

The Evolution of Chrome Databases

By Ryan Benson

Each "Chrome window" below represents an SQLite database file Chrome uses to store data. The first tab in each window is the name of the file in the 'Default' folder inside the 'Chrome' directory. Each of the next 30 tabs correspond to a version of Chrome. The dark gray bars are the names of tables inside the database. An example of actual data is in brackets after the column description.

The Chrome data folder is located:

WinXP: <userdir>\Local Settings\Application Data\Google\Chrome\User Data\Default\
Vista/7: <userdir>\AppData\Local\Google\Chrome\User Data\Default\
Linux: <userdir>/.config/google-chrome/Default/
OS X: <userdir>/Library/Application Support/Google/Chrome/Default/

Red items are data entered by a user (such as autofill items or URLs) or retrieved from a website (like cookies or page content).
Green items are time related. Different color shades correspond to different timestamp formats.
Blue items are counts, codes, or other Chrome-internal descriptive items an end user is not likely to see.
Yellow items show how Chrome links fields together internally (such as primary and foreign keys).

History X 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30

The History database tracks webpage visits that occurred in the past three months and all file downloads.

urls	
id (INTEGER)	Primary Key [1]
url (LONGVARCHAR)	Page URL [http://tools.google.com/chrome/intl/en-US/welcome.html]
title (LONGVARCHAR)	Page Title [Getting Started]
visit_count (INTEGER)	Total number of times this URL has been visited. [1]
typed_count (INTEGER)	Number of times this URL has been manually entered in the URL bar. [0]
hidden (INTEGER)	Hidden entries should not be shown in typical UI or queries; this is usually for subframes. Boolean (0 = no, 1 = yes)
last_visit_time (INTEGER)	WebKit Timestamp [1301488000671600]
visits	
id (INTEGER)	Primary Key [1]
url (INTEGER)	Foreign Key to urls.id [1]
from_visit (INTEGER)	visits.id of previous page [0]
visit_time (INTEGER)	WebKit Timestamp [13014489130705790]
visit_duration (INTEGER)	Number of Milliseconds [21815000]
transition (INTEGER)	Transition Code [805306374]
is_indexed (BOOLEAN)	If a page has been indexed. Boolean (0 = no, 1 = yes)
visit_source	
id (INTEGER)	Foreign Key to visits.id [1]
source (INTEGER)	(0 = Synced, 1 = Browsed, 2 = Added by Extension, 3 = Imported from Firefox, 4 = Imported from IE, 5 = Imported from Safari)
downloads	
id (INTEGER)	Primary Key [1]
url (LONGVARCHAR)	Download URL [http://upload.wikimedia.org/wikipedia/commons/7/7a/Hard_disk.jpg]
full_path (LONGVARCHAR)	Local Path [C:\Users\Ryan\Documents\Downloads\Hard_disk.jpg]
target_path (LONGVARCHAR)	C:\Users\Ryan\Downloads\Disk.jpg
current_path (LONGVARCHAR)	C:\Users\Ryan\Downloads\Disk.jpg
start_time (INTEGER)	Unix Timestamp [1370406510]
end_time (INTEGER)	Unix Timestamp [1372310048]
received_bytes (INTEGER)	Size in Bytes [1387537]
total_bytes (INTEGER)	Size in Bytes [1387537]
state (INTEGER)	0 = In Progress, 1 = Complete, 2 = Cancelled, 3 = Interrupted
opened (INTEGER)	If this download has ever been opened from the browser. Boolean (0 = no, 1 = yes)
danger_type (INTEGER)	Codes 1-8 (0 is safe)
interrupt_reason (INTEGER)	Codes 1-50 (0 if successful)
etag (VARCHAR)	42abd14e45e859fc176c509f1b9c032
last_modified (VARCHAR)	Tue, 17 Jul 2012 21:36:38 GMT
downloads_url_chains	
id (INTEGER)	Foreign Key to downloads.id. Not unique (all downloads in a chain will have same id) [1]
url (LONGVARCHAR)	http://upload.wikimedia.org/wikipedia/commons/7/7a/Hard_disk.jpg
chain_index (INTEGER)	Number of link in the download chain (starts at 0) [0]

Since Chrome's release in 2008, there have been **30 different versions** of the browser. Chrome started with a quarterly release cycle, but in 2010 began releasing a new stable version every six weeks.

Transition Codes:
The last (least significant) 8 bits describe how Chrome got to this page. The first (most significant) 24 bits represent qualifiers that further describe the transition[1]. Bitwise AND the transition value with 0xFF to get one of the following codes:

0 Link	6 Top Level / Start page
1 Typed	7 Form Submit
2 Bookmark	8 Reload / Restore
3 Auto Subframe	9 Keyword
4 Manual Subframe	10 Keyword Generated
5 Omnibar Generated	

Chrome **doesn't record visits that were browsed in Chrome** (code 1) in *visit_source.source*. This is to save space, as the vast majority of visits would be of this type[2].

Danger Type Codes:
0 = Not Dangerous
1 = Dangerous
2 = Dangerous URL
3 = Dangerous Content
4 = Content May Be Malicious
5 = Uncommon Content
6 = Dangerous Host User Validated
7 = Dangerous Host
8 = Potentially Unwanted [3]

The **Archived History** database has nearly the same schema as **History**, but tracks webpage activity older than 3 months. It lacks the download related tables, as the History database tracks all downloads, even older ones. This **same SQL query can be used on both the History and Archived History** files.

There is no set structure for the *cookies.value* column; websites can store whatever text they want. Some cookies are simple timestamps or true/false values. Other **cookies have internal structures**, like the Google Analytics cookie in the example.

The **History Index** files were removed completely from Chrome in v30. As part of the upgrade to v30, Chrome deleted all existing History Index files[4].

As of Chrome v30, there are **16 autofill related tables** in Web Data. The two listed here are typically the most useful; the other are omitted for brevity's sake.

SQL for basic information on:
URLs: SELECT urls.id, urls.url, urls.title, urls.visit_count, urls.typed_count, urls.last_visit_time, visits.visit_time, visits.from_visit, visits.transition FROM urls, visits WHERE urls.id = visits.url
Downloads (v1 - 25): SELECT downloads.id, downloads.url, downloads.full_path, downloads.start_time, downloads.received_bytes, downloads.total_bytes, downloads.state FROM downloads
Downloads (v26 - 30): SELECT downloads.id, downloads_url_chains.url, downloads.target_path, downloads.start_time, downloads.end_time, downloads.received_bytes, downloads.total_bytes, downloads.state, downloads.opened, downloads.danger_type, downloads.interrupt_reason FROM downloads,downloads_url_chains WHERE downloads.id = downloads_url_chains.id

Cookies X 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30

The Cookies database tracks all HTTP cookies and their metadata in one location.

cookies	
host_key (TEXT)	Domain the cookie is for [.google.com]
path (TEXT)	Path the cookie is valid for [/chrome/intl/en-US/]
name (TEXT)	Cookie Name [utma]
value (TEXT)	Cookie Value [21104190.1290377463.1370406401.1370406401.1370406401.1]
last_access_utc (INTEGER)	WebKit Timestamp [13015988978767200]
creation_utc (INTEGER)	WebKit Timestamp [12998806882322090]
expires_utc (INTEGER)	WebKit Timestamp [13061878882000000]
secure (INTEGER)	Only send cookie over SSL. Boolean (0 = no, 1 = yes)
httponly (INTEGER)	If the cookie is HttpOnly (and not accessible by the client). Boolean (0 = no, 1 = yes)
persistent (INTEGER)	Whether the cookie is a session cookie [0] or a persistent cookie [1]
has_expires (INTEGER)	If the cookie has an expiration time (session cookies will not). Boolean (0 = no, 1 = yes)
priority (INTEGER)	Used to decide which cookies to purge and which to keep (0 = low, 1 = medium (default), 2 = high)

SQL for basic cookie information: SELECT cookies.creation_utc, cookies.last_access_utc, cookies.host_key, cookies.path, cookies.name, cookies.value FROM cookies

History Index YYYY-MM 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30

Each History Index file tracks a page's URL, title, and text content for a given month.

pages_content	
docid (INTEGER)	Primary Key [7]
c0url	http://www.obsidianforensics.com/blog/hindsight/
c1title	[Obsidian Forensics » Announcing Hindsight - A New Free Chrome Forensics Tool]
c2body	Indexed text from page [Announcing Hindsight - A New Free Chrome Forensics Tool Ryan / 2 comments / Hindsight is a free tool...]
info	
time (INTEGER)	WebKit Timestamp [1301488000671600]

Web Data X 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30

The Web Data database records text a user enters into web forms to let Chrome to automatically fill in similar future forms.

autofill	
pair_id (INTEGER)	Primary Key [1]
name (VARCHAR)	Name of the input as specified in the HTML [Keyword]
value (VARCHAR)	The literal contents of the text field [Chrome forensics]
value_lower (VARCHAR)	The contents of the text field made lower case. [chrome forensics]
count (INTEGER)	How many times the user has entered the string value in a field of name name. [2]
autofill_dates	
pair_id (INTEGER)	Foreign Key to autofill.pair_id. Not unique. Chrome stores the first use and last use timestamps for a pair_id. [1]
date_created (INTEGER)	Unix Timestamp [1372310048]

SQL for basic autofill information: SELECT autofill.value, autofill.name, autofill.count, autofill_dates.date_created FROM autofill, autofill_dates WHERE autofill.pair_id = autofill_dates.pair_id

References
[1]. http://src.chromium.org/viewvc/chrome/trunk/src/content/public/common/page_transition_types_list.h
[2]. http://src.chromium.org/svn/trunk/src/chrome/browser/history/visit_database.cc
[3]. http://src.chromium.org/chrome/trunk/src/content/public/browser/download_danger_type.h
[4]. <http://src.chromium.org/viewvc/chrome?revision=212459&view=revision>