATA_SSL_STATE_COUNT
520	4		
524			

0x15	CACHE_HEADER_DATA_NOTIFICATION_FILTER	bit flags to filter cache notifications
0x16	CACHE_HEADER_DATA_ROOT_LEAK_OFFSET	file offset of first leak entry
0x1B	CACHE_HEADER_DATA_ROOT_GROUPLIST_OFFSET	file offset of first GROUP_LIST_ENTRY, else zero

8.5. Groups and group list

URL record group list or group offset
set for http:// with flags 0x00000004 set. The value does not change for different items (0x00004008)
Although the flag is set in some host/file it contains 0x00000000

offset	size	value	description
0	8		Group identifier 0 if group entry is not used, -1 if group entry is an index entry
If group identifier == -1			
8	4		Group entry offset The offset is relative to the start of the file or 0 if not available
Else if group identifier != 0			
8	4		Group flags
Common			
12	4		Group type
16	8		Disk usage

offset	size	value	description
			Value in bytes
24	8		Disk quota Value in 1024 bytes (KiB)
If group identifier == -1			
8	4		First available (free) group entry offset The offset is relative to the start of the file or 0 if not available
Else if group identifier != 0			
8	4		Group data offset The offset is relative to the start of the file or 0 if not available
Common			
32	8		Unknown Empty values

Value	Identifier	Description
0x01	CACHEGROUP_FLAG_NONPURGEABLE	
0x02	CACHEGROUP_FLAG_FLUSHURL_ONDELETE	

Offset	Size	Description
0x00	GROUPNAME_MAX_LENGTH bytes	group name
0x78	GROUP_OWNER_STORAGE_SIZE dwords	owner storage
0x88	dword	in allocated entry: zero in free entry: file offset of next free GROUP_DATA_ENTRY, else zero

8.6. Hash algorithm

Define the hash pad table as:

```
uint8_t hash_pad_table[ 256 ] = {
        0x01, 0x0e, 0x6e, 0x19, 0x61, 0xae, 0x84, 0x77,
        0x8a, 0xaa, 0x7d, 0x76, 0x1b, 0xe9, 0x8c, 0x33,
        0x57, 0xc5, 0xb1, 0x6b, 0xea, 0xa9, 0x38, 0x44,
        0x1e, 0x07, 0xad, 0x49, 0xbc, 0x28, 0x24, 0x41,
        0x31, 0xd5, 0x68, 0xbe, 0x39, 0xd3, 0x94, 0xdf,
        0x30, 0x73, 0x0f, 0x02, 0x43, 0xba, 0xd2, 0x1c,
        0x0c, 0xb5, 0x67, 0x46, 0x16, 0x3a, 0x4b, 0x4e,
        0xb7, 0xa7, 0xee, 0x9d, 0x7c, 0x93, 0xac, 0x90,
        0xb0, 0xa1, 0x8d, 0x56, 0x3c, 0x42, 0x80, 0x53,
        0x9c, 0xf1, 0x4f, 0x2e, 0xa8, 0xc6, 0x29, 0xfe,
        0xb2, 0x55, 0xfd, 0xed, 0xfa, 0x9a, 0x85, 0x58,
        0x23, 0xce, 0x5f, 0x74, 0xfc, 0xc0, 0x36, 0xdd,
        0x66, 0xda, 0xff, 0xf0, 0x52, 0x6a, 0x9e, 0xc9,
        0x3d, 0x03, 0x59, 0x09, 0x2a, 0x9b, 0x9f, 0x5d,
        0xa6, 0x50, 0x32, 0x22, 0xaf, 0xc3, 0x64, 0x63,
        0x1a, 0x96, 0x10, 0x91, 0x04, 0x21, 0x08, 0xbd,
        0x79, 0x40, 0x4d, 0x48, 0xd0, 0xf5, 0x82, 0x7a,
        0x8f, 0x37, 0x69, 0x86, 0x1d, 0xa4, 0xb9, 0xc2,
        0xc1, 0xef, 0x65, 0xf2, 0x05, 0xab, 0x7e, 0x0b,
        0x4a, 0x3b, 0x89, 0xe4, 0x6c, 0xbf, 0xe8, 0x8b,
        0x06, 0x18, 0x51, 0x14, 0x7f, 0x11, 0x5b, 0x5c,
        0xfb, 0x97, 0xe1, 0xcf, 0x15, 0x62, 0x71, 0x70,
        0x54, 0xe2, 0x12, 0xd6, 0xc7, 0xbb, 0x0d, 0x20,
        0x5e, 0xdc, 0xe0, 0xd4, 0xf7, 0xcc, 0xc4, 0x2b,
        0xf9, 0xec, 0x2d, 0xf4, 0x6f, 0xb6, 0x99, 0x88,
        0x81, 0x5a, 0xd9, 0xca, 0x13, 0xa5, 0xe7, 0x47,
        0xe6, 0x8e, 0x60, 0xe3, 0x3e, 0xb3, 0xf6, 0x72,
        0xa2, 0x35, 0xa0, 0xd7, 0xcd, 0xb4, 0x2f, 0x6d,
        0x2c, 0x26, 0x1f, 0x95, 0x87, 0x00, 0xd8, 0x34,
        0x3f, 0x17, 0x25, 0x45, 0x27, 0x75, 0x92, 0xb8,
        0xa3, 0xc8, 0xde, 0xeb, 0xf8, 0xf3, 0xdb, 0x0a,
        0x98, 0x83, 0x7b, 0xe5, 0xcb, 0x4c, 0x78, 0xd1
}
```

What do the first 4 bytes contain?

```
uint8_t hash_data[ 4 ];

hash_data[ 0 ] = hash_pad_table[ url_string[ 0 ] ];
hash_data[ 1 ] = hash_pad_table[ url_string[ 1 ] ];
hash_data[ 2 ] = hash_pad_table[ url_string[ 2 ] ];
hash_data[ 3 ] = hash_pad_table[ url_string[ 3 ] ];

for( string_index = 1;
     string_index < string_lenght;
     string_index++ )
{
          if( url_string[ string_index ] == 0 )
          {
                    break;
          }
          if( ( url_string[ string_index ] == (uint8_t) '/' )
           && ( url_string[ string_index + 1 ] == 0 ) )
          {
                    break;
          }
          hash_data[ 0 ] ^= url_string[ 0 ];
          hash_data[ 1 ] ^= url_string[ 1 ];
          hash_data[ 2 ] ^= url_string[ 2 ];
          hash_data[ 3 ] ^= url_string[ 3 ];


          hash_data[ 0 ] = hash_pad_table[ hash_data[ 0 ] ];
          hash_data[ 1 ] = hash_pad_table[ hash_data[ 1 ] ];
          hash_data[ 2 ] = hash_pad_table[ hash_data[ 2 ] ];
          hash_data[ 3 ] = hash_pad_table[ hash_data[ 3 ] ];
}

hash_value   = hash_data[ 3 ];
hash_value <<= 8;
hash_value  |= hash_data[ 2 ];
hash_value <<= 8;
hash_value  |= hash_data[ 1 ];
hash_value <<= 8;
hash_value  |= hash_data[ 0 ];
```

Appendix A. References

[JONES03]
Title: Forensic Analysis of Internet Explorer Activity Files
Author(s): Keith J. Jones
URL: http://sourceforge.net/projects/odessa/

Title: Microsoft Windows Browser Cache system info
URL: http://www.conknet.com/~w_kranz/mswinbrz.txt

Title: Reverse Engineering Index.dat
URL: http://www.latenighthacking.com/projects/2003/reIndexDat/

[CHAPPELL10]
Title: The INDEX.DAT File Format
Author(s): Geoff Chappell
URL: http://www.geoffchappell.com/studies/windows/ie/wininet/api/urlcache/hashkey.htm?tx=20,78,83,84,88

Title: The Hash Algorithm for URL Caching
Author(s): Geoff Chappell
URL: http://www.geoffchappell.com/studies/windows/ie/wininet/api/urlcache/hashkey.htm?tx=20,78,83,84,88

[BUNTING]
Title: Understanding index.dat Files
Author(s): Captain Stephen M. Bunting
URL:
http://web.archive.org/web/20090605202325/http://128.175.24.251/forensics/index_dat1.htm
http://web.archive.org/web/20090605200839/http://128.175.24.251/forensics/index_dat2.htm

[MSDN]
Title: Microsoft Developer Network
URL: http://msdn.microsoft.com/

Appendix B. GNU Free Documentation License

Version 1.3, 3 November 2008

0. PREAMBLE

The purpose of this License is to make a manual, textbook, or other functional and useful document "free" in the sense of freedom: to assure everyone the effective freedom to copy and redistribute it, with or without modifying it, either commercially or noncommercially. Secondarily, this License preserves for the author and publisher a way to get credit for their work, while not being considered responsible for modifications made by others.

This License is a kind of "copyleft", which means that derivative works of the document must themselves be free in the same sense. It complements the GNU General Public License, which is a copyleft license designed for free software.

We have designed this License in order to use it for manuals for free software, because free software needs free documentation: a free program should come with manuals providing the same freedoms that the software does. But this License is not limited to software manuals; it can be used for any textual work, regardless of subject matter or whether it is published as a printed book. We recommend this License principally for works whose purpose is instruction or reference.

1. APPLICABILITY AND DEFINITIONS

This License applies to any manual or other work, in any medium, that contains a notice placed by the copyright holder saying it can be distributed under the terms of this License. Such a notice grants a world-wide, royalty-free license, unlimited in duration, to use that work under the conditions stated herein. The "Document", below, refers to any such manual or work. Any member of the public is a licensee, and is addressed as "you". You accept the license if you copy, modify or distribute the work in a way requiring permission under copyright law.

A "Modified Version" of the Document means any work containing the Document or a portion of it, either copied verbatim, or with modifications and/or translated into another language.

A "Secondary Section" is a named appendix or a front-matter section of the Document that deals exclusively with the relationship of the publishers or authors of the Document to the Document's overall subject (or to related matters) and contains nothing that could fall directly within that overall subject. (Thus, if the Document is in part a textbook of mathematics, a Secondary Section may not explain any mathematics.) The relationship could be a matter of historical connection with the subject or with related matters, or of legal, commercial, philosophical, ethical or political position regarding them.

The "Invariant Sections" are certain Secondary Sections whose titles are designated, as being those of Invariant Sections, in the notice that says that the Document is released under this License. If a section does not fit the above definition of Secondary then it is not allowed to be designated as Invariant. The Document may contain zero Invariant Sections. If the Document does not identify any Invariant Sections then there are none.

The "Cover Texts" are certain short passages of text that are listed, as Front-Cover Texts or Back-Cover Texts, in the notice that says that the Document is released under this License. A Front-Cover Text may be at most 5 words, and a Back-Cover Text may be at most 25 words.

A "Transparent" copy of the Document means a machine-readable copy, represented in a format whose specification is available to the general public, that is suitable for revising the document straightforwardly with generic text editors or (for images composed of pixels) generic paint programs or (for drawings) some widely available drawing editor, and that is suitable for input to text formatters or for automatic translation to a variety of formats suitable for input to text formatters. A copy made in an otherwise Transparent file format whose markup, or absence of markup, has been arranged to thwart or discourage subsequent modification by readers is not Transparent. An image format is not Transparent if used for any substantial amount of text. A copy that is not "Transparent" is called "Opaque".

Examples of suitable formats for Transparent copies include plain ASCII without markup, Texinfo input format, LaTeX input format, SGML or XML using a publicly available DTD, and standardconforming simple HTML, PostScript or PDF designed for human modification. Examples of transparent image formats include PNG, XCF and JPG. Opaque formats include proprietary formats that can be read and edited only by proprietary word processors, SGML or XML for which the DTD and/or processing tools are not generally available, and the machine-generated HTML, PostScript or PDF produced by some word processors for output purposes only.

The "Title Page" means, for a printed book, the title page itself, plus such following pages as are needed to hold, legibly, the material this License requires to appear in the title page. For works in formats which do not have any title page as such, "Title Page" means the text near the most prominent appearance of the work's title, preceding the beginning of the body of the text.

The "publisher" means any person or entity that distributes copies of the Document to the public.

A section "Entitled XYZ" means a named subunit of the Document whose title either is precisely XYZ or contains XYZ in parentheses following text that translates XYZ in another language. (Here XYZ stands for a specific section name mentioned below, such as "Acknowledgements", "Dedications", "Endorsements", or "History".) To "Preserve the Title" of such a section when you modify the Document means that it remains a section "Entitled XYZ" according to this definition.

The Document may include Warranty Disclaimers next to the notice which states that this License applies to the Document. These Warranty Disclaimers are considered to be included by reference in this License, but only as regards disclaiming warranties: any other implication that these Warranty Disclaimers may have is void and has no effect on the meaning of this License.

2. VERBATIM COPYING

You may copy and distribute the Document in any medium, either commercially or noncommercially, provided that this License, the copyright notices, and the license notice saying this License applies to the Document are reproduced in all copies, and that you add no other conditions whatsoever to those of this License. You may not use technical measures to obstruct or control the reading or further copying of the copies you make or distribute. However, you may accept compensation in exchange for copies. If you distribute a large enough number of copies you must also follow the conditions in section 3.

You may also lend copies, under the same conditions stated above, and you may publicly display copies.

3. COPYING IN QUANTITY

If you publish printed copies (or copies in media that commonly have printed covers) of the Document, numbering more than 100, and the Document's license notice requires Cover Texts, you must enclose the copies in covers that carry, clearly and legibly, all these Cover Texts: Front-Cover Texts on the front cover, and Back-Cover Texts on the back cover. Both covers must also clearly and legibly identify you as the publisher of these copies. The front cover must present the full title with all words of the title equally prominent and visible. You may add other material on the covers in addition. Copying with changes limited to the covers, as long as they preserve the title of the Document and satisfy these conditions, can be treated as verbatim copying in other respects.

If the required texts for either cover are too voluminous to fit legibly, you should put the first ones listed (as many as fit reasonably) on the actual cover, and continue the rest onto adjacent pages.

If you publish or distribute Opaque copies of the Document numbering more than 100, you must either include a machine-readable Transparent copy along with each Opaque copy, or state in or with each Opaque copy a computer-network location from which the general network-using public has access to download using public-standard network protocols a complete Transparent copy of the Document, free of added material. If you use the latter option, you must take reasonably prudent steps, when you begin distribution of Opaque copies in quantity, to ensure that this Transparent copy will remain thus accessible at the stated location until at least one year after the last time you distribute an Opaque copy (directly or through your agents or retailers) of that edition to the public.

It is requested, but not required, that you contact the authors of the Document well before redistributing any large number of copies, to give them a chance to provide you with an updated version of the Document.

4. MODIFICATIONS

You may copy and distribute a Modified Version of the Document under the conditions of sections 2 and 3 above, provided that you release the Modified Version under precisely this License, with the Modified Version filling the role of the Document, thus licensing distribution and modification of the Modified Version to whoever possesses a copy of it. In addition, you must do these things in the Modified Version:

- A. Use in the Title Page (and on the covers, if any) a title distinct from that of the Document, and from those of previous versions (which should, if there were any, be listed in the History section of the Document). You may use the same title as a previous version if the original publisher of that version gives permission.
- B. List on the Title Page, as authors, one or more persons or entities responsible for authorship of the modifications in the Modified Version, together with at least five of the principal authors of the Document (all of its principal authors, if it has fewer than five), unless they release you from this requirement.
- C. State on the Title page the name of the publisher of the Modified Version, as the publisher.
- D. Preserve all the copyright notices of the Document.
- E. Add an appropriate copyright notice for your modifications adjacent to the other copyright notices.
- F. Include, immediately after the copyright notices, a license notice giving the public permission to use the Modified Version under the terms of this License, in the form shown in the Addendum below.
- G. Preserve in that license notice the full lists of Invariant Sections and required Cover Texts given in the Document's license notice.
- H. Include an unaltered copy of this License.
- I. Preserve the section Entitled "History", Preserve its Title, and add to it an item stating at least the title, year, new authors, and publisher of the Modified Version as given on the Title Page. If there is no section Entitled "History" in the Document, create one stating the title, year, authors, and publisher of the Document as given on its Title Page, then add an item describing the Modified Version as stated in the previous sentence.
- J. Preserve the network location, if any, given in the Document for public access to a Transparent copy of the Document, and likewise the network locations given in the Document for previous versions it was based on. These may be placed in the "History" section. You may omit a network location for a work that was published at least four years before the Document itself, or if the original publisher of the version it refers to gives permission.
- K. For any section Entitled "Acknowledgements" or "Dedications", Preserve the Title of the section, and preserve in the section all the substance and tone of each of the contributor acknowledgements and/or dedications given therein.
- L. Preserve all the Invariant Sections of the Document, unaltered in their text and in their titles. Section numbers or the equivalent are not considered part of the section titles.
- M. Delete any section Entitled "Endorsements". Such a section may not be included in the

Modified Version.

- N. Do not retitle any existing section to be Entitled "Endorsements" or to conflict in title with any Invariant Section.
- O. Preserve any Warranty Disclaimers.

If the Modified Version includes new front-matter sections or appendices that qualify as Secondary Sections and contain no material copied from the Document, you may at your option designate some or all of these sections as invariant. To do this, add their titles to the list of Invariant Sections in the Modified Version's license notice. These titles must be distinct from any other section titles.

You may add a section Entitled "Endorsements", provided it contains nothing but endorsements of your Modified Version by various parties—for example, statements of peer review or that the text has been approved by an organization as the authoritative definition of a standard.

You may add a passage of up to five words as a Front-Cover Text, and a passage of up to 25 words as a Back-Cover Text, to the end of the list of Cover Texts in the Modified Version. Only one passage of Front-Cover Text and one of Back-Cover Text may be added by (or through arrangements made by) any one entity. If the Document already includes a cover text for the same cover, previously added by you or by arrangement made by the same entity you are acting on behalf of, you may not add another; but you may replace the old one, on explicit permission from the previous publisher that added the old one.

The author(s) and publisher(s) of the Document do not by this License give permission to use their names for publicity for or to assert or imply endorsement of any Modified Version.

5. COMBINING DOCUMENTS

You may combine the Document with other documents released under this License, under the terms defined in section 4 above for modified versions, provided that you include in the combination all of the Invariant Sections of all of the original documents, unmodified, and list them all as Invariant Sections of your combined work in its license notice, and that you preserve all their Warranty Disclaimers.

The combined work need only contain one copy of this License, and multiple identical Invariant Sections may be replaced with a single copy. If there are multiple Invariant Sections with the same name but different contents, make the title of each such section unique by adding at the end of it, in parentheses, the name of the original author or publisher of that section if known, or else a unique number. Make the same adjustment to the section titles in the list of Invariant Sections in the license notice of the combined work.

In the combination, you must combine any sections Entitled "History" in the various original documents, forming one section Entitled "History"; likewise combine any sections Entitled "Acknowledgements", and any sections Entitled "Dedications". You must delete all sections Entitled "Endorsements".

6. COLLECTIONS OF DOCUMENTS

You may make a collection consisting of the Document and other documents released under this License, and replace the individual copies of this License in the various documents with a single copy that is included in the collection, provided that you follow the rules of this License for verbatim copying of each of the documents in all other respects.

You may extract a single document from such a collection, and distribute it individually under this License, provided you insert a copy of this License into the extracted document, and follow this

License in all other respects regarding verbatim copying of that document.

7. AGGREGATION WITH INDEPENDENT WORKS

A compilation of the Document or its derivatives with other separate and independent documents or works, in or on a volume of a storage or distribution medium, is called an "aggregate" if the copyright resulting from the compilation is not used to limit the legal rights of the compilation's users beyond what the individual works permit. When the Document is included in an aggregate, this License does not apply to the other works in the aggregate which are not themselves derivative works of the Document.

If the Cover Text requirement of section 3 is applicable to these copies of the Document, then if the Document is less than one half of the entire aggregate, the Document's Cover Texts may be placed on covers that bracket the Document within the aggregate, or the electronic equivalent of covers if the Document is in electronic form. Otherwise they must appear on printed covers that bracket the whole aggregate.

8. TRANSLATION

Translation is considered a kind of modification, so you may distribute translations of the Document under the terms of section 4. Replacing Invariant Sections with translations requires special permission from their copyright holders, but you may include translations of some or all Invariant Sections in addition to the original versions of these Invariant Sections. You may include a translation of this License, and all the license notices in the Document, and any Warranty Disclaimers, provided that you also include the original English version of this License and the original versions of those notices and disclaimers. In case of a disagreement between the translation and the original version of this License or a notice or disclaimer, the original version will prevail.

If a section in the Document is Entitled "Acknowledgements", "Dedications", or "History", the requirement (section 4) to Preserve its Title (section 1) will typically require changing the actual title.

9. TERMINATION

You may not copy, modify, sublicense, or distribute the Document except as expressly provided under this License. Any attempt otherwise to copy, modify, sublicense, or distribute it is void, and will automatically terminate your rights under this License.

However, if you cease all violation of this License, then your license from a particular copyright holder is reinstated (a) provisionally, unless and until the copyright holder explicitly and finally terminates your license, and (b) permanently, if the copyright holder fails to notify you of the violation by some reasonable means prior to 60 days after the cessation.

Moreover, your license from a particular copyright holder is reinstated permanently if the copyright holder notifies you of the violation by some reasonable means, this is the first time you have received notice of violation of this License (for any work) from that copyright holder, and you cure the violation prior to 30 days after your receipt of the notice.

Termination of your rights under this section does not terminate the licenses of parties who have received copies or rights from you under this License. If your rights have been terminated and not permanently reinstated, receipt of a copy of some or all of the same material does not give you any rights to use it.

10. FUTURE REVISIONS OF THIS LICENSE

The Free Software Foundation may publish new, revised versions of the GNU Free Documentation License from time to time. Such new versions will be similar in spirit to the present version, but may differ in detail to address new problems or concerns. See http://www.gnu.org/copyleft/.

Each version of the License is given a distinguishing version number. If the Document specifies that a particular numbered version of this License "or any later version" applies to it, you have the option of following the terms and conditions either of that specified version or of any later version that has been published (not as a draft) by the Free Software Foundation. If the Document does not specify a version number of this License, you may choose any version ever published (not as a draft) by the Free Software Foundation. If the Document specifies that a proxy can decide which future versions of this License can be used, that proxy's public statement of acceptance of a version permanently authorizes you to choose that version for the Document.

11. RELICENSING

"Massive Multiauthor Collaboration Site" (or "MMC Site") means any World Wide Web server that publishes copyrightable works and also provides prominent facilities for anybody to edit those works. A public wiki that anybody can edit is an example of such a server. A "Massive Multiauthor Collaboration" (or "MMC") contained in the site means any set of copyrightable works thus published on the MMC site.

"CC-BY-SA" means the Creative Commons Attribution-Share Alike 3.0 license published by Creative Commons Corporation, a not-for-profit corporation with a principal place of business in San Francisco, California, as well as future copyleft versions of that license published by that same organization.

"Incorporate" means to publish or republish a Document, in whole or in part, as part of another Document.

An MMC is "eligible for relicensing" if it is licensed under this License, and if all works that were first published under this License somewhere other than this MMC, and subsequently incorporated in whole or in part into the MMC, (1) had no cover texts or invariant sections, and (2) were thus incorporated prior to November 1, 2008.

The operator of an MMC Site may republish an MMC contained in the site under CC-BY-SA on the same site at any time before August 1, 2009, provided the MMC is eligible for relicensing.