Mobile Forensics: Extracting Data from WhatsApp

hackers-arise.com/mobile-forensics-extracting-data-from-whatsapp/

Welcome back, aspiring digital investigators!

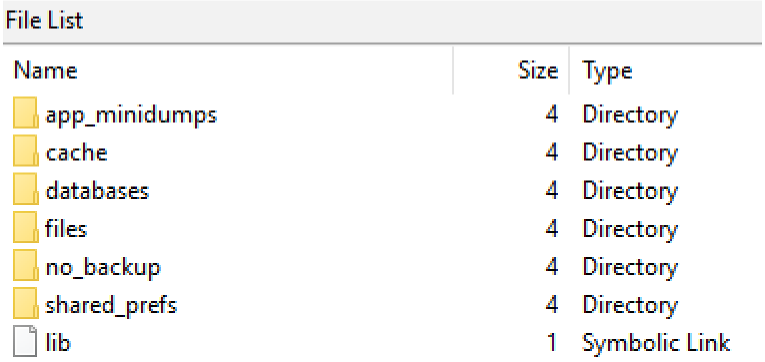
Today we will take a look at WhatsApp forensics. WhatsApp is one of those apps that are both private and routine for many users. People treat chats like a private conversation, and because it feels comfortable, users often share things there that they would not say on public social networks. That's why WhatsApp is so critical for digital forensics. The app stores conversations, media, timestamps, group membership information and metadata that can help reconstruct events, identify contacts and corroborate timelines in criminal and cyber investigations.

At Hackers-Arise we offer professional digital forensics services that support cybercrime investigations and fraud examinations. WhatsApp forensics is done to find reliable evidence. The data recovered from a device can show who communicated with whom, when messages were sent and received, what media was exchanged, and often which account owned the device. That information is used to link suspects and verify statements. It also maps movements when combined with location artifacts that investigators and prosecutors can trust.

You will see how WhatsApp keeps its data on different platforms and what those files contain.

WhatsApp Artifacts on Android Devices

On Android, WhatsApp stores most of its private application data inside the device's user data area. In a typical layout you will find the app's files under a path such as `/data/data/com.whatsapp/` (or equivalently `/data/user/0/com.whatsapp/` on many devices). Those directories are not normally accessible without elevated privileges. To read them directly you will usually need superuser (root) access on the device or a physical dump of the file system obtained through lawful and technically appropriate means. If you do not have root or a physical image, your options are restricted to logical backups or other extraction methods which may not expose the private WhatsApp databases.

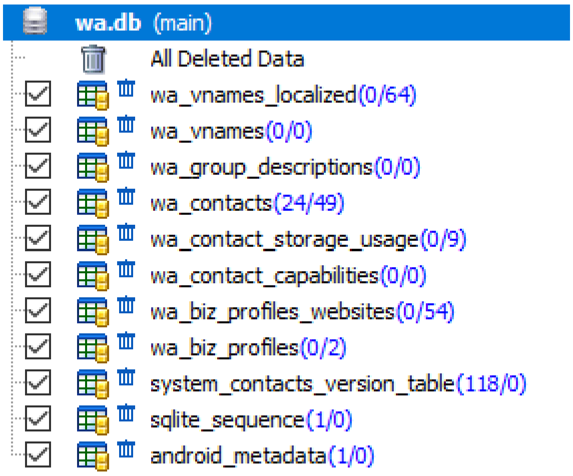


File List

Name	Size	Type
app_minidumps	4	Directory
cache	4	Directory
databases	4	Directory
files	4	Directory
no_backup	4	Directory
shared_prefs	4	Directory
lib	1	Symbolic Link

Source: Group-IB

Two files deserve immediate attention on Android: `wa.db` and `msgstore.db`. Both are SQLite databases and together they form the core of WhatsApp evidence.

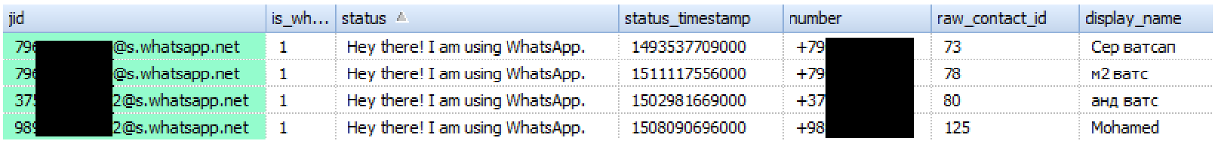


wa.db (main)

- All Deleted Data
- ☑ wa_vnames_localized(0/64)
- ☑ wa_vnames(0/0)
- ☑ wa_group_descriptions(0/0)
- ☑ wa_contacts(24/49)
- ☑ wa_contact_storage_usage(0/9)
- ☑ wa_contact_capabilities(0/0)
- ☑ wa_biz_profiles_websites(0/54)
- ☑ wa_biz_profiles(0/2)
- ☑ system_contacts_version_table(118/0)
- ☑ sqlite_sequence(1/0)
- ☑ android_metadata(1/0)

Source: Group-IB

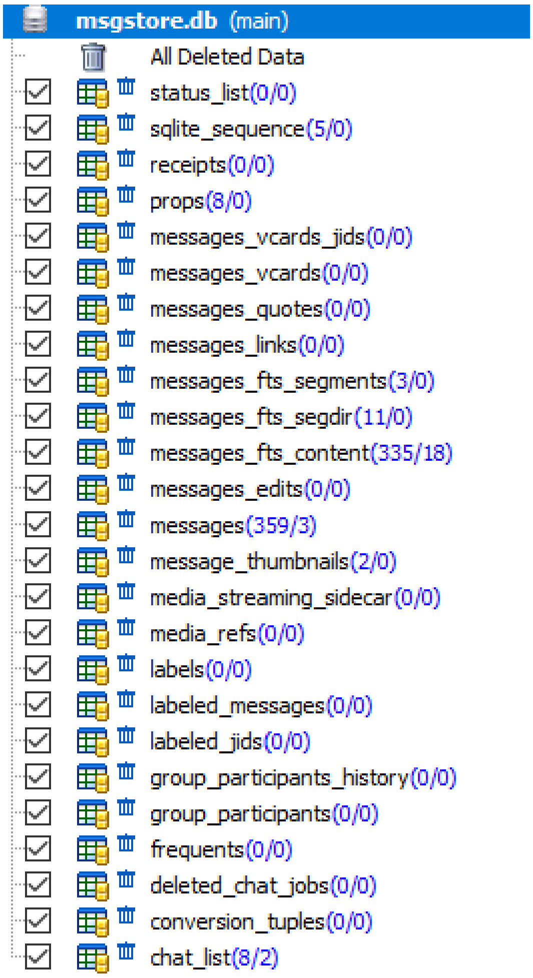
`wa.db` is the contacts database. It lists the WhatsApp user's contacts and typically contains phone numbers, display names, status strings, timestamps for when contacts were created or changed, and other registration metadata. You will usually open the file with a SQLite browser or query it with `sqlite3` to inspect tables. The key tables investigators look for are the table that stores contact records (often named `wa_contacts` or similar), `sqlite_sequence` which holds auto-increment counts and gives you a sense of scale, and `android_metadata` which contains localization info such as the app language.



jid	is_wh...	status	status_timestamp	number	raw_contact_id	display_name
79[redacted]@s.whatsapp.net	1	Hey there! I am using WhatsApp.	1493537709000	+79[redacted]	73	Сер ватсап
79[redacted]@s.whatsapp.net	1	Hey there! I am using WhatsApp.	1511117556000	+79[redacted]	78	м2 ватс
37[redacted]2@s.whatsapp.net	1	Hey there! I am using WhatsApp.	1502981669000	+37[redacted]	80	анд ватс
98[redacted]2@s.whatsapp.net	1	Hey there! I am using WhatsApp.	1508090696000	+98[redacted]	125	Mohamed

Source: Group-IB

`Wa.db` is essentially the address book for WhatsApp. It has names, numbers and a little context for each contact.

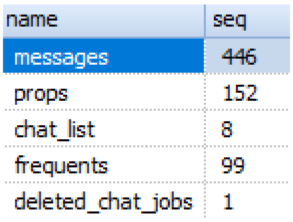


msgstore.db (main)
All Deleted Data
☑ status_list(0/0)
☑ sqlite_sequence(5/0)
☑ receipts(0/0)
☑ props(8/0)
☑ messages_vcards_jids(0/0)
☑ messages_vcards(0/0)
☑ messages_quotes(0/0)
☑ messages_links(0/0)
☑ messages_fts_segments(3/0)
☑ messages_fts_segdir(11/0)
☑ messages_fts_content(335/18)
☑ messages_edits(0/0)
☑ messages(359/3)
☑ message_thumbnails(2/0)
☑ media_streaming_sidecar(0/0)
☑ media_refs(0/0)
☑ labels(0/0)
☑ labeled_messages(0/0)
☑ labeled_jids(0/0)
☑ group_participants_history(0/0)
☑ group_participants(0/0)
☑ frequents(0/0)
☑ deleted_chat_jobs(0/0)
☑ conversion_tuples(0/0)
☑ chat_list(8/2)

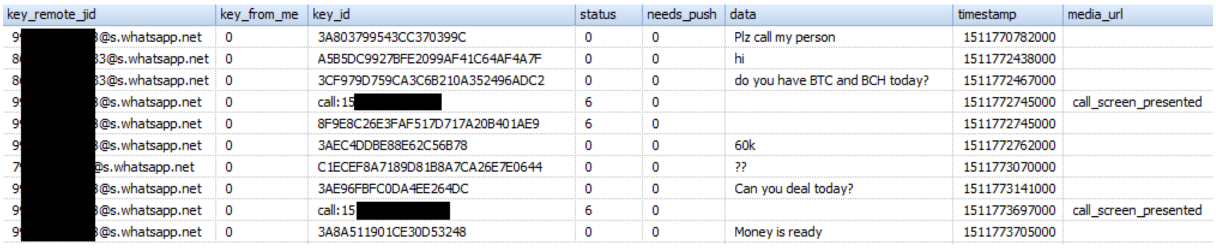
Source: Group-IB

`msgstore.db` is the message store. This database contains sent and received messages, timestamps, message status, sender and receiver identifiers, and references to media files. In many WhatsApp versions you will find tables that include a general information table (often named `sqlite_sequence`), a full-text index table for message content (`message_fts_content` or similar), the main `messages` table which usually contains the message body and metadata, `messages_thumbnails` which catalogs images and their timestamps, and a `chat_list` table that stores conversation entries.

Be aware that WhatsApp evolves and field names change between versions. Newer schema versions may include extra fields such as `media_enc_hash`, `edit_version`, or `payment_transaction_id`. Always inspect the schema before you rely on a specific field name.



name	seq
messages	446
props	152
chat_list	8
frequents	99
deleted_chat_jobs	1

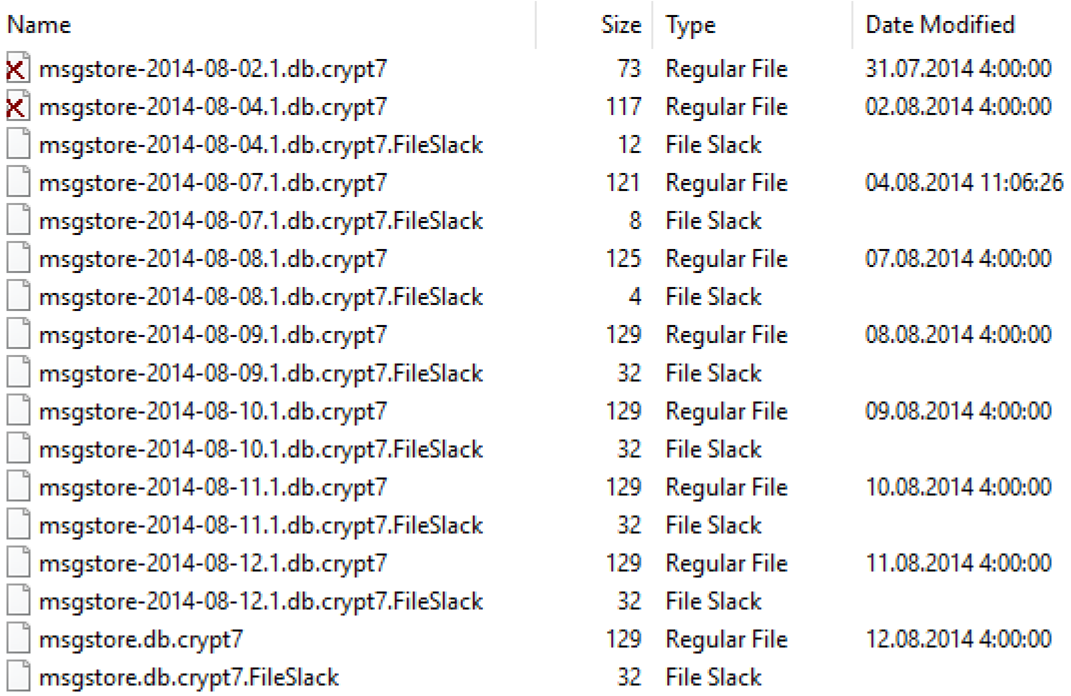


key_remote_jid	key_from_me	key_id	status	needs_push	data	timestamp	media_url
9[redacted]@s.whatsapp.net	0	3A803799543CC370399C	0	0	Plz call my person	1511770782000	
8[redacted]33@s.whatsapp.net	0	A5B5DC9927BFE2099AF41C64AF4A7F	0	0	hi	1511772438000	
8[redacted]33@s.whatsapp.net	0	3CF979D759CA3C6B210A352496ADC2	0	0	do you have BTC and BCH today?	1511772467000	
9[redacted]@s.whatsapp.net	0	call:15[redacted]	6	0		1511772745000	call_screen_presented
9[redacted]@s.whatsapp.net	0	8F9E8C26E3FAF517D717A20B401AE9	6	0		1511772745000	
9[redacted]@s.whatsapp.net	0	3AEC4DDBE88E62C56B78	0	0	60k	1511772762000	
7[redacted]@s.whatsapp.net	0	C1ECEF8A7189D81B8A7CA26E7E0644	0	0	??	1511773070000	
9[redacted]@s.whatsapp.net	0	3AE96FBFC0DA4EE264DC	0	0	Can you deal today?	1511773141000	
9[redacted]@s.whatsapp.net	0	call:15[redacted]	6	0		1511773697000	call_screen_presented
9[redacted]@s.whatsapp.net	0	3A8A511901CE30D53248	0	0	Money is ready	1511773705000	

Source: Group-IB

On many Android devices WhatsApp also keeps encrypted backups in a public storage location, typically under `/data/media/0/WhatsApp/Databases/` (the virtual SD card)

or `/mnt/sdcard/WhatsApp/Databases/` for physical SD cards. Those backup files look like `msgstore.db.cryptXX`, where `XX` indicates the cryptographic scheme version.



Name	Size	Type	Date Modified
☒ msgstore-2014-08-02.1.db.crypt7	73	Regular File	31.07.2014 4:00:00
☒ msgstore-2014-08-04.1.db.crypt7	117	Regular File	02.08.2014 4:00:00
msgstore-2014-08-04.1.db.crypt7.FileSlack	12	File Slack	
msgstore-2014-08-07.1.db.crypt7	121	Regular File	04.08.2014 11:06:26
msgstore-2014-08-07.1.db.crypt7.FileSlack	8	File Slack	
msgstore-2014-08-08.1.db.crypt7	125	Regular File	07.08.2014 4:00:00
msgstore-2014-08-08.1.db.crypt7.FileSlack	4	File Slack	
msgstore-2014-08-09.1.db.crypt7	129	Regular File	08.08.2014 4:00:00
msgstore-2014-08-09.1.db.crypt7.FileSlack	32	File Slack	
msgstore-2014-08-10.1.db.crypt7	129	Regular File	09.08.2014 4:00:00
msgstore-2014-08-10.1.db.crypt7.FileSlack	32	File Slack	
msgstore-2014-08-11.1.db.crypt7	129	Regular File	10.08.2014 4:00:00
msgstore-2014-08-11.1.db.crypt7.FileSlack	32	File Slack	
msgstore-2014-08-12.1.db.crypt7	129	Regular File	11.08.2014 4:00:00
msgstore-2014-08-12.1.db.crypt7.FileSlack	32	File Slack	
msgstore.db.crypt7	129	Regular File	12.08.2014 4:00:00
msgstore.db.crypt7.FileSlack	32	File Slack	

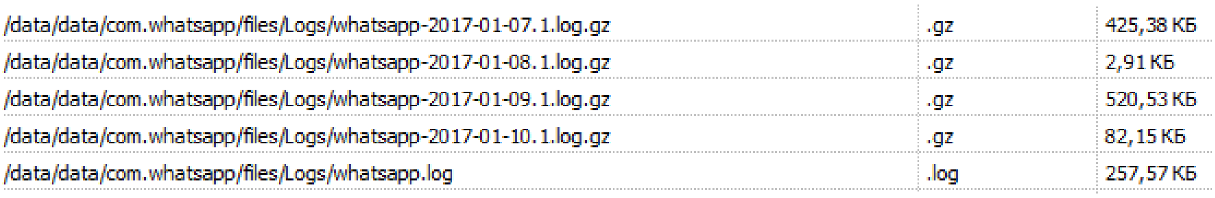
Source: Group-IB

The `msgstore.db.cryptXX` files are an encrypted copy of `msgstore.db` intended for device backups. To decrypt them you need a cryptographic key that WhatsApp stores privately on the device, usually somewhere like `/data/data/com.whatsapp/files/`. Without that key, those encrypted backups are not readable.

Other important Android files and directories to examine include the preferences and registration XMLs in `/data/data/com.whatsapp/shared_prefs/`. The file `com.whatsapp_preferences.xml` often contains profile details and configuration values. A fragment of such a file may show the phone number associated with the account, the app version, a profile message such as “Hey there! I am using WhatsApp.” and the account display name. The `registration.RegisterPhone.xml` file typically contains registration metadata like the phone number and regional format.

The `axolotl.db` file in `/data/data/com.whatsapp/databases/` holds cryptographic keys (used in the Signal/Double Ratchet protocol implementation) and account identification data. `chatsettings.db` contains app settings. Logs are kept under `/data/data/com.whatsapp/files/Logs/` and may include `whatsapp.log` as well as compressed rotated backups like `whatsapp-YYYY-MM-DD.1.log.gz`

These logs can reveal app activity and errors that may be useful for timing or troubleshooting analysis.



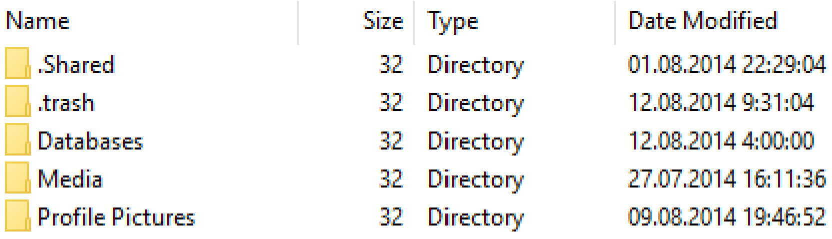
/data/data/com.whatsapp/files/Logs/whatsapp-2017-01-07.1.log.gz	.gz	425,38 КБ
/data/data/com.whatsapp/files/Logs/whatsapp-2017-01-08.1.log.gz	.gz	2,91 КБ
/data/data/com.whatsapp/files/Logs/whatsapp-2017-01-09.1.log.gz	.gz	520,53 КБ
/data/data/com.whatsapp/files/Logs/whatsapp-2017-01-10.1.log.gz	.gz	82,15 КБ
/data/data/com.whatsapp/files/Logs/whatsapp.log	.log	257,57 КБ

Source: Group-IB

Media is often stored in the media tree on internal or external storage:

`/data/media/0/WhatsApp/Media/WhatsApp Images/` for images,

`/data/media/0/WhatsApp/Media/WhatsApp Voice Notes/` for voice messages (usually Opus format), `WhatsApp Audio`, `WhatsApp Video`, and `WhatsApp Profile Photos`.



Name	Size	Type	Date Modified
.Shared	32	Directory	01.08.2014 22:29:04
.trash	32	Directory	12.08.2014 9:31:04
Databases	32	Directory	12.08.2014 4:00:00
Media	32	Directory	27.07.2014 16:11:36
Profile Pictures	32	Directory	09.08.2014 19:46:52

Source: Group-IB

Within the app's private area you may also find cached profile pictures under `/data/data/com.whatsapp/cache/Profile Pictures/` and avatar thumbnails under `/data/data/com.whatsapp/files/Avatars/`. Some avatar thumbnails use a `.j` extension while actually being JPEG files. Always validate file signatures rather than trusting extensions.

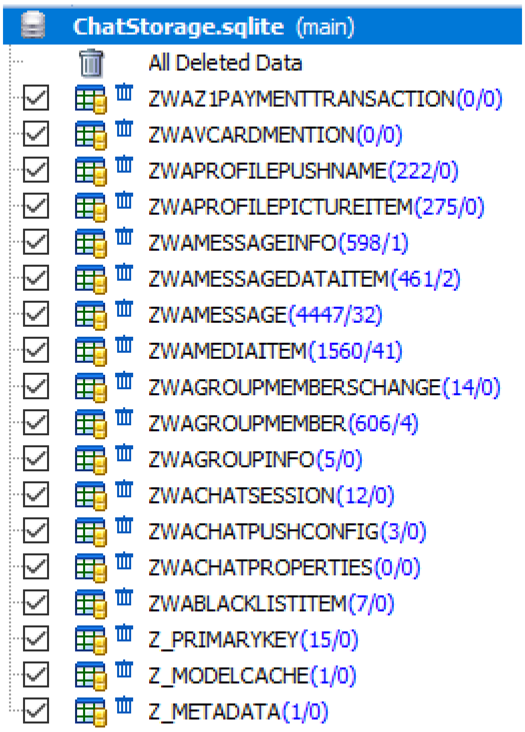
If the device uses an SD card, a WhatsApp directory at the card's root may store copies of shared files (`/mnt/sdcard/WhatsApp/.Share/`), a trash folder for deleted content (`/mnt/sdcard/WhatsApp/.trash/`), and the `Databases` subdirectory with encrypted backups and media subfolders mirroring those on internal storage. The presence of deleted files or `.trash` folders can be a fruitful source of recovered media.

A key complication on Android is manufacturer or custom-ROM behavior. Some vendors add features that change where app data is stored. For example, certain Xiaomi phones implement a "Second Space" feature that creates a second user workspace. WhatsApp in the second workspace stores its data under a different user ID path such as `/data/user/10/com.whatsapp/databases/wa.db` rather than the usual `/data/user/0/com.whatsapp/databases/wa.db`

As things evolve and change, you need to validate the actual paths on the target device rather than assuming standard locations.

WhatsApp Artifacts on iOS Devices

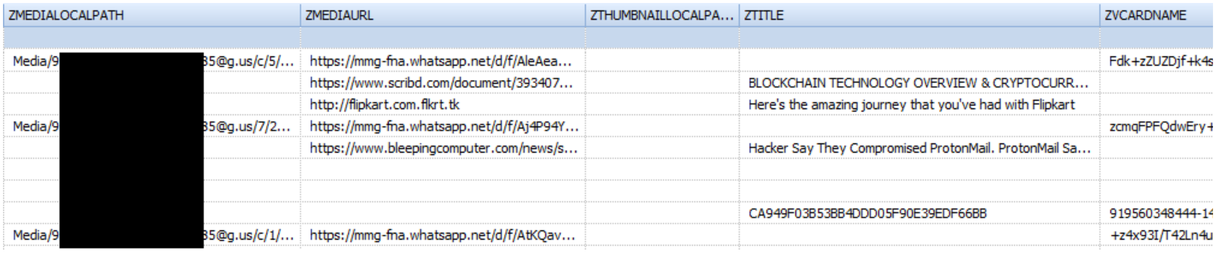
On iOS, WhatsApp tends to centralize its data into a few places and is commonly accessible via device backups. The main application database is often `ChatStorage.sqlite` located under a shared group container such as `/private/var/mobile/Applications/group.net.whatsapp.WhatsApp.shared/` (some forensic tools display this as `AppDomainGroup-group.net.whatsapp.WhatsApp.shared`).



ChatStorage.sqlite (main)
All Deleted Data
☑ ZWAZ1PAYMENTTRANSACTION(0/0)
☑ ZWAVCARDMENTION(0/0)
☑ ZWAPROFILEPUSHNAME(222/0)
☑ ZWAPROFILEPICTUREITEM(275/0)
☑ ZWAMESSAGEINFO(598/1)
☑ ZWAMESSAGEDATAITEM(461/2)
☑ ZWAMESSAGE(4447/32)
☑ ZWAMEDIAITEM(1560/41)
☑ ZWAGROUPMEMBERSCHANGE(14/0)
☑ ZWAGROUPMEMBER(606/4)
☑ ZWAGROUPINFO(5/0)
☑ ZWACHATSESSION(12/0)
☑ ZWACHATPUSHCONFIG(3/0)
☑ ZWACHATPROPERTIES(0/0)
☑ ZWABLACKLISTITEM(7/0)
☑ Z_PRIMARYKEY(15/0)
☑ Z_MODELCACHE(1/0)
☑ Z_METADATA(1/0)

Source: Group-IB

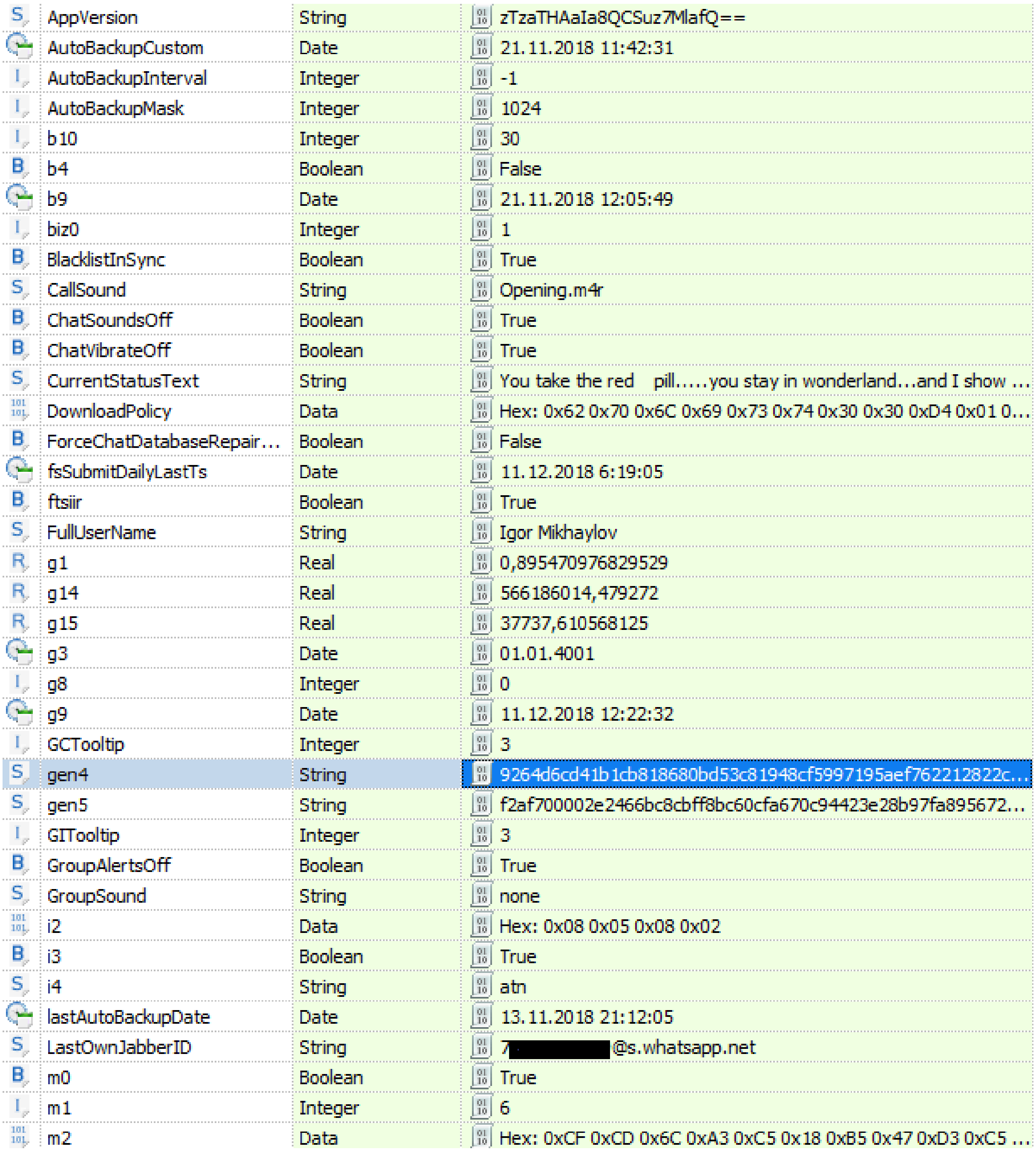
Within `ChatStorage.sqlite` the most informative tables are often `ZWAMESSAGE`, which stores message records, and `ZWAMEDIAITEM`, which stores metadata for attachments and media items. Other tables like `ZWAPROFILEPUSHNAME` and `ZWAPROFILEPICTUREITEM` map WhatsApp identifiers to display names and avatars. The table `Z_PRIMARYKEY` typically contains general database metadata such as record counts.



ZMEDIALOCALPATH	ZMEDIAURL	ZTHUMBNAILLOCALPA…	ZTITLE	ZVCARDNAME
Media/9[redacted]5@g.us/c/5/…	https://mmg-fna.whatsapp.net/d/f/AleAea…			Fdk+zZUZDjf+k4s
[redacted]	https://www.scribd.com/document/393407…		BLOCKCHAIN TECHNOLOGY OVERVIEW & CRYPTOCURR…	
[redacted]	http://flipkart.com.flkrt.tk		Here's the amazing journey that you've had with Flipkart	
Media/9[redacted]5@g.us/7/2…	https://mmg-fna.whatsapp.net/d/f/Aj4P94Y…			zcmqFPFQdwEry+
[redacted]	https://www.bleepingcomputer.com/news/s…		Hacker Say They Compromised ProtonMail. ProtonMail Sa…	
[redacted]				
[redacted]				
[redacted]			CA949F03B53BB4DDD05F90E39EDF66BB	919560348444-1
Media/9[redacted]5@g.us/c/1/…	https://mmg-fna.whatsapp.net/d/f/AtKQav…			+z4x93I/T42Ln4u

Source: Group-IB

iOS also places supporting files in the group container. `BackedUpKeyValue.sqlite` can contain cryptographic keys and data useful for identifying account ownership. `ContactsV2.sqlite` stores contact details: names, phone numbers, profile statuses and WhatsApp IDs. A simple text file like `consumer_version` may hold the app version and `current_wallpaper.jpg` (or `wallpaper` in older versions) contains the background image used in WhatsApp chats. The `blockedcontacts.dat` file lists blocked numbers, and `pw.dat` can hold an encrypted password. Preference plists such as `net.whatsapp.WhatsApp.plist` or `group.net.whatsapp.WhatsApp.shared.plist` store profile settings.



S	AppVersion	String	zTzaTHAaIa8QCSuz7MlafQ==
	AutoBackupCustom	Date	21.11.2018 11:42:31
I	AutoBackupInterval	Integer	-1
I	AutoBackupMask	Integer	1024
I	b10	Integer	30
B	b4	Boolean	False
	b9	Date	21.11.2018 12:05:49
I	biz0	Integer	1
B	BlacklistInSync	Boolean	True
S	CallSound	String	Opening.m4r
B	ChatSoundsOff	Boolean	True
B	ChatVibrateOff	Boolean	True
S	CurrentStatusText	String	You take the red pill.....you stay in wonderland...and I show ...
101 101	DownloadPolicy	Data	Hex: 0x62 0x70 0x6C 0x69 0x73 0x74 0x30 0x30 0xD4 0x01 0...
B	ForceChatDatabaseRepair...	Boolean	False
	fsSubmitDailyLastTs	Date	11.12.2018 6:19:05
B	ftsiir	Boolean	True
S	FullUserName	String	Igor Mikhaylov
R	g1	Real	0,895470976829529
R	g14	Real	566186014,479272
R	g15	Real	37737,610568125
	g3	Date	01.01.4001
I	g8	Integer	0
	g9	Date	11.12.2018 12:22:32
I	GCTooltip	Integer	3
S	gen4	String	9264d6cd41b1cb818680bd53c81948cf5997195aef762212822c...
S	gen5	String	f2af700002e2466bc8cbff8bc60cfa670c94423e28b97fa895672...
I	GITooltip	Integer	3
B	GroupAlertsOff	Boolean	True
S	GroupSound	String	none
101 101	i2	Data	Hex: 0x08 0x05 0x08 0x02
B	i3	Boolean	True
S	i4	String	atn
	lastAutoBackupDate	Date	13.11.2018 21:12:05
S	LastOwnJabberID	String	7[redacted]@s.whatsapp.net
B	m0	Boolean	True
I	m1	Integer	6
101 101	m2	Data	Hex: 0xCF 0xCD 0x6C 0xA3 0xC5 0x18 0xB5 0x47 0xD3 0xC5 ...

Source: Group-IB

Media thumbnails, avatars and message media are stored under paths like `/private/var/mobile/Applications/group.net.whatsapp.WhatsApp.shared/Media/Profile/` and `/private/var/mobile/Applications/group.net.whatsapp.WhatsApp.shared/Message/Media/`. WhatsApp logs, for example `calls.log` and `calls.backup.log`, often survive in the Documents or Library/Logs folders and can help establish call activity.

Because iOS devices are frequently backed up through iTunes or Finder, you can often extract WhatsApp artefacts from a device backup rather than needing a full file system image. If the backup is unencrypted and complete, it may include the `ChatStorage.sqlite` file and associated media. If the backup is encrypted you will need the backup password or legal access methods to decrypt it. In practice, many investigators create a forensic backup and then examine the WhatsApp databases with a SQLite viewer or a specialized forensic tool that understands WhatsApp schema differences across versions.

Practical Notes For Beginners

From the databases and media files described above you can recover contact lists, full or partial chat histories, timestamps in epoch format (commonly Unix epoch in milliseconds on Android), message status (sent, delivered, read), media filenames and hashes, group membership, profile names and avatars, blocked contacts, and even application logs and version metadata. It helps us understand who communicated with whom, when messages were exchanged, whether media were transferred, and which accounts were configured on the device.

For beginners, a few practical cautions are important to keep in mind. First, always operate on forensic images or copies of extracted files. Do not work directly on the live device unless you are performing an approved, controlled acquisition and you have documented every action. Second, use reliable forensic tools to open SQLite databases. If you are parsing fields manually, confirm timestamp formats and time zones. Third, encrypted backups require the device’s key to decrypt. The key is usually stored in the private application area on Android, and without it you cannot decode the .cryptXX files. Fourth, deleted chats and files are not always gone, as databases may leave records or media may remain in caches or on external storage. Yet recovery is never guaranteed and depends on many factors including the time since deletion and subsequent device activity.

When you review message tables, map the message ID fields to media references carefully. Many WhatsApp versions use separate tables for media items where the actual file is referenced by a media ID or filename. Thumbnail tables and media directories will help you reconstruct the link between a textual message and the file that accompanied it. Pay attention to the presence of additional fields in newer app versions. These may contain payment IDs, edit history or encryption metadata. Adapt your queries accordingly.

Finally, because WhatsApp and operating systems change over time, always inspect the schema and file timestamps on the specific evidence you have. Do not assume field names or paths are identical between devices or app versions. Keep a list of the paths and filenames you find so you can reproduce your process and explain it in reports.

Summary

WhatsApp forensics is a rich discipline. On Android the primary artifacts are the wa.db contacts database, the msgstore.db message store and encrypted backups such as msgstore.db.cryptXX, together with media directories, preference XMLs and cryptographic key material in the app private area. On iOS the main artifact is ChatStorage.sqlite and a few supporting files in the app group container and possibly contained in a device backup. To retrieve and interpret these artifacts you must have appropriate access to the device or an image and know where to look for the app files on the device you are examining. Also, be comfortable inspecting SQLite databases and be prepared to decrypt backups where necessary.

If this kind of work interests you and you want to understand how real mobile investigations are carried out, you can also join our three-day mobile forensics course. The training walks you through the essentials of Android and iOS, explains how evidence is stored on modern devices, and shows you how investigators extract and analyze that data during real cases. You will work with practical labs that involve hidden apps, encrypted communication, and devices that may have been rooted or tampered with.

Learn more:

https://hackersarise.thinkific.com/courses/mobile-forensics