# Database Schemas

## Atii\_schema

Contains data from the Anti-Human Trafficking Intelligence Initiative related to crypto. The tables are:

* atii\_tbl: table with crypto and onion addresses
* crypto\_tbl: table with crypto, darknet and crypto money type
* darknet\_url\_tbl: table with CSE urls detected in the darknet

The table atii\_tbl may be connected with the crypto\_tbl.

## Geocomply\_schema

Contains the GeoComply data with info about IP’s. The tables are:

* geoguard\_anonymizers\_tbl: table with information about anonymous IP’s such as if it is vpn, tor, residential proxy, etc
* geoguard\_geoip\_tbl: table with information about geolocation of IP’s such as latitude, longitude, country, etc.

Geocomply data can be merged with any date and IP to get the geolocation. Is important to highlight that the query key is date and IP.

## Main

Contains the CRC data. The tables are:

* crc\_data\_tbl: Main table of CRC data, containing date, ip\_addresses, guid and others
* crc\_geoguard\_tbl: table with CRC IP’s, with IP’s privacy identified using the anonymizers table
* crc\_geoip\_outliers\_tbl: table with list of IP’s that are private and cannot be geolocated
* crc\_geoip\_tbl: table with CRC IP’s geolocated using the geocomply schema
* ips\_blacklist\_tbl: table with CRC list of blacklisted IP’s

## Helpers

Data that helps a process or calculation for Light House. The tables are:

* au\_ips\_tbl: table with the list of Australian IP’s. The source is IP2Location.
* au\_words: table with the list of Australian slangs and locations. The source is Warren.
* dim\_crc\_username\_tbl: table with the list of usernames from CRC table. The source is crc\_data\_tbl.

## IP2Location

Contains from IP2Location data. The table is:

* ip2location\_tbl: table with information about geolocation of IP’s such as latitude, longitude, country, etc.

IP2Location data can be merged with any IP to get the geolocation. Is important to highlight that the query key IP.

## Playground

Contains experimental tables with data that could finally be used in production. These tables are temporary and deleted after they are not used.

## Archives

Contains data used in past projects and not used in current projects. We are keeping them as they may be used again or we need them as backup, but eventually all the tables will be deleted.