# Executive Summary

This forensic analysis aimed to examine WhatsApp on an Android device, following the NIST 800-101 guidelines for mobile forensics. (Richard Ayers (NIST), 2014) The investigation focused on identifying and extracting key forensic artifacts from WhatsApp’s file system, including chat histories, contact details, cryptographic keys, and media metadata. Through a structured forensic approach, various tools were used to analyse the application’s internal architecture, databases, and security measures.

The analysis revealed that WhatsApp stores critical user data in multiple SQLite databases, with msgstore.db containing chat messages and timestamps, wa.db storing contact details and group memberships, and axolotl.db managing cryptographic keys for end-to-end encryption. Additionally, media files and shared locations were found

in media.db and location.db, which may provide further forensic details. The investigation also confirmed that WhatsApp employs a client-server architecture with encrypted communication via the Signal Protocol, preventing unauthorized access to message content.

# Methodology

For the forensic analysis of the WhatsApp mobile app on an android device, to ensure a sound analysis was done, the NIST 800-101 guidelines for mobile device forensics were used. The guideline places an importance on the Identification, Acquisition, Analysis and Reporting on mobile devices, to ensure forensic integrity and data reliability. (Richard Ayers (NIST), 2014). For the selection of academic and industry references, only recent reputable sources that provided in depth details on the analysis of WhatsApp’s data storage architecture and primary databases, were used.

# Identification and Analysis of App Files and Artifacts

The first phase, in accordance with NIST guidelines, was to identifying potential artifacts within WhatsApp's Android file system. The main databases analysed were msgstore.db, wa.db and axolotl.db**.** msgstore.db was chosen, as it stores all chat messages, including timestamps, sender and receiver metadata, and message content, wa.db was chosen, as it stores contact information like user contacts, group memberships, and aids in the linking of phone numbers to user identities, and axolotl.db was chosen as it stores cryptographic keys and data related to the Signal Protocol. (Anglano, 2014) (K. Kaushik and Y. Katara, 2022)

# Forensic Analysis Methods and Tools

The first tool used was a macOS M1 chip laptop, were all the software needed, were downloaded onto. After Identifying the artifacts needed for the forensic analysis, the next step was to extract and analyse them. The WhatsApp apk (Version 2.25.5.17) was downloaded to observe how the app behaves.

The WhatsApp Apk was then run on Android Studio, an application which allows users to run virtual android devices (AVD). These AVD’s allow for a dynamic analysis of the app, in a controlled environment. (see Appendix F for screenshot) Once WhatsApp, was successfully running on the android VM, th APKTool was used to deconstruct the app’s binary into readable components, and examine its internal structure, to explore how WhatsApp stores and processes data. Additionally, **J**ADX was then used to decompile Dalvik Executable (DEX) files, to allow for exploring WhatsApp's Java code, and identify how data is handled and encrypted. These decompilers were chosen, as they provide a comprehensive breakdown of WhatsApp’s binary and source code structure. (see Appendix G for screenshot)

Once the APK was decompiled, Android Debug Bridge (ADB), was used to interact with the Android VM. ADB was essential, as it allowed data to be extracted from the live VM, which gave access to cloud-based backups and locally stored databases.(see Appendix E for android debug bridge in action)

WhatsApp stores local data in .db files in databases/, which can then be viewed with SQLite Database Browser. SQLite Database was used to parse and analyse the entire database, which included msgstore.db and wa.db. SQLite was used because it made extracting chat messages, timestamps and contact information easier, although some of the data was still encrypted.(see Appendix for SQLite in action).

# Test Environment

Since direct analysis of a mobile device could compromise data integrity, a virtualized test environment was used to create a controlled, safe environment for the analysis.

# App Investigation

WhatsApp is one of the most popular instant messaging applications worldwide. It offers text, voice, and multimedia communication, boasting over one billion users worldwide. It implements default end-to-end encryption using the Signal Protocol, ensuring that only communicating users can access message content, to enhance privacy and security. (Rastogi, 2017)

**Sources of Forensic Artifacts**

# Source Code

The first artifact explored, was the WhatsApp apk data, which was decompiled from Dalvik Executable (DEX) files, with JADX, to allow for the analysis of WhatsApp’s Java source code.

The decompiled Java code, was then viewed with the JadxGUI, which provided, a clear view of the application logic layer of the application, as it showed different packages like the androidx package, which ensures things like backward compatibility; the car.app package, which integrates WhatsApp functionality into automative environments like android auto, for a safe messaging experience while driving; the biometric package which allows biometric authentication methods like fingerprints and facial recognition; the ui package which is a collection of components and layouts that define WhatsApp’s aesthetics and interactive elements. (see Appendix A for screenshot of JadxGUI).

# Databases

In the /data/data/com.whatsapp/databases/ directory, there are 30 files associated with WhatsApp’s internal databases, however for this analysis the databases analysed are:

**axolotl.db:** this a 647 KB database, which stores cryptographic keys and data related to the Signal Protocol, which WhatsApp uses for end to end encryption. It stores tables like the sessions table, which keeps track of user sessions and keeps track of keys created when users send messages. The senders\_key table, keeps track of the senders\_id, and the senders\_account\_id, and basically details concerning the sender, and an identities table which stores the identities of the user and contacts, which whatsapp uses for end-to-end encryption.(see

Appendix B for screenshots)

**msgstore.db:** which is a 3.3MB database, which contains, tables storing everything involving messages, user's chat history, and stores messages exchanged with contacts. In the chat table, it stores information about, when messages were sent, which user sent them, when the last message was sent, an unseen message count and a timestamp which keeps track of when the user last read messages, and more. (see Appendix C for screenshots).

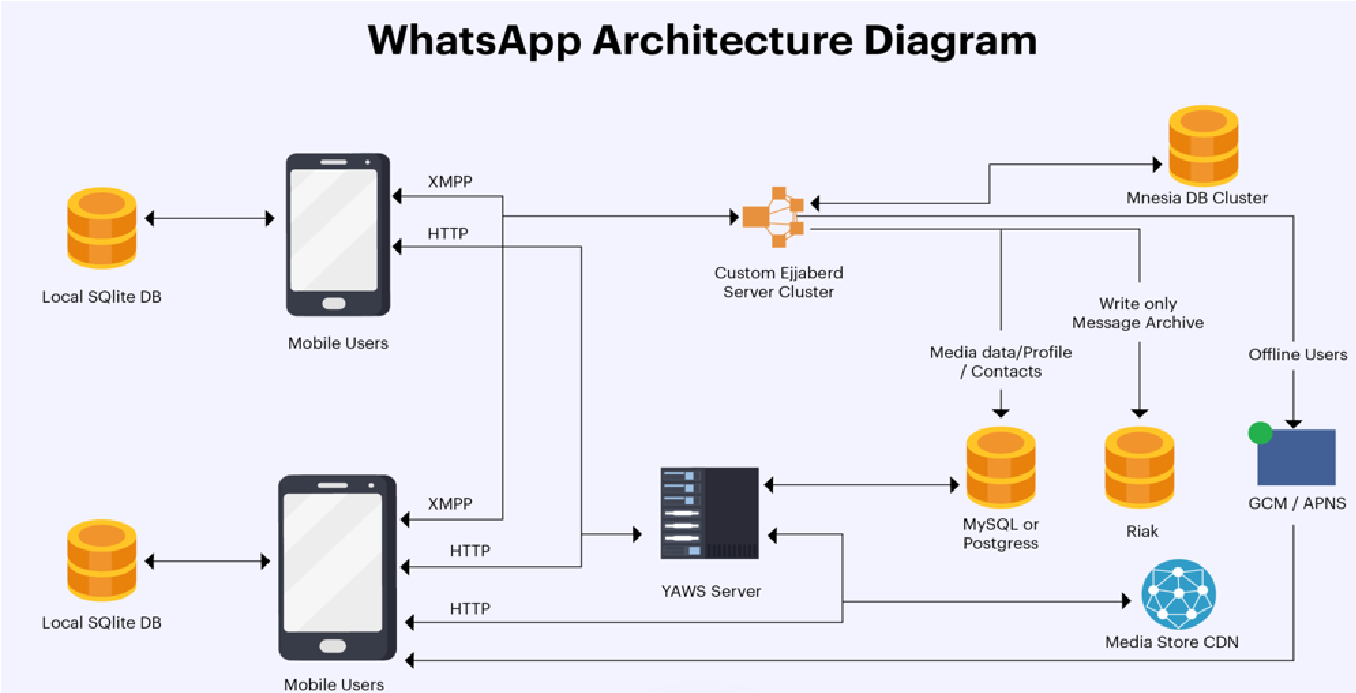
**wa.db:** this is a 602KB database, which stores information about the user’s WhatsApp contacts. Contacts details, are stored in the wa\_contacts database, which stores details like contacts display names, family names, if they’ve been reported for spam, and if they have a profile photo or not. This database also has a group of tables, which keep tracks of groups the user belongs to, as well as a table, which keeps track of the user id of blocked users on the application. (See Appendix D for screenshots).

Other notable databases in this directory are the **location.db** which stores logs shared location data, which includes things like real-time location sharing sessions; the media.db which keeps records of media files sent and received, including their locations on the device's storage; and the payments.db which handles data related to WhatsApp’s payment features, storing things like transaction history and payment credentials. (See Appendix E for screenshots).

# Architecture

WhatsApp uses a client-server architecture, which utilizes end-to-end encryption, for secure communication. The architecture consists of the clientlayer, business logic layer, and the database layer. The client-layer is responsible for handling user interactions, while the business logic layer handles the processing of messages and communication between users, while the database layer stores user information, meta data for user accounts and message data. (Cressler, 2021)

# WhatsApp Architecture Diagram

Each mobile app, hosts a local SQLite database for storing conversations, the client then uses HTTP WebSockets to send and retrieve data

*Figure 1(Cressler, 2021)* like images and

videos from the YAWS web server, while XMP is used to send those multimedia files to other users.

When an XMPP message is sent, it gets sent to Whatsapp’s custom Ejabberd server, which runs on BEAM and FreeBSD. The Ejabberd server saves the message in a Mnesia database table, after which it gets put in a queue. When the receiving user opens the app, the message in the queue gets routed through the Ejabberd server and delivered to the recipient, and once successfully delivered, the message then gets deleted from the Mnesia database. (Cressler, 2021).

# App Behaviour

Running WhatsApp on an Android emulator within a VM on an M1 Mac presents performance and compatibility challenges, as the app interacts with system services but struggles with Google Play-dependent features such as notifications and backups, which then led to having to create a virtual android device that uses the google API, rather than one that uses Google Play.

WhatsApp stores critical data in /data/data/com.whatsapp/, whilst Media files are stored separately, however access to the database was restricted due to Android’s security policies, so Root permissions were required for database extraction.

# Use Cases

WhatsApp is a widely adopted mobile messaging application with over 800 million users, however beyond personal mobile messaging, WhatsApp offers the ability for free one-on-one or group calls over the internet, with secure private connections. (WhatsApp, 2025)

WhatsApp offers file sharing for things like PDF’s, word documents, spreadsheets, and slides etc. (WhatsApp, 2025)

WhatsApp offers businesses the ability to broadcast messages to sent promotions, reminders and updates to customers who subscribe to them. Businesses also use WhatsApp to provide real time support with things like AIpowered chatbots. (WhatsApp, 2025)

# Recommendations

Research has shown, there are methods to decrypt network traffic and obtain forensic artifacts related to WhatsApp's calling feature, including phone numbers, server IPs, call durations, and termination details, with tools like

Wireshark. (W. Ahmed, 2021)

Forensic examiners can also analyse media files, and message details shared through WhatsApp to extract metadata, such as timestamps, geolocation data, device information, groups users’ belong to, etc. This process involves accessing dedicated folders where media files are stored and using tools to retrieve metadata, providing insights into the origin and history of the files. (W. Ahmed, 2021)

# Conclusion

The forensic analysis of WhatsApp on an Android device provided valuable insights into the application's data storage mechanisms, encryption practices, and forensic artifacts. By following the NIST 800-101 guidelines, the investigation ensured forensic integrity and data reliability. Key findings revealed that WhatsApp stores critical user data across multiple databases,

like msgstore.db for chat history, wa.db for contacts, axolotl.db for encryption keys and media.db for media files.

# Limitations

Despite the comprehensive approach, limitations were encountered, as running WhatsApp in an android emulator presented performance issues, particularly when it came to loading media and backups, and while SQLite Database Browser allowed the extraction of chat data, certain files were encrypted, requiring decryption keys, that may not always be available.

# Implications

Understanding WhatsApp’s data storage architecture will enable investigators to extract and analyse relevant forensic evidence in criminal investigations, cybercrime cases, and incident response scenarios. The identification of things like media metadata, timestamps, group membership, and encryption keys allow for digital reconstruction of user interactions, which can be pivotal in investigations*.*

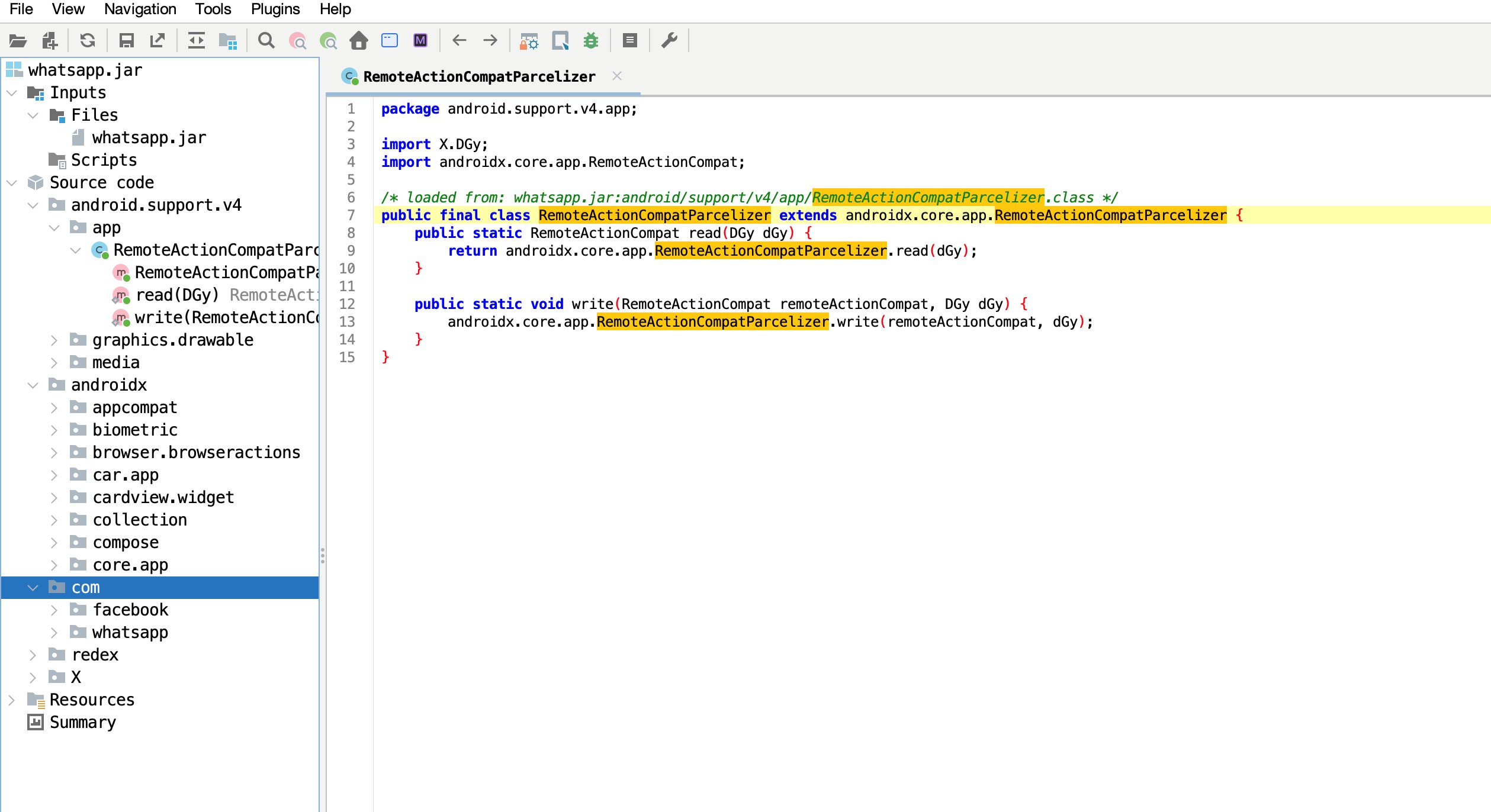
# Future Work

Further research can be done on network traffic analysis, which could provide additional forensic insights, by monitoring packet data and metadata, IP addresses could be traced and network tracking could also provide more context to call logs, and message timestamps, even if the message content remains encrypted.

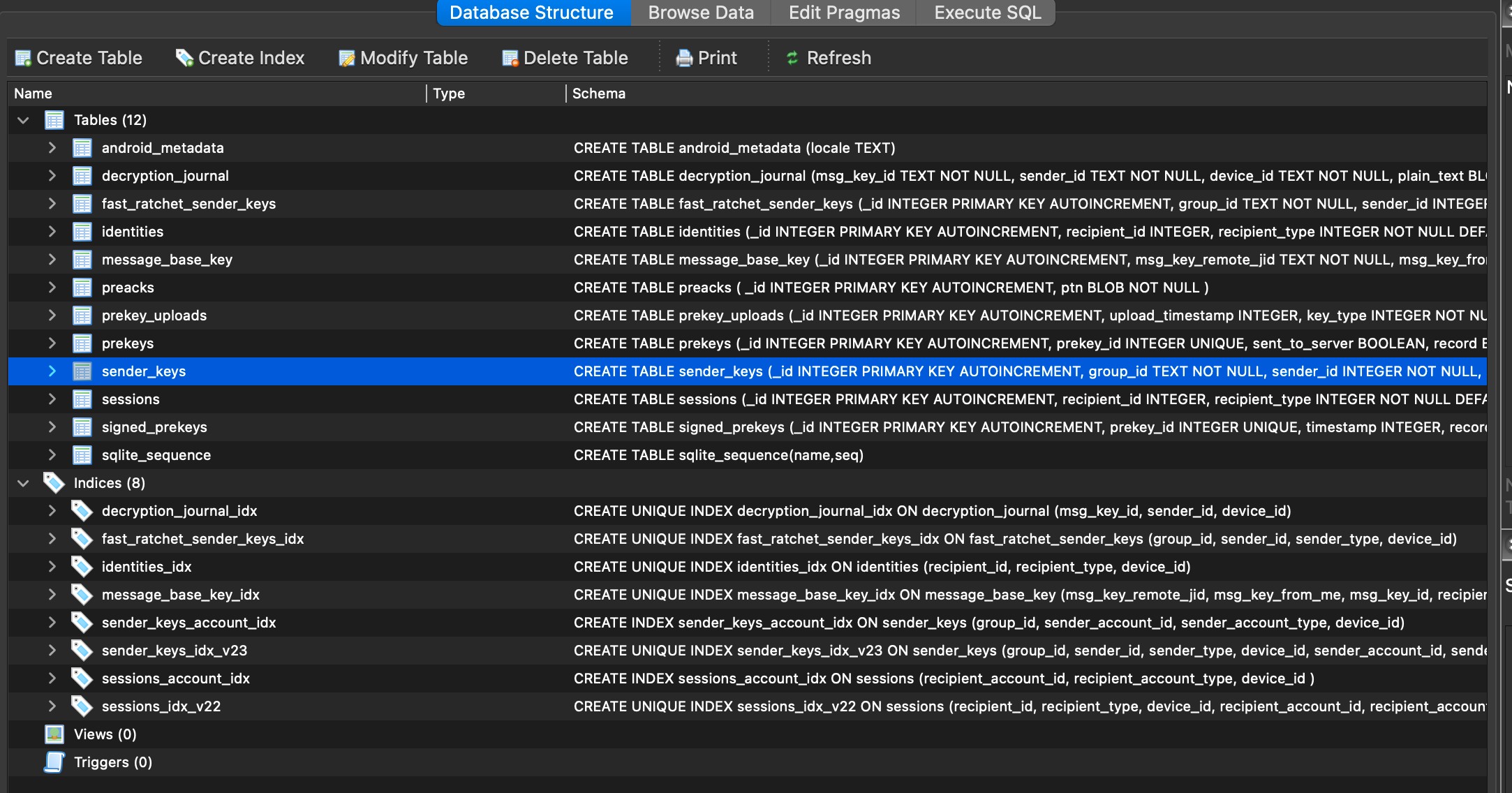
Expanding the study to include the analysis of WhatsApp across other platforms would provide a comparative analysis of WhatsApp data storage across different operating systems.

# Appendix

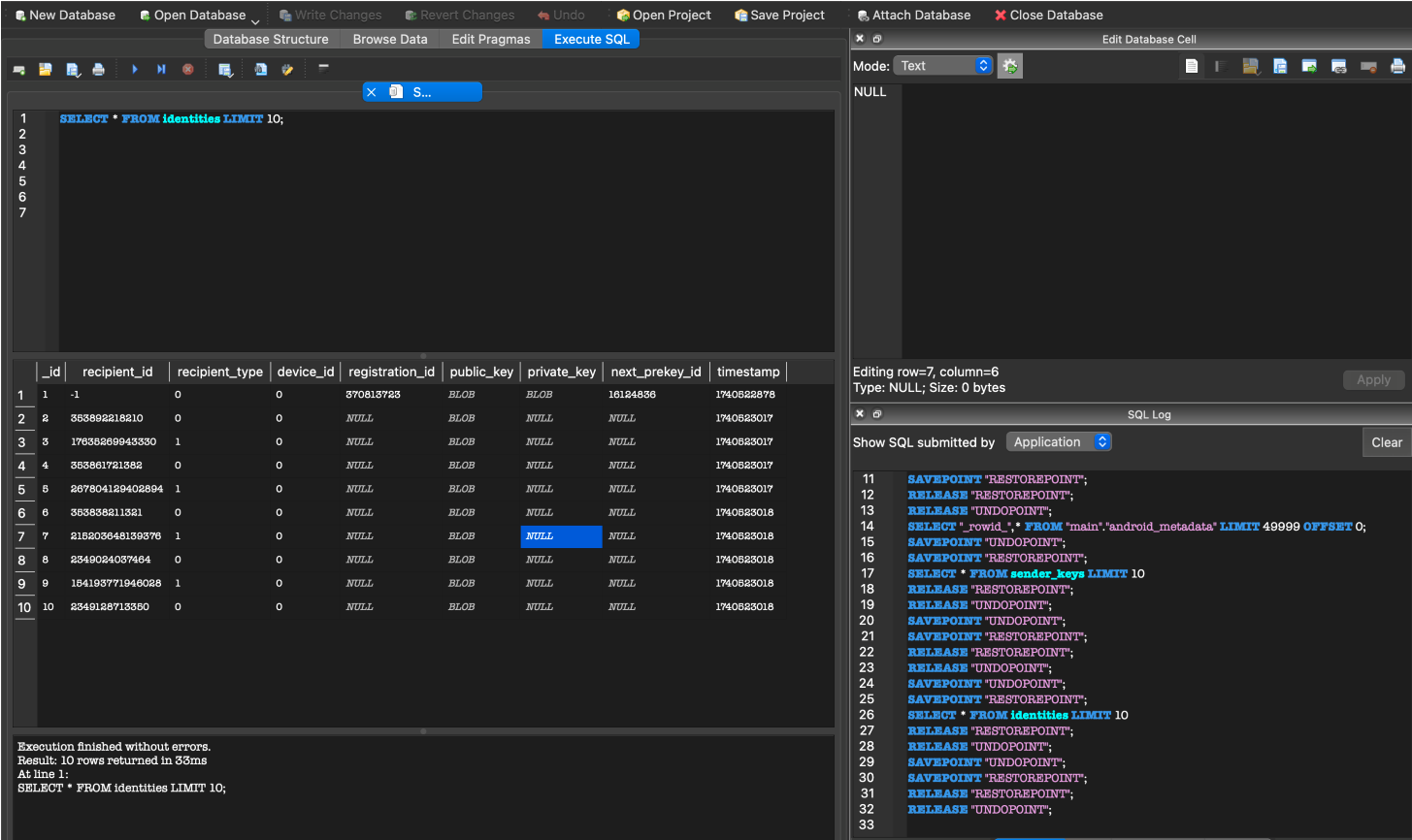
1. Screenshot of different java files in the JadxGUI



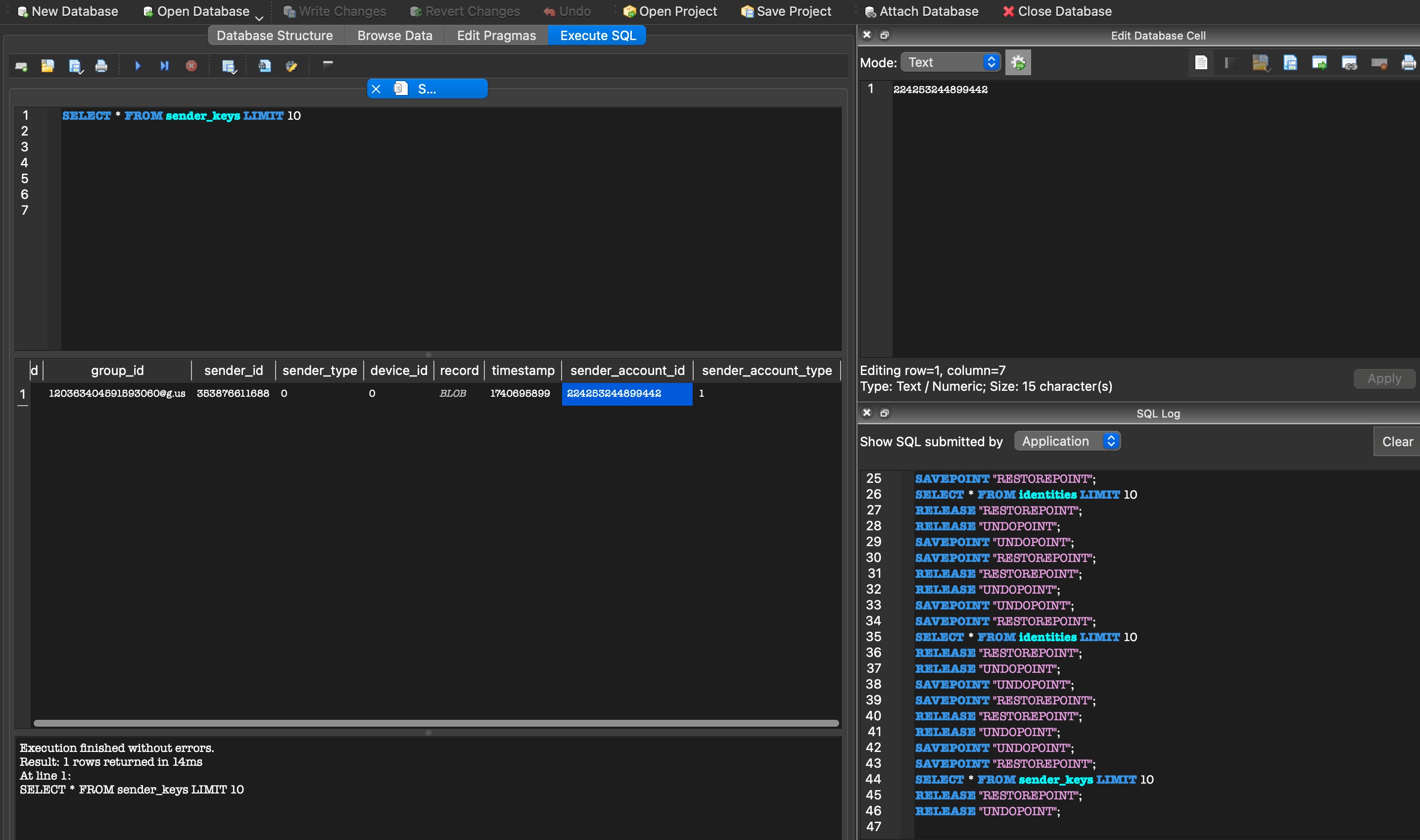
1. axolotl.db structure



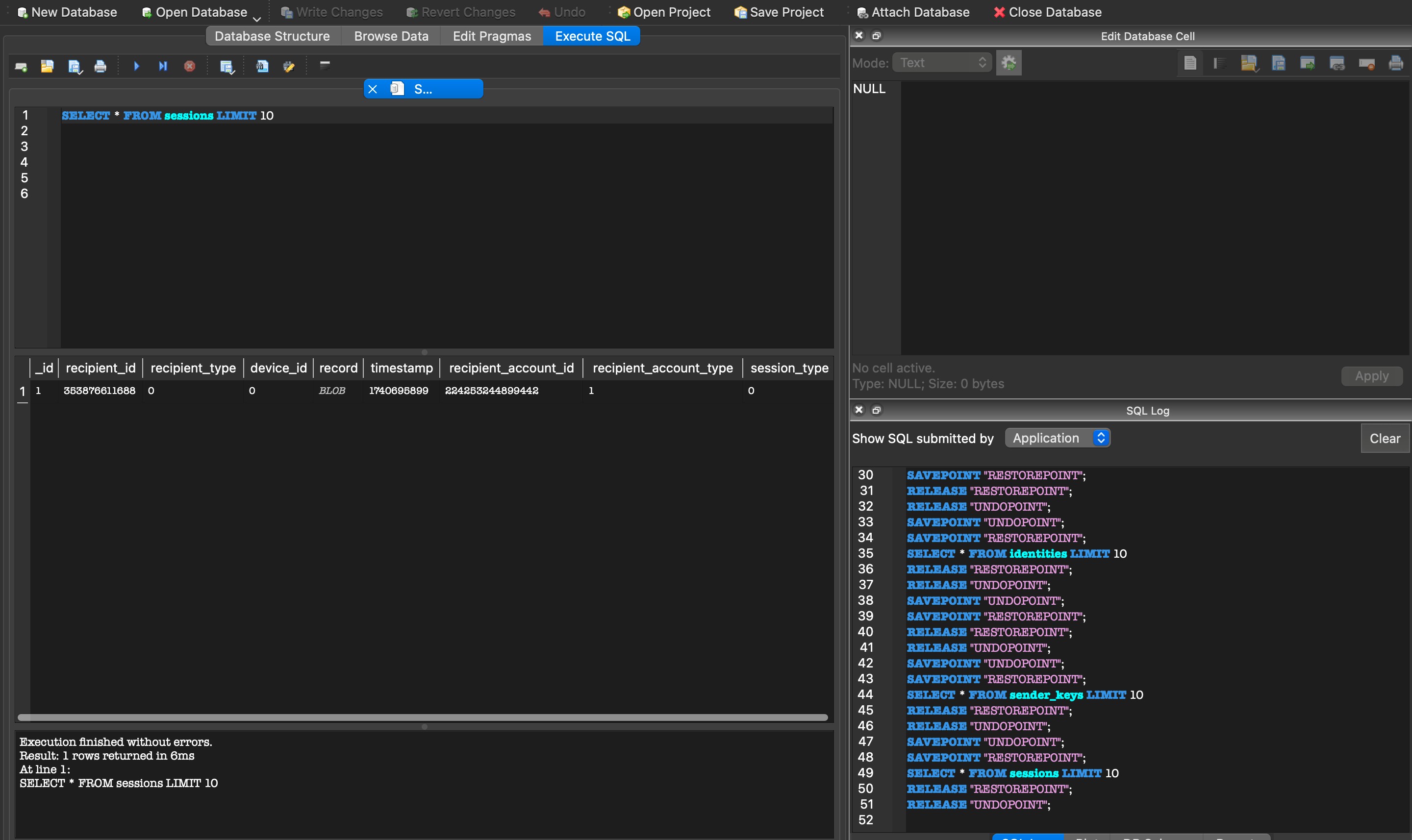
Identities table axolotl.db



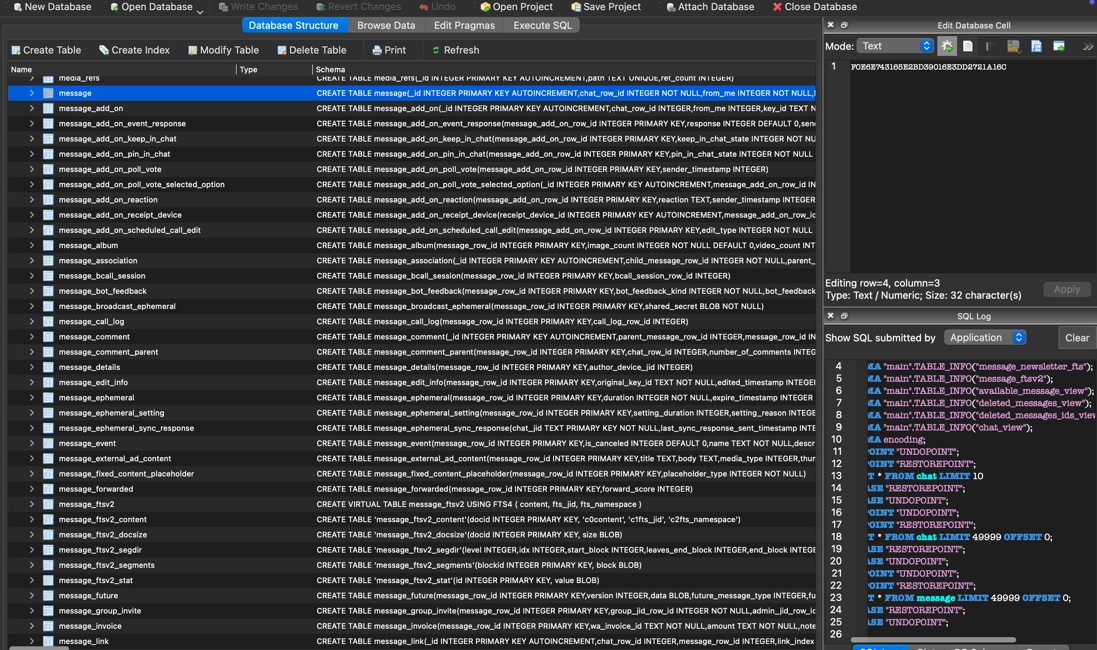
Senders\_key table axolotl.db

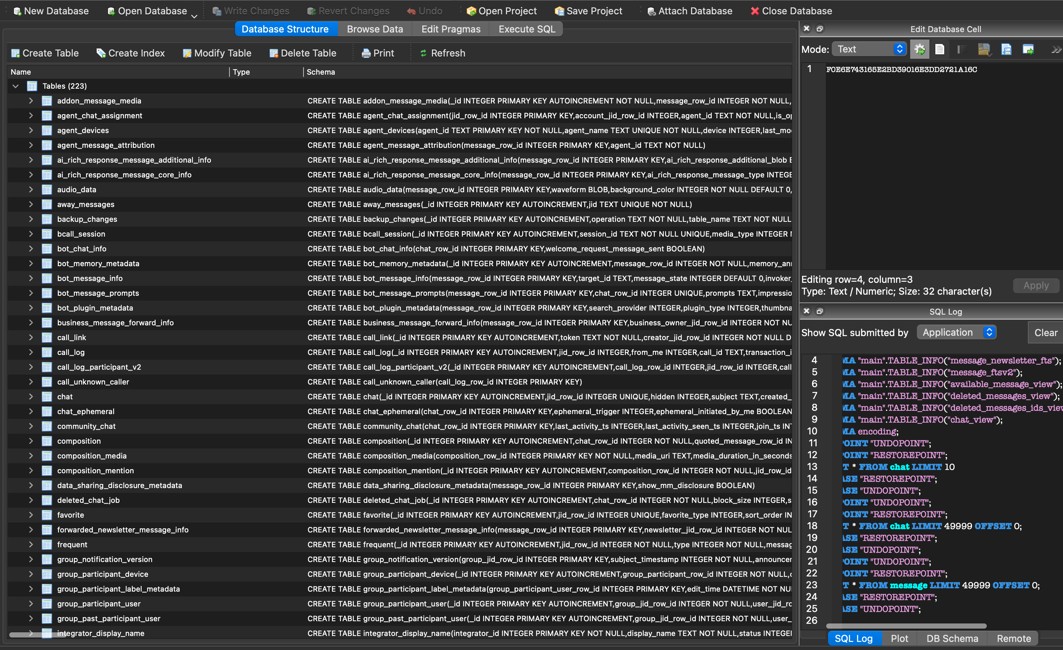


sessions table from axolotl.db

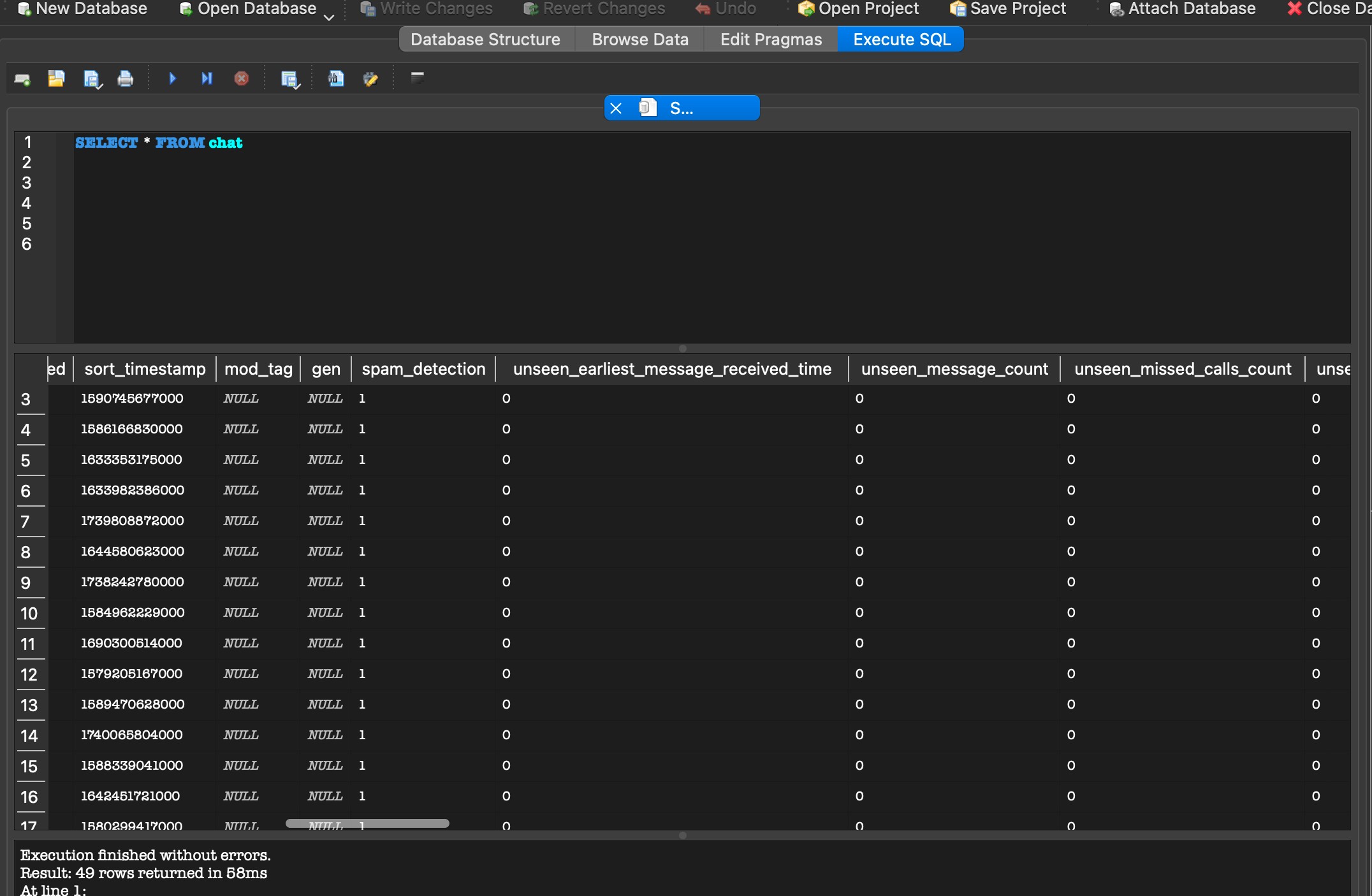
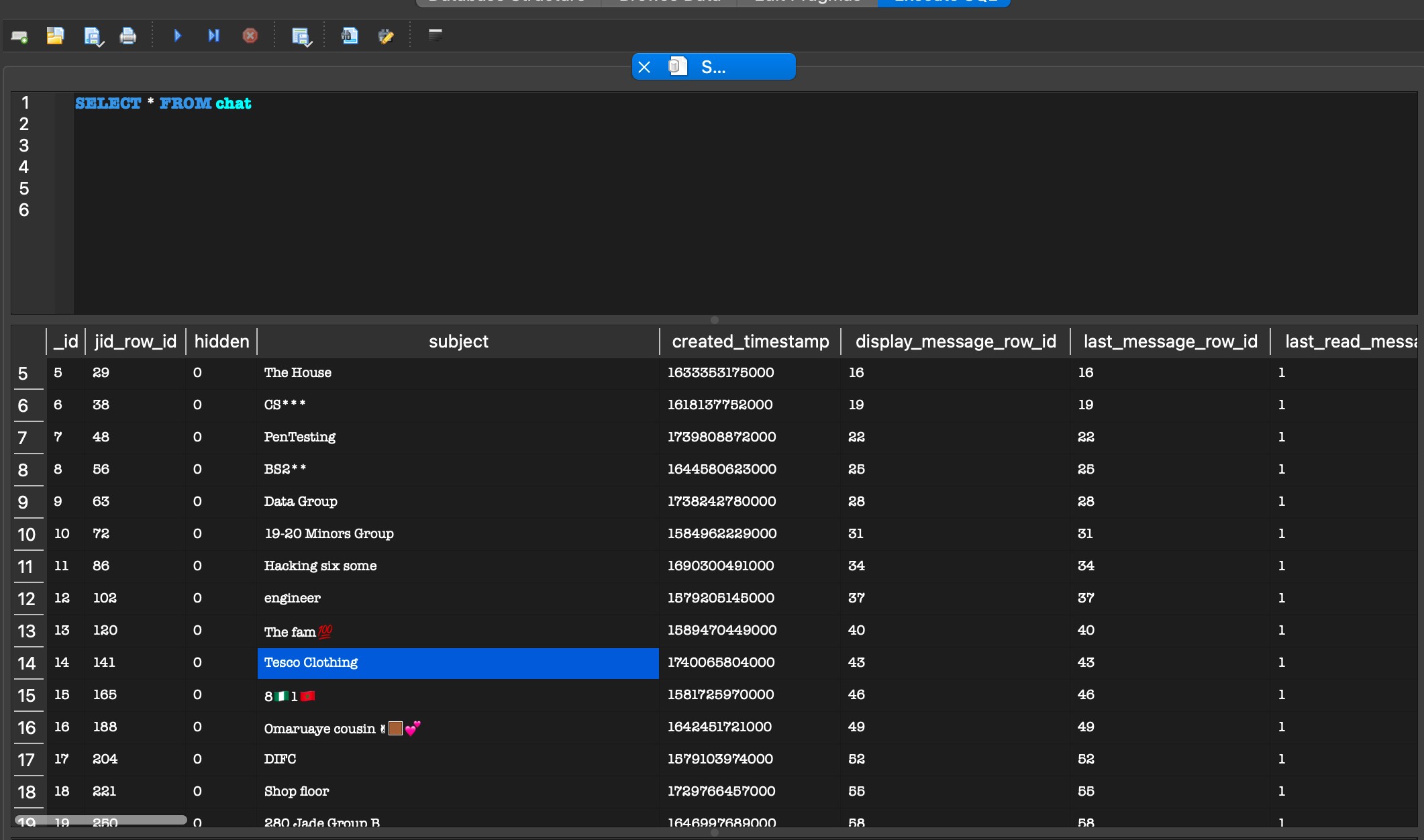


1. Msgstore.db database structure



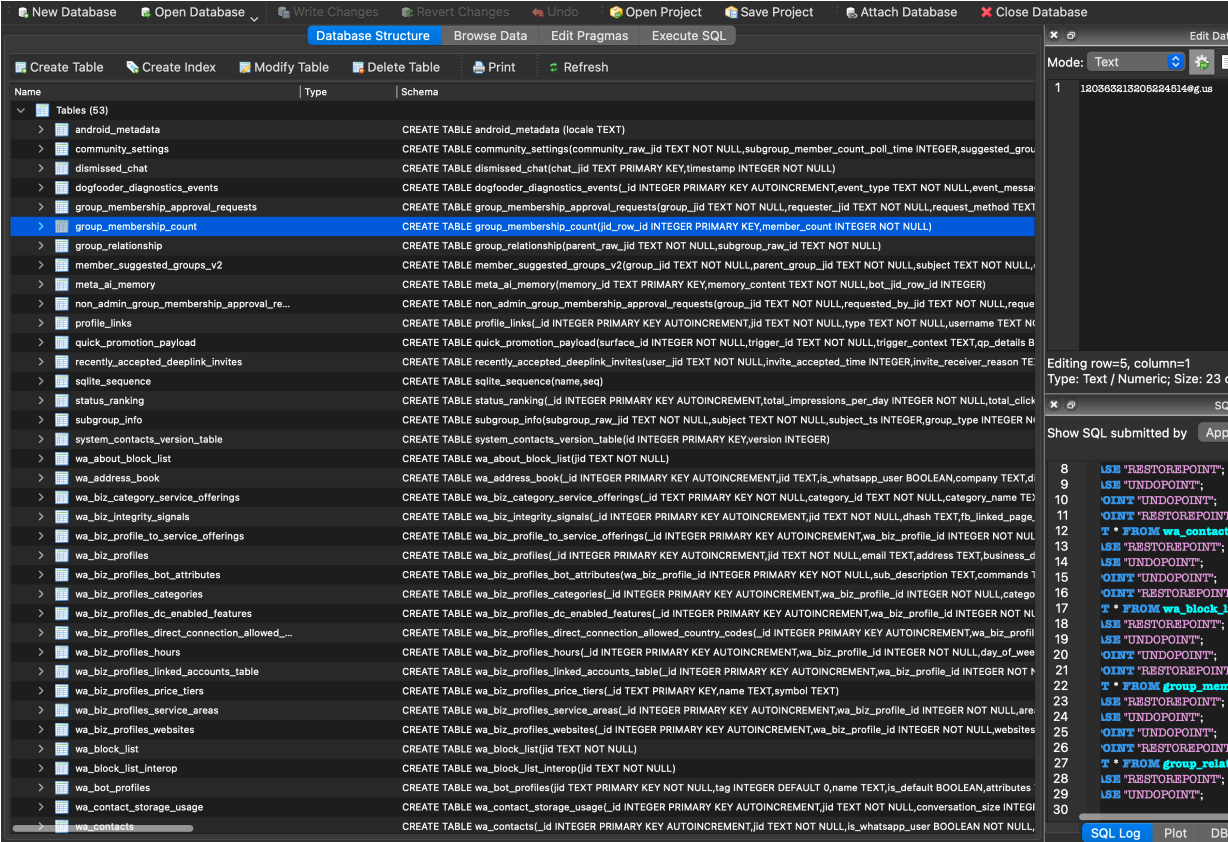


Chat table from msgstore.db

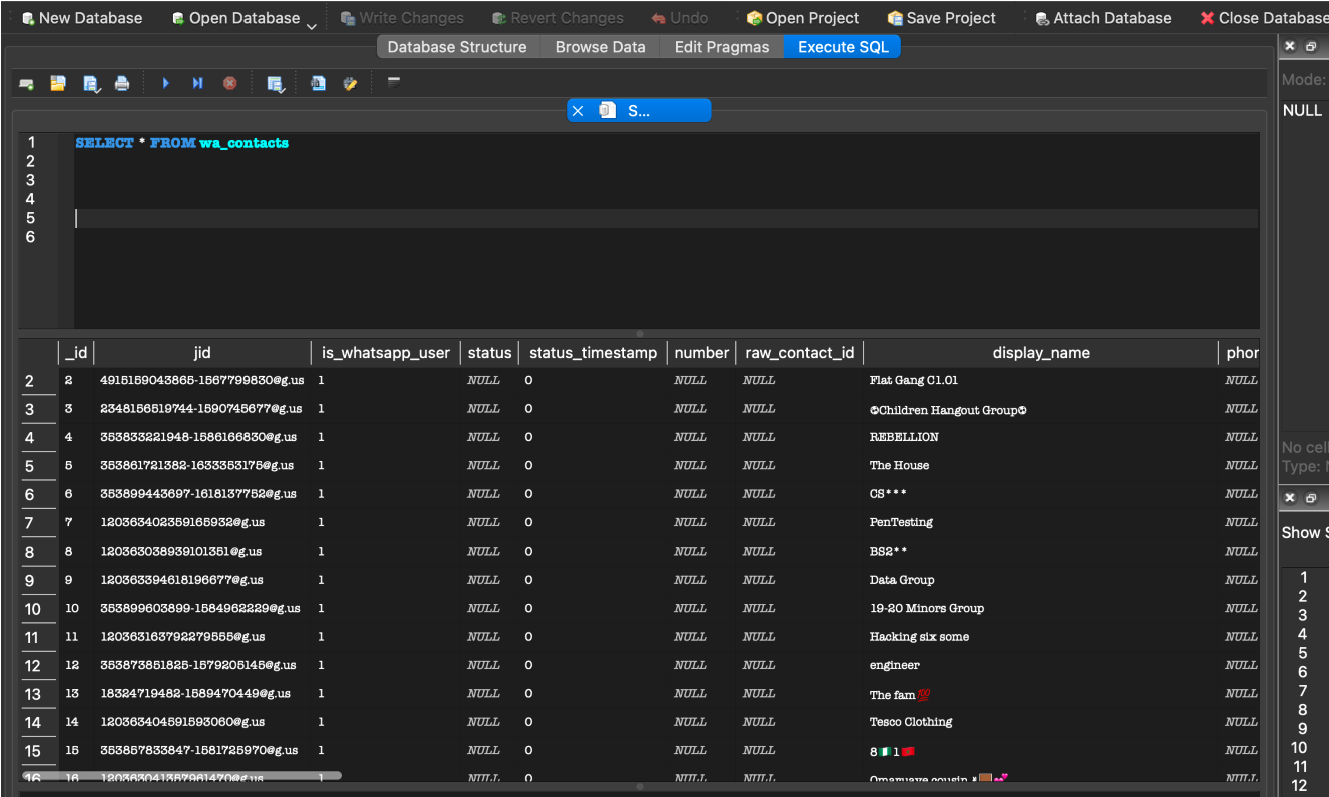




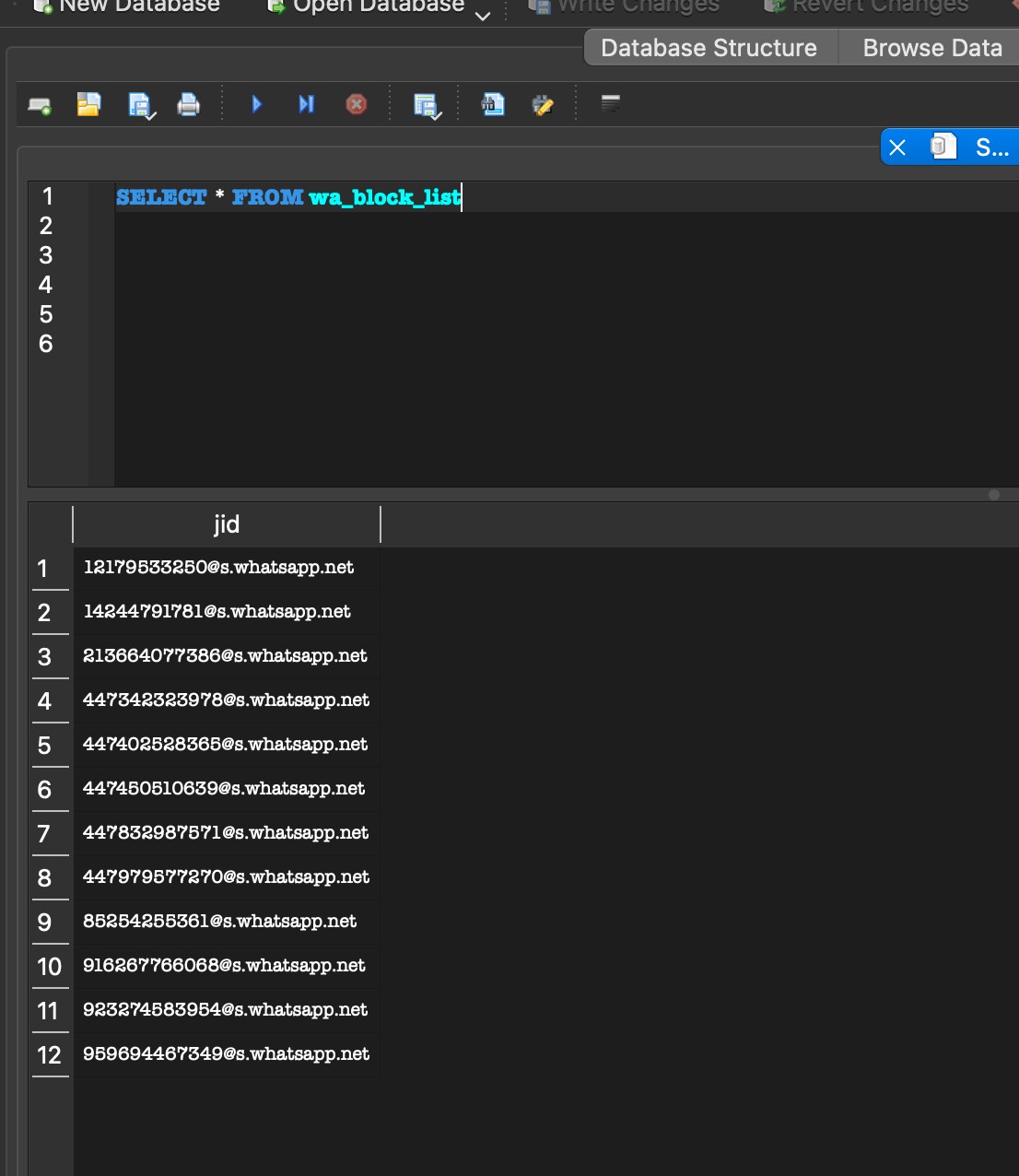
1. Wa\_db structure



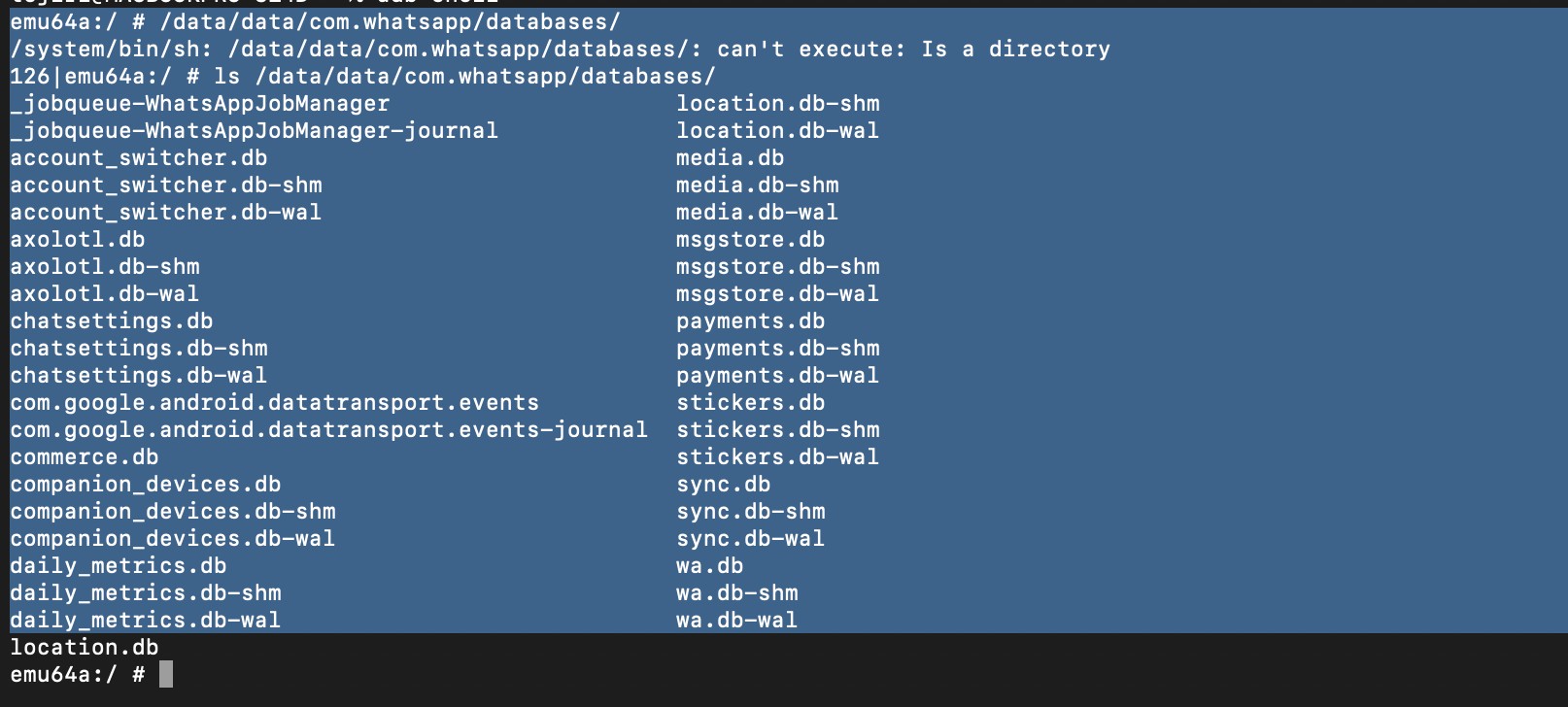
wa\_contacts db from the wa.db



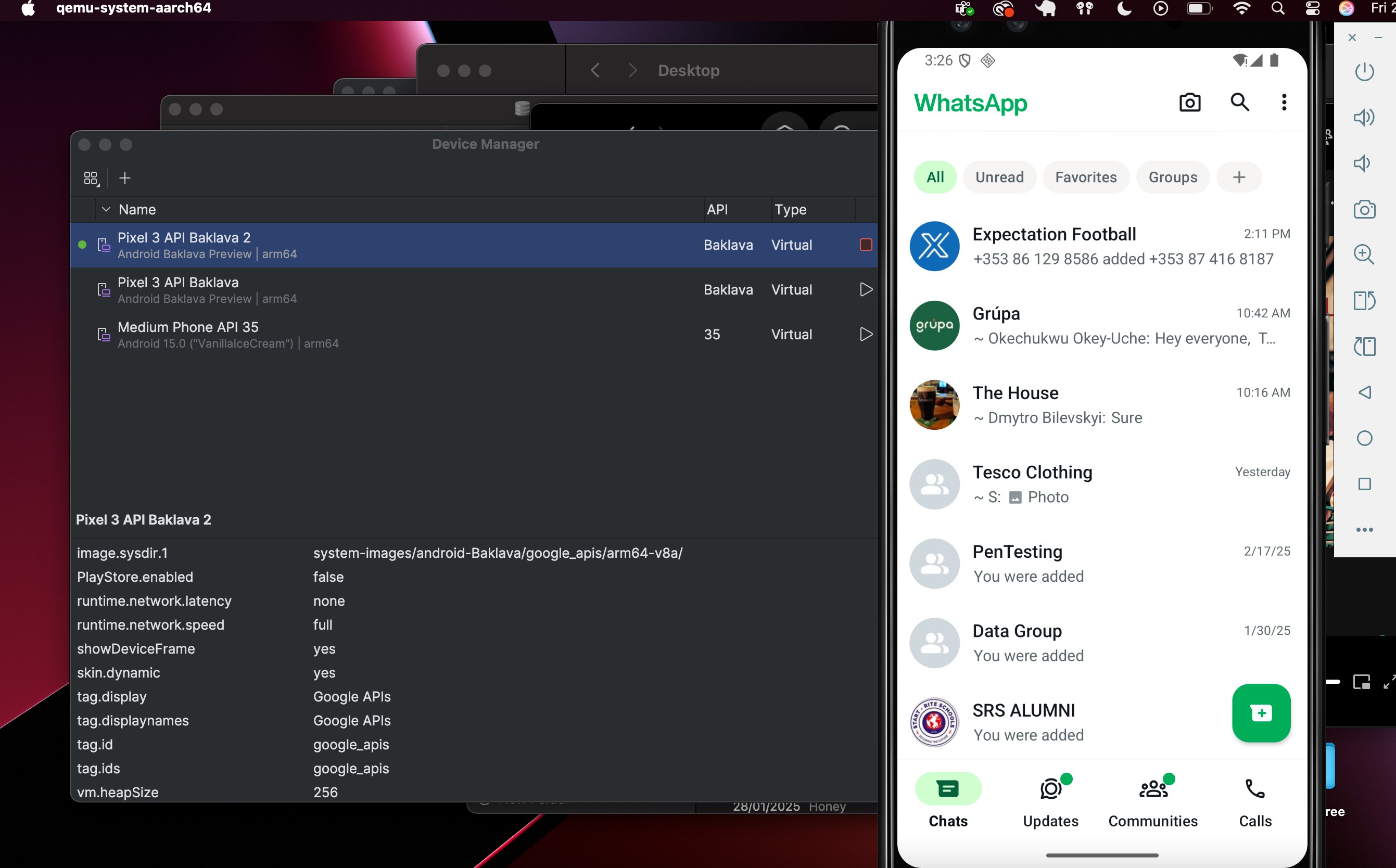
Wa\_block\_list from wa.db



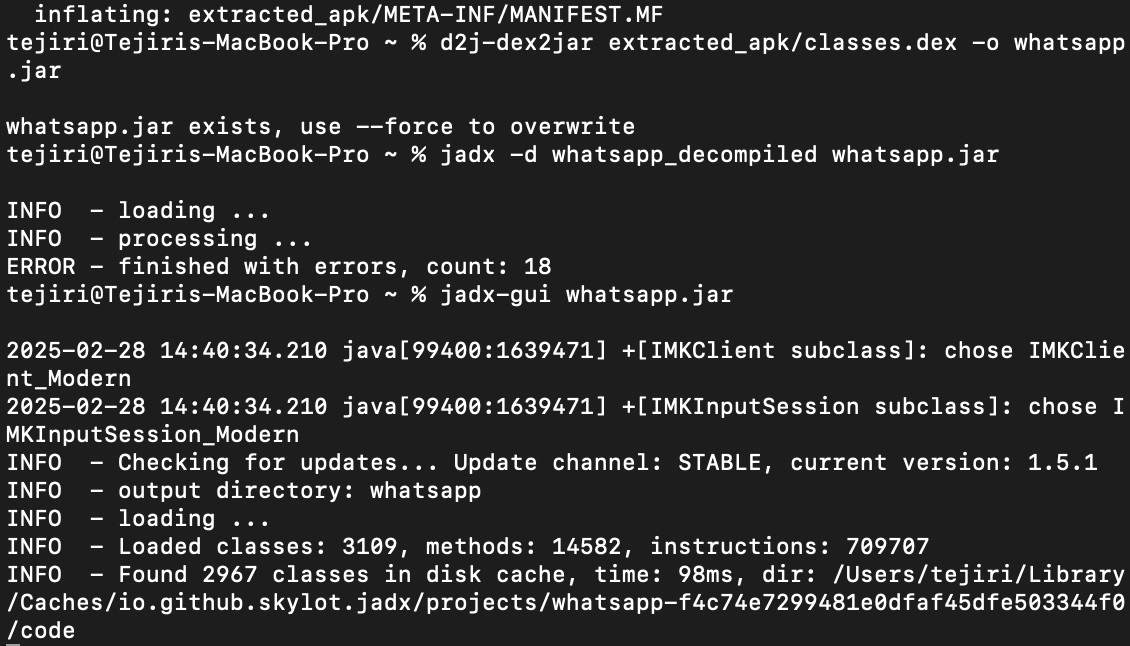
1. List of WhatsApp Databases



1. Android Studio running whatsapp



1. Decompiling the dex2jar file



# Bibliography

Anglano, C., 2014. *Forensic analysis of WhatsApp Messenger on Android smartphones,* Alessandria, Italy: ScienceDirect.

Cressler, C., 2021. *Understanding WhatsApp's Architecture & System Design.*

[Online]

Available at: http://cometchat.com/blog/whatsapps-architecture-and-systemdesign

K. Kaushik and Y. Katara, 2022. *Forensic Analysis of WhatsApp chat data,* Noida, India: IEEE.

Rastogi, N., 2017. *WhatsApp security and role of metadata in preserving privacy,* s.l.: Cornell University.

Richard Ayers (NIST), S. B. (. W. J. (., 2014. *NIST SP 800-101 Rev. 1 Guidelines on Mobile Device Forensics,* s.l.: NIST.

W. Ahmed, F. S. A. R. J. F. I. a. L. A., 2021. *WhatsApp Network Forensics: Discovering the IP Addresses of Suspects,* Paris, France: IEEE Institute of Electrical and Electronics Engineers .

WhatsApp, 2025. *Discover common WhatsApp Business Platform use cases that help drive customer and business growth.* [Online]

Available at: https://business.whatsapp.com/products/business-platform

(Mohammad Shadeed \*, 2022)