THREAT HUNTER RESEARCH

2023 Annual Data Breach Risk Report

An annual view of verified data exposure, illegal trading channels and the response priorities that emerged in 2023.

Original publication: 2024-01-19
Research team: Threat Hunter Research Team
Source report: 32 pages

Original report text

This text version is reconstructed based on the 32-page original PDF, retaining the report narrative, chapters and research scope; the cover, repeated table of contents and purely decorative pages are not repeated. Localized figures are placed in context throughout this web edition, with the original PDF retained for reference.

Preface

In 2023, various industries will be fully digitized, accelerating business innovation and development, while ensuring the data security of a large number of digital assets and cloud businesses has become a "dilemma" faced by many companies. The impact and losses caused by data leakage will further expand, causing serious impact on key areas of various industries.

Threat Hunter released the "2023 Data Breach Risk Annual Report", which conducts a detailed analysis of the data breach risk profile in 2023, the threat-actor data trading market, etc. The data shows:

1. In 2023, more than 19,500 effective data leakage incidents will be monitored, analyzed and verified across the entire network, involving finance, logistics,

More than 20 industries such as aviation and travel, e-commerce, and automobiles;

2. The financial industry surpassed the logistics industry and became the industry with the largest number of citizens' personal information leakage incidents in 2023;

The number of citizen personal information leakage incidents in the aviation industry has also increased significantly, ranking among the top three industries for citizens' personal information leakage for the first time;

3. From the perspective of data leakage channels, they are mainly concentrated in more hidden and convenient anonymous group chats and darknet channels; it is worth mentioning

Yes, more than 50% of data transactions in cases of citizen personal information leakage occurred at night, and more than 30% occurred on non-working days.

Related noun definitions:

1. Data leakage intelligence: Threat Hunter captured "unauthorized personal/organizational sensitive information through anonymous group chats, dark web and other channels"

"Publicly traded or used" intelligence information may include historical data, duplicate data, etc., and is often of huge magnitude;

2. Data leakage incidents: Threat Hunter security research experts analyze and verify samples of data leakage intelligence to confirm

Data breaches that are believed to be real and valid;

3. Citizen personal information: refers to citizens' personal identity information, including but not limited to name, ID number, date of birth, phone number, etc.

Phone number, home address, bank account information, etc.;

4. Historical personal information: refers to personal information that has been leaked before this data leakage incident, a lot of historical personal information

Collected and integrated into a social engineering library by threat actors;

5. Enterprise sensitive code: refers to the enterprise's core codes, algorithms, technologies, passwords and other sensitive information, including software source code.

Code, database structure, API keys, access credentials, encryption algorithms, etc.;

6. Corporate sensitive information: refers to corporate confidential documents and sensitive information, including but not limited to contracts, business plans, financial reports

Tables, market research reports, customer lists, etc.;

7. Darknet: refers to a hidden network that ordinary netizens cannot search and access through conventional means. They need to use some specific software.

Configuration or authorization is required to log in;

8. Private group: A group that can only be entered through an invitation link/administrator's approval. Generally, outsiders cannot monitor or enter.

The group chat.

Data Breach Risk Overview 2023

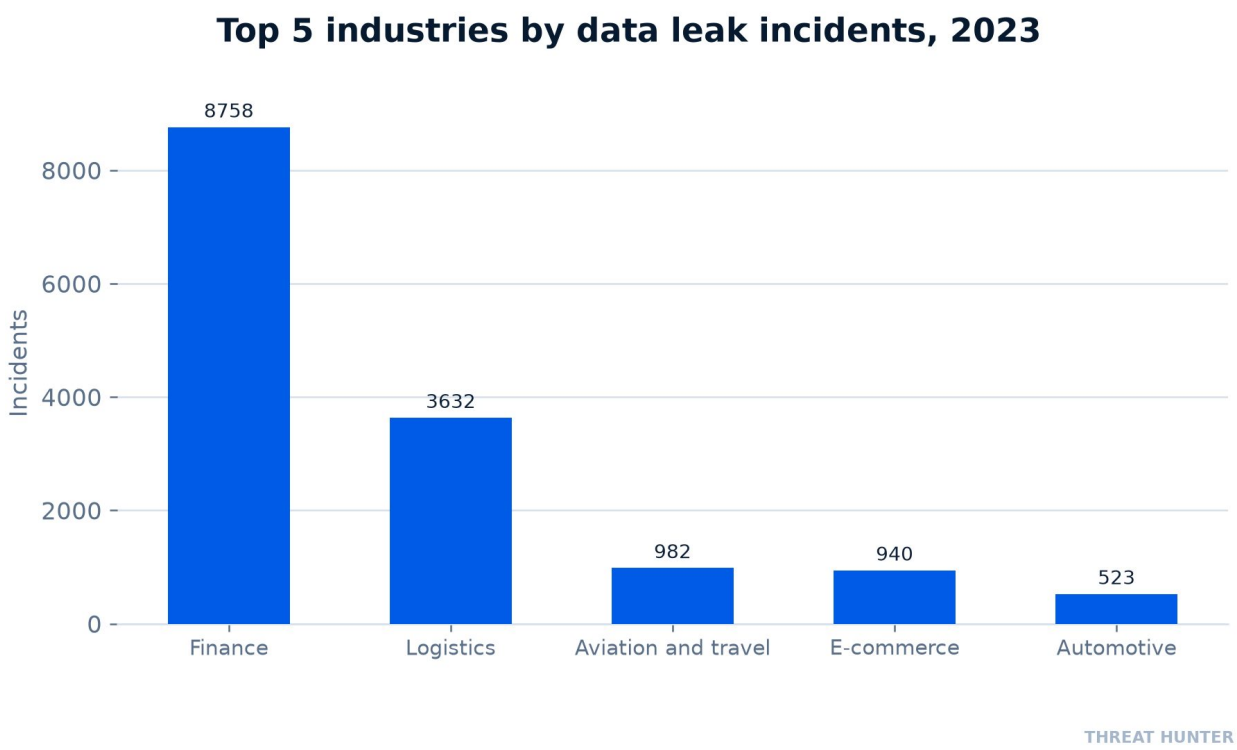
1. Overview of data breach risks in 2023

1.1 There were more than 19,500 monitored data leakage incidents in 2023, and industries such as finance, logistics, and aviation are the hardest hit areas by data leakages.

According to data from the Threat Hunter data leakage risk monitoring platform, among the nearly 150 million pieces of intelligence monitored across the entire network in 2023, more than 19,500 data leakage incidents were analyzed and verified as valid. From the perspective of industry distribution, data breaches in 2023 involve more than 20 industries, and the top five industries with the number of data breaches are finance, logistics, aviation and travel, e-commerce, and automobiles.

1.2 The financial industry ranked first with 8,758 data breaches, and the aviation industry jumped to third

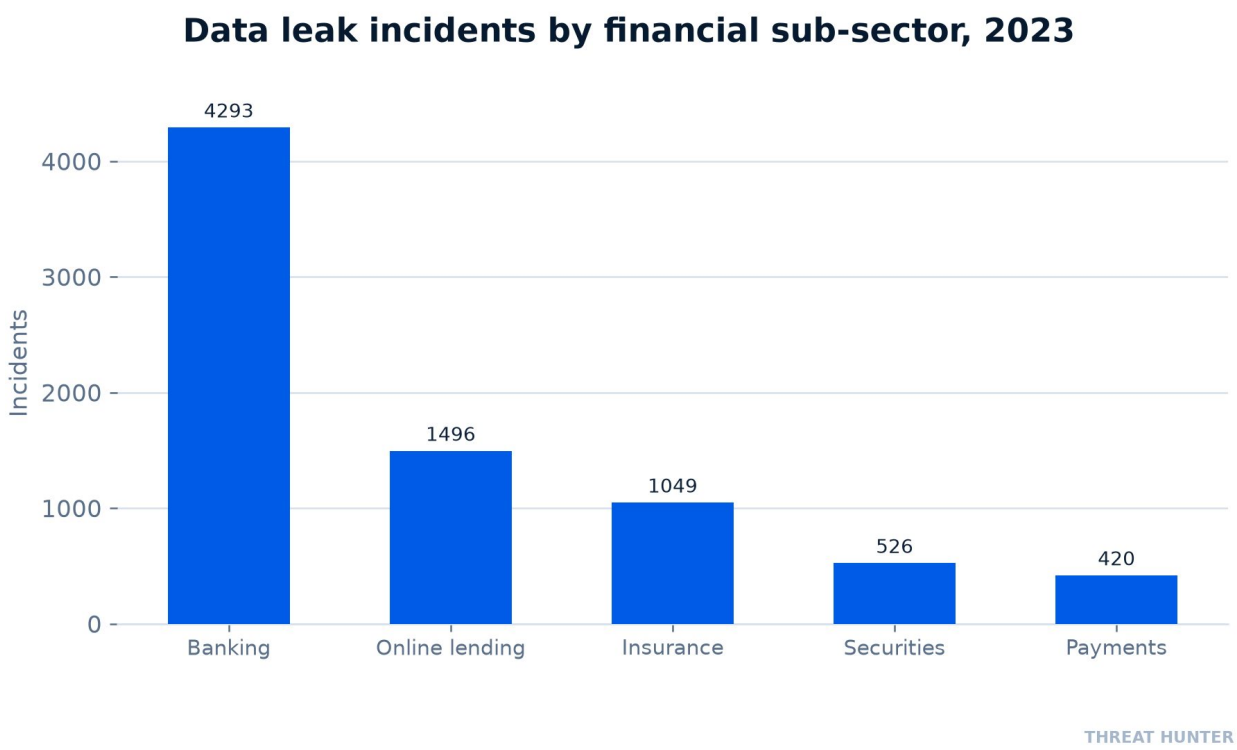
In 2023, the financial industry will still be the hardest hit area for personal information leakage, with 8,758 data leakage incidents involving information of high-net-worth individuals in banking, insurance, securities and other industries. This is mainly due to the higher value of profits used by downstream threat actors for marketing promotion and fraud.



Top 5 industries by data leak incidents in 2023: finance leads with 8,758

Source: Threat Hunter original report, page 6. Redrawn in English from the original figure.

From the perspective of financial sub-sectors, the banking industry has the largest number of data breaches, with a total of 4,293 incidents occurring throughout the year, followed by the online lending, insurance, securities and payment industries.

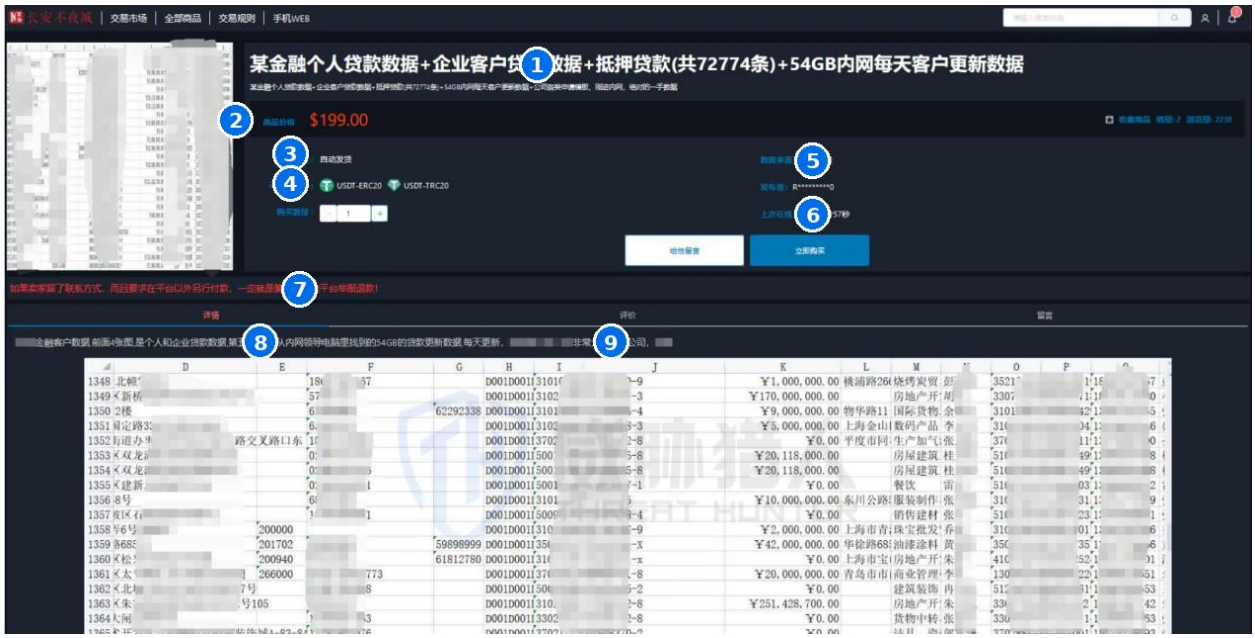


Data leak incidents by financial sub-sector in 2023: banking leads with 4,293

Source: Threat Hunter original report, page 7. Redrawn in English from the original figure.

Typical data breach cases in the financial industry:

Taking a financial company as an example, its user loan information was leaked on a well-known Chinese dark web, with more than 72,000 pieces of data leaked. The number is still updated and increasing every day. threat actors sell it for $199. So far, it has completed 2 transactions and been viewed more than 2,200 times.



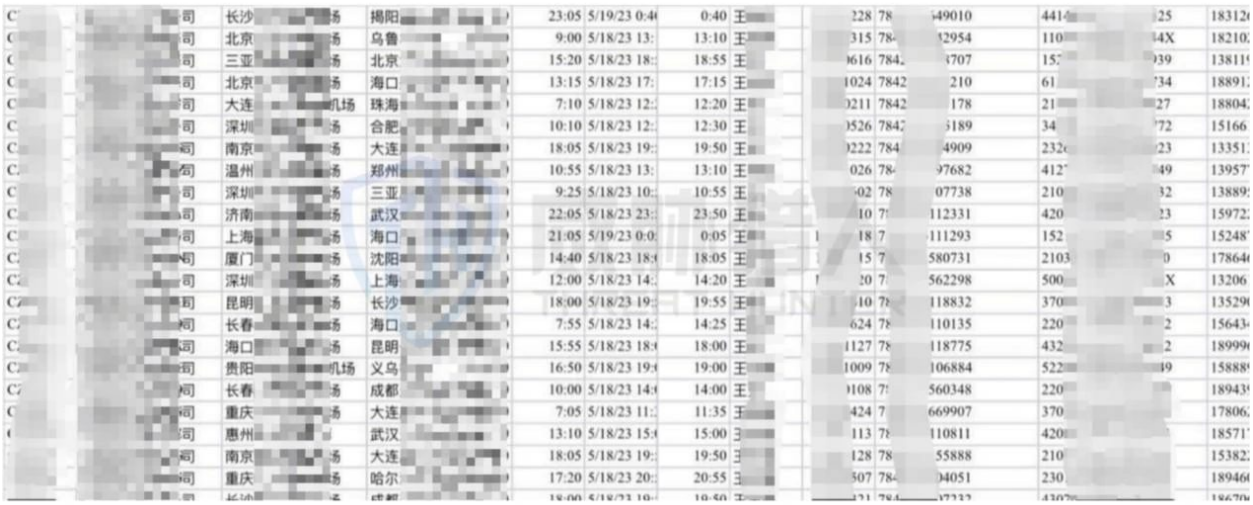
A darknet listing selling a financial firm's loan data and 54 GB of intranet customer data for USD 199

Source: Threat Hunter original report, page 8. Original image retained; numbered translations added by Threat Hunter.

Figure guide

1	A financial firm's personal loan data + corporate loan data + mortgage loans (72,774 records) + 54 GB of daily-updated customer data from its intranet
2	Price
3	Product type: automatic delivery
4	Accepted payment: USDT (ERC20 / TRC20)
5	Data source: self-penetrated
6	Last online
7	If the seller leaves contact details and asks for payment outside the platform, it is a scam; report it to the platform
8	The first four images are the firm's personal and corporate loan data; the fifth is 54 GB of daily-updated loan data found on an executive's intranet computer
9	A very large listed company

The investigation found that hackers crawled the financial company's data through attacks and penetrations. If the company fails to perceive the risk of data leakage in a timely manner and lacks effective defense and disposal measures, it will cause hard-to-eliminate economic losses and brand reputation impacts.



Sample of the leaked loan data posted with the listing (masked in the original)

Source: Threat Hunter original report, page 9. Original image retained.

It is worth mentioning that the aviation industry ranks third and has become one of the areas hardest hit by data breaches. As the COVID-19 epidemic improves, the tourism industry is recovering significantly, consumer demand has been strongly released, and the amount of passenger information such as air ticket information and hotel information has also increased significantly.

Threat Hunter security researchers have observed that in 2023, "cancellation and rebooking" frauds occurred frequently due to flight information leaks. After cybercriminal groups accurately obtained the passenger's name, ID number, flight number and other information, they carried out online fraud by changing the ticket, and the fraud success rate was high.

Typical data breach cases in the aviation industry:

The Threat Hunter data leakage risk monitoring platform captured sensitive information such as "real-time tickets without registration" released by threat actors in an encrypted-messaging data trading group "Hacker Orders", including passenger names, phone numbers, ID numbers, flight information and other sensitive information.



A chat with a data seller about the price of flight passenger data

Source: Threat Hunter original report, page 9. Original image retained; numbered translations added by Threat Hunter.

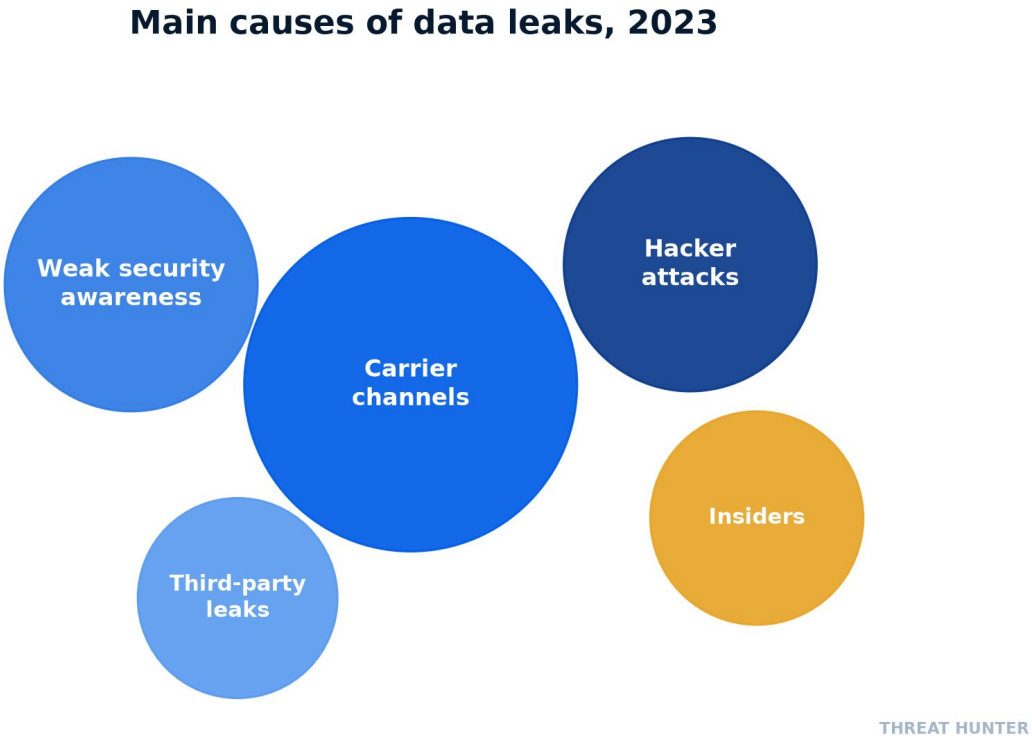
Figure guide

1	What are the prices for flight data now?
2	That expensive?
3	It has to be real data
4	Do you think takeaway data selling for 2–3 yuan is real?
5	Still, not as low as 2–3 yuan
6	Before, I got it from people for about 10 yuan each
7	Depends how many you take
8	My lowest price is 13
9	None of mine is fake

Further analysis found that the passenger data involved multiple airlines, and was initially determined to be leaked by a common partner of multiple airlines. It is understood that the transaction price of personal information of aviation and travel users mostly ranges from 13 to 15 yuan, with the highest price being 20 yuan.

1.3 Causes of data leakage include operator channel leakage, insider leakage, hacker attacks, etc.

Judging from the specific causes of data leaks, the causes of data leaks in 2023 include operator channel leaks, insider leaks, hacker attacks, security awareness issues, etc. Among them, the number of data leakage incidents caused by operator channel leakage is the largest.



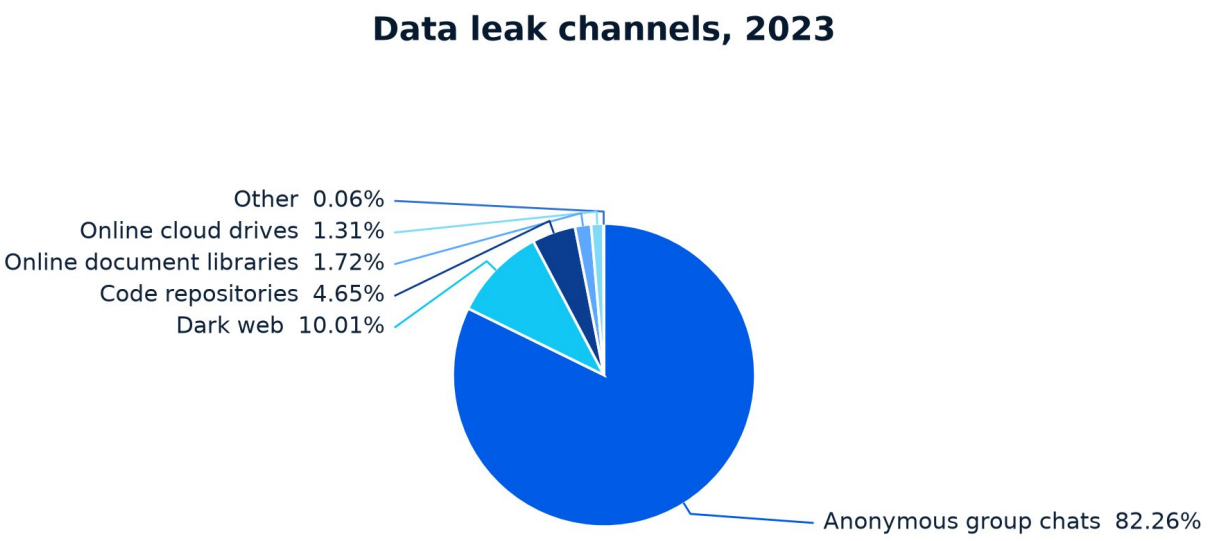
Main causes of data leaks in 2023: carrier channels, insiders, hacking, weak security awareness and third parties

Source: Threat Hunter original report, page 10. Redrawn in English from the original figure.

- Operator channel: threat actors obtain access data of specified web pages, installation data of specified applications, receiving and sending data of specified text messages and other information through channels such as operators' insiders or illegal agents;
- Insider leaks: Internal employees of the company, driven by interests, use methods such as data export and manual photography to obtain sensitive customer information and then sell it;
- Hacker attack: External hackers use crawlers, scanning, penetration and other methods to attack enterprise systems and network assets, and exploit enterprise network vulnerabilities to steal data on a large scale;
- Security awareness issues: During the work process, internal employees of the company unconsciously upload important files, documents, codes, etc. within the company to public network environments such as network disk libraries and code hosting platforms, resulting in the leakage of sensitive information;
- Third-party leakage: A third party that has a cooperative relationship with an enterprise has the authority to access certain sensitive data of the enterprise. However, due to problems such as irregular management, these sensitive data are leaked to threat actors through the third party;

1.4 Anonymous group chats and the dark web are the main channels for data leakage, accounting for 92%

Among the data breaches monitored by Threat Hunter in 2023, more than 92% occurred in anonymous group chats and on the dark web, of which



Data leak channels in 2023: anonymous group chats account for 82.26% and the dark web for 10.01%

Source: Threat Hunter original report, page 11. Redrawn in English from the original figure.

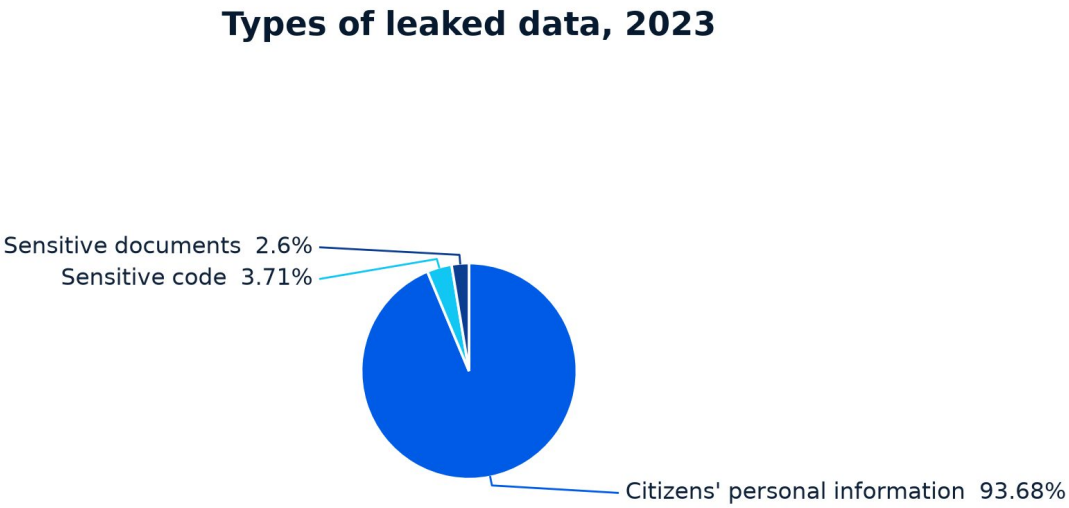
82.26% is concentrated in anonymous group chats, and 10.01% occurs on the dark web. The main reason is that anonymous group chats and dark web channels are concealed.

It is highly reliable and difficult to trace back to threat actors themselves. It is the preferred channel for threat actors to communicate and trade. In addition, Threat Hunter also detected data leakage incidents in code warehouses (such as code-hosting and API-collaboration platforms), network disk libraries and other channels.

As of December 2023, Threat Hunter data leakage monitoring intelligence covered nearly 20,000 channels/group chats on encrypted messaging apps, and risk incidents of citizens' personal information leakage were found in more than 1,700 channels/group chats.

1.5 "Citizen personal information" will still be the main type of data leakage in 2023, accounting for more than 90%

Judging from the types of data leaks, there are three main types of data leaks in 2023: 18,347 cases of citizen personal information (93.68%), 727 cases of sensitive code (3.71%), and 510 cases of sensitive documents (2.6%).



Types of leaked data in 2023: citizens' personal information accounts for 93.68%

Source: Threat Hunter original report, page 12. Redrawn in English from the original figure.

Among them, the data leakage of citizens’ personal information accounted for the highest proportion, reaching 93.68%. As the main data type of downstream cybercrime ecosystem, the following will provide a detailed analysis of the data leakage of citizens’ personal information.

1.6 Characteristics and trends of citizens’ personal information leakage in 2023

1.6.1 More than 80% of citizens’ personal information containing “phone numbers” was leaked, mainly used for targeted marketing/fraudulent crimes

In 2023, there were more than 18,000 leaks of citizens' personal information, of which more than 80% of the leaked information fields contained "phone numbers", which were mainly used by downstream marketing/fraud gangs to send text messages and telemarketing calls to relevant phone numbers to further commit illegal and malicious activity acts.

Among citizens’ personal information leakage incidents in 2023, data field combinations such as “name + phone number + ID number + bank card number” appeared most frequently, with nearly 900 related data leakage incidents.

Judging from the high frequency of such data field combinations, on the one hand, downstream threat actor groups can carry out targeted malicious activity based on complete data fields and citizen portrait information, making the overall malicious activity success rate and revenue higher;

On the other hand, some upstream threat-actor groups can only obtain the "phone number" field due to limited technical means. The threat-actor groups will use a variety of methods to splice personal information and complete data field information, thereby increasing the value of the data.

Leaked fields top 10:

Leaked field	Incidents
Phone number	14,425
Name	12,731
ID number	7,095
Address	5,614
Bank card number	2,702
Waybill number	2,297
Product details	2,260
SMS content	1,910
Carrier	799
Time	459

Top 10 leaked data fields in 2023: phone numbers appear in 14,425 incidents

Source: Threat Hunter original report, page 13. Table re-typeset in English from the original.

1.6.2 There is a frequent phenomenon of threat actors re-splicing "historical personal data information" and selling it multiple times.

After Threat Hunter investigated the trafficking of citizens' personal information, it was discovered that the black market for data trading has the phenomenon of historical data information being re-integrated by threat actors and sold multiple times. Although the data has been used by buyers for malicious purposes before, and the effect of secondary malicious activity is obviously reduced, threat actors will sell it at a lower price and make profits.



姓名	身份证号	出生年	性别	地区	地区	运营商	地址	短信
*	身份证34082211	1997	男	安徽	合肥	**	安徽省,安庆市	【**证券】尊敬的用户，您通过**证券买入理财1500.00元。
*	身份证34262511	1988	女	安徽	合肥	**	安徽省,马鞍山	【**证券】尊敬的用户，您通过**证券零用钱转入理财金95.00元。
*	身份证34240111	1989	女	安徽	六安	**	安徽省,六安市	【**证券】尊敬的用户，您通过**证券零用钱转入理财金5000.00元。
*	身份证34122511	1993	男	安徽	阜阳	**	安徽省,阜阳市	【**证券】尊敬的用户，您通过**证券零用钱转入理财金4800.00元。
*	身份证34212611	1974	男	安徽	亳州	**	安徽省,亳州市	【**证券】尊敬的用户，您通过**证券零用钱转入理财金100.00元。
*	身份证34122311	1989	女	安徽	亳州	**	安徽省,亳州市	【**证券】尊敬的用户，您通过**证券零用钱转入理财金1000.00元。
*	身份证34262311	1987	男	安徽	亳州	**	安徽省_巢湖地	【**证券】尊敬的用户，您通过**证券零用钱转入理财金3449.40元。
*	身份证34120411	1997	男	安徽	合肥	**	安徽省,阜阳市	【**证券】尊敬的用户，您通过**证券零用钱转入理财金1658.00元。
*	[illegible]	[illegible]	[illegible]	[illegible]	亳州	**	安徽省_亳州市	【**证券】尊敬的用户，您通过**证券零用钱转入理财金100.00元。
*	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]州	**	安徽省,池州市	【**证券】尊敬的用户，您通过**证券买入理财30000.00元。
*	身份证34082711	1991	男	安徽	合肥	**	安徽省,安庆市	【**证券】尊敬的用户，您通过**证券零用钱转入理财金330.00元。
*	身份证34242611	1985	女	安徽	合肥	**	安徽省,六安市	【**银行】您尾号2946的账户于12月09日16:06理财收益兑付(收)4924.40元
*	身份证34262311	1988	男	安徽	芜湖	**	安徽省,芜湖市	【**证券】尊敬的用户，您通过**证券零用钱转入理财金8000.00元。
*	身份证41112311	1997	男	安徽	芜湖	**	河南省,漯河市	【**证券】尊敬的用户，您通过**证券零用钱转入理财金95.00元。
*	身份证34010211	1969	女	安徽	合肥	**	安徽省,合肥市	【**银行】您尾号6009的账户于11月21日16:42理财本金兑付(收)20000.00元
*	身份证22030211	1985	男	安徽	合肥	**	吉林省,四平市	【**证券】尊敬的用户，您通过**证券买入理财100000.00元。
*	身份证34012211	1988	男	安徽	合肥	**	安徽省,合肥市	【**证券】尊敬的用户，您通过**证券零用钱转入理财金5.00元。
*	身份证37078511	1983	男	安徽	合肥	**	山东省,潍坊市	【**证券】尊敬的用户，您通过**证券买入理财50000.00元。
*	身份证34260111	1992	女	安徽	合肥	**	安徽省,合肥市	【**银行】您尾号4199的账户于12月27日17:02理财收益兑付(收)3089.59元
*	身份证34052111	1986	男	安徽	合肥	**	安徽省,马鞍山	【**证券】尊敬的用户，您通过**证券买入理财10000.00元。
*	身份证43068211	1988	男	安徽	合肥	**	湖南省,岳阳市	【**证券】尊敬的用户，您通过**证券零用钱转入理财金1900.00元。

Sample of re-spliced personal data offered for sale (masked in the original)

Source: Threat Hunter original report, page 14. Original image retained.

First, threat actors obtain preliminary citizens' personal information (such as phone numbers) through some illegal channels, and then use social engineering databases containing historical leaked information and other channels to refine the data and supplement the integrity of data fields, thereby increasing the value and profitability of citizens' personal information. Specific methods include:

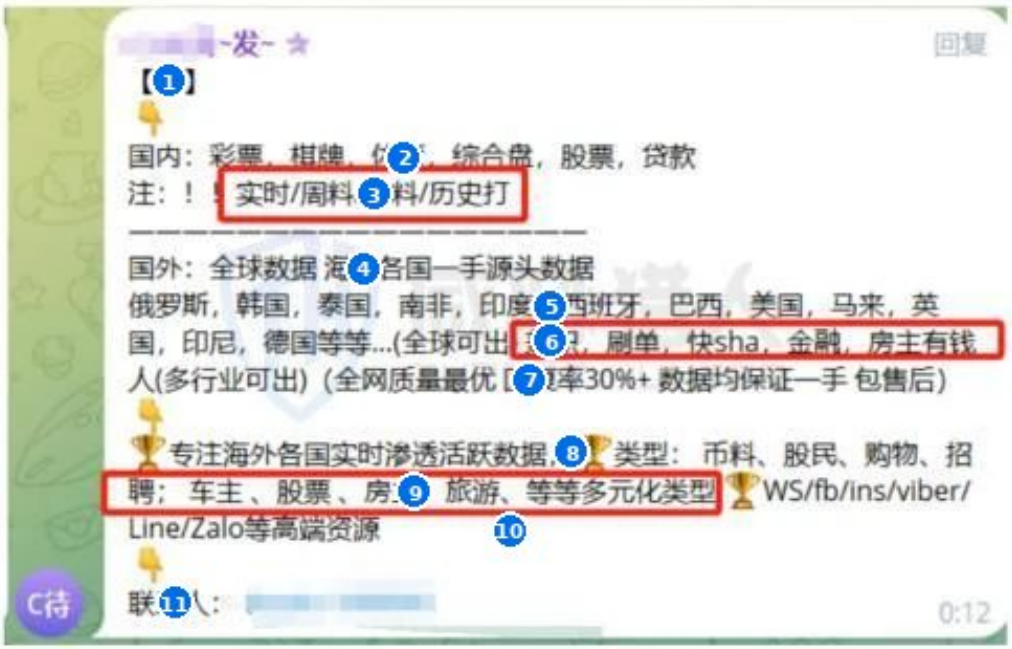
1. Query the specific location and operator information of the corresponding phone number through public channels such as browsers;

2. Use the social engineering database to further query the identity information of the account owner through the phone number;

3. For users of specific platform websites, intermediaries perform verification queries, clean the data and filter out invalid phone numbers;

4. Use public interfaces provided by a phone-platform vendor and automated script tools to fully automatically detect whether the number is registered.

Checking built-in messaging and video-call functions to identify iOS users makes the data more complete and increase the value of the data.



A seller advertising domestic and international personal data by industry and country

Source: Threat Hunter original report, page 15. Original image retained; numbered translations added by Threat Hunter.

Figure guide

1	[Supply]
2	Domestic: lottery, card games, sports, mixed betting, stocks, loans
3	Note: real-time, weekly, monthly and historical data
4	International: first-hand source data from countries worldwide
5	Russia, South Korea, Thailand, South Africa, India, Spain, Brazil, US, Malaysia, UK,
6	Indonesia, Germany, etc. Worldwide: part-time order brushers, quick-loan users, finance, wealthy homeowners
7	Many industries; best quality on the network, 30%+ reply rate, all first-hand with after-sales
8	Focus on real-time active data from other countries: crypto holders, stock investors, shoppers, job seekers
9	car owners, stock investors, homeowners, travellers and more
10	Contacts on several international messaging and social apps
11	Contact



A number-checking service profile offering active-number screening and a verification-code API

Source: Threat Hunter original report, page 15. Original image retained; numbered translations added by Threat Hunter.

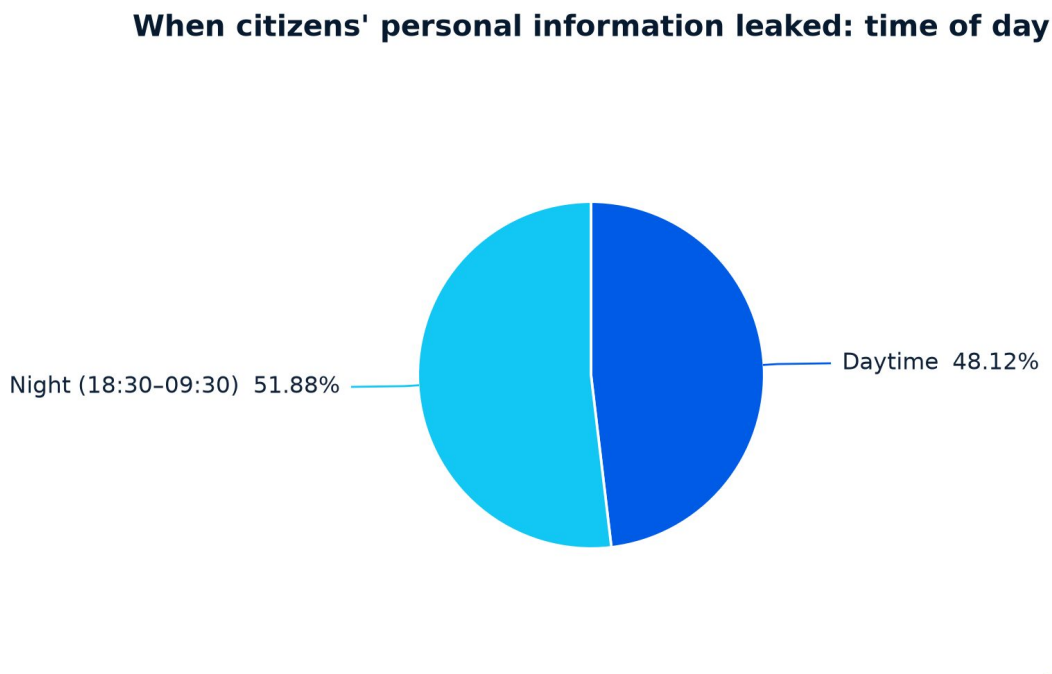
Figure guide

1	"Canghai International" number checking // active-number screening
2	Recently online
3	Worldwide (domestic and international) number checking: filter invalid numbers, screen active ones
4	Targeted fine screening // verification-code API // permanent customer service via a messaging app
5	Bio

1.6.3 More than 50% of citizens' personal information is traded at night, and more than 30% is traded on non-working days.

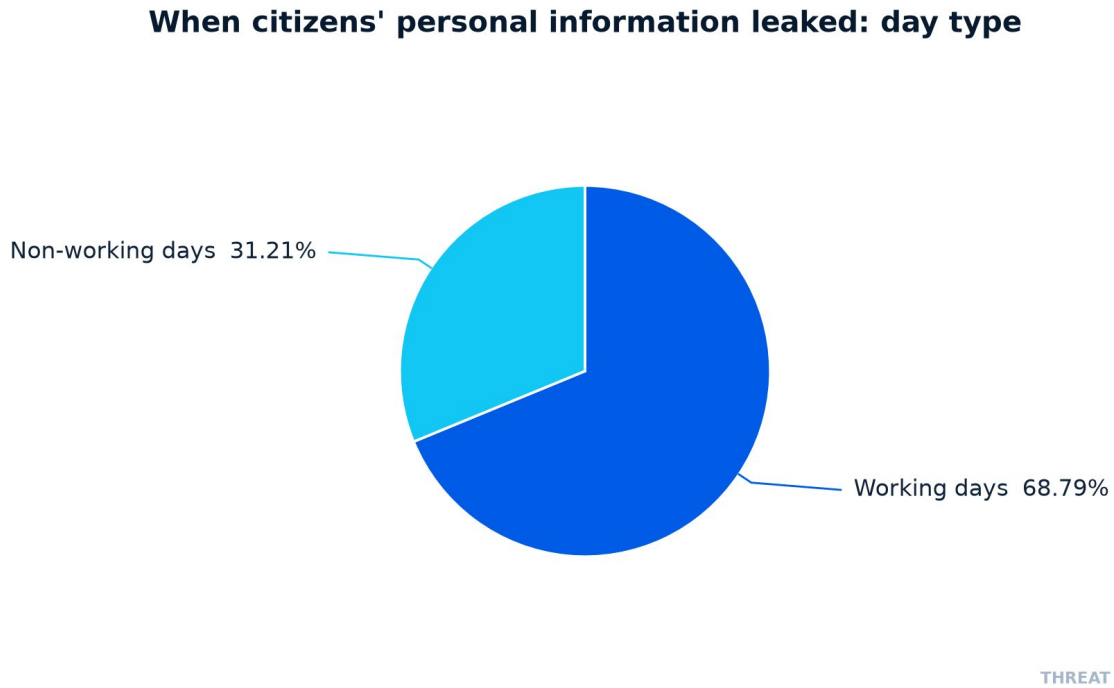
Threat Hunter research statistics found that among data transaction times in citizen personal information leakage incidents in 2023, the number of incidents that occurred on non-working days (weekends and holidays) was as high as 31.21%, and the number of incidents that occurred at night accounted for as high as 51.88%, more than half.

(Night: 18:30 to 09:30 the next day)



Time of day when citizens' personal information was leaked in 2023: 51.88% at night

Source: Threat Hunter original report, page 15. Redrawn in English from the original figure.



Day of the week when citizens' personal information was leaked in 2023: 31.21% on non-working days

Source: Threat Hunter original report, page 15. Redrawn in English from the original figure.

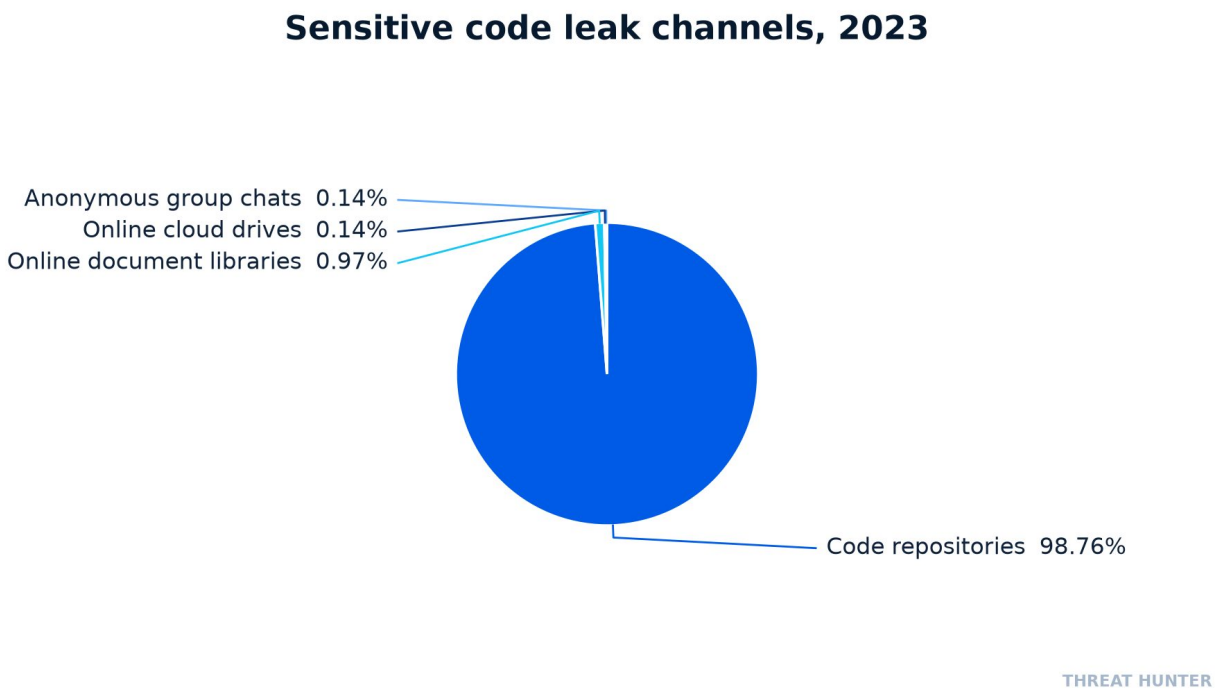
When the defense is at its weakest, it is difficult for enterprises to quickly sense and respond in a timely manner when a data breach occurs, so that they miss the best response opportunity, which has a significant impact on corporate funds, brand reputation and business competition.

1.7 In addition to citizens' personal information leaks, there were more than 1,200 leaks of sensitive code, data files and other information in 2023

Among the types of data breaches in 2023, in addition to citizens' personal information, there are also information types such as corporate sensitive codes and sensitive data files.

1.7.1 The main channels for sensitive code leakage are code warehouses, including database passwords, source code and other information

Among them, more than 727 cases of corporate sensitive code leaks were discovered, mainly concentrated on code-hosting platforms, accounting for 98.76% of the total.

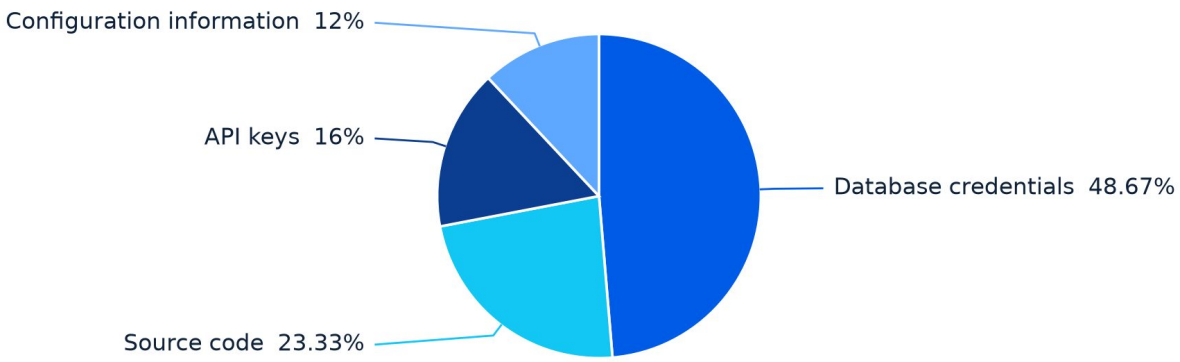


Sensitive code leak channels in 2023: code repositories account for 98.76%

Source: Threat Hunter original report, page 16. Redrawn in English from the original figure.

The content types of sensitive code leaks are mainly database account passwords and source code information, including intranet and public network account passwords, etc. Since a large amount of internal sensitive information is involved, the impact and losses caused by sensitive code leaks cannot be underestimated.

Information types in sensitive code leaks, 2023



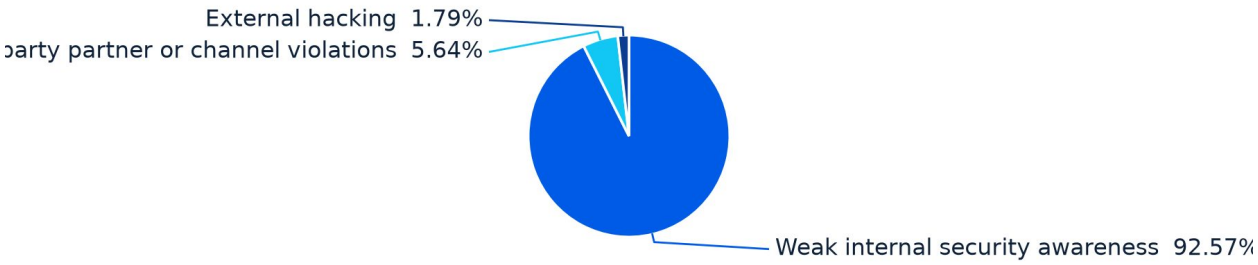
THREAT HUNTER

Types of information in sensitive code leaks, 2023: database credentials lead at 48.67%

Source: Threat Hunter original report, page 17. Redrawn in English from the original figure.

The reasons for the leakage of sensitive code are mainly attributed to the security awareness of internal employees, leakage of third-party cooperation and external hacker attacks. This is different from the reasons for malicious purposes transactions of citizens' personal information data.

Causes of sensitive code leaks, 2023



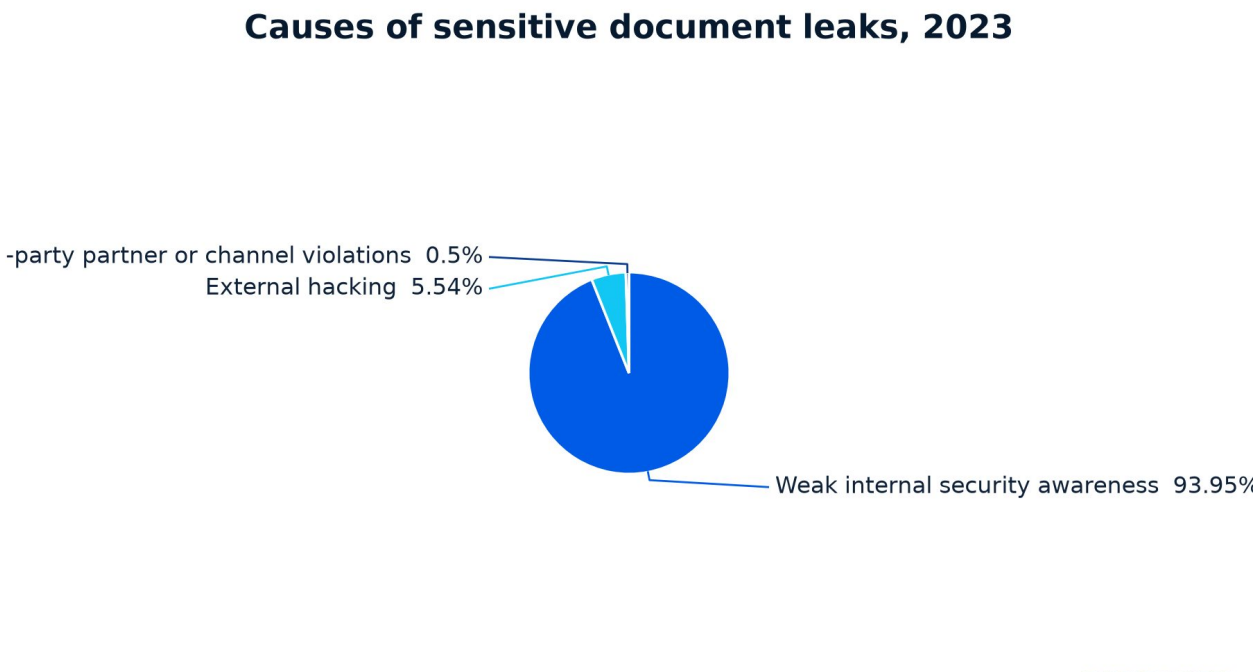
THREAT HUNTER

Causes of sensitive code leaks in 2023: weak internal security awareness accounts for 92.57%

Source: Threat Hunter original report, page 17. Redrawn in English from the original figure.

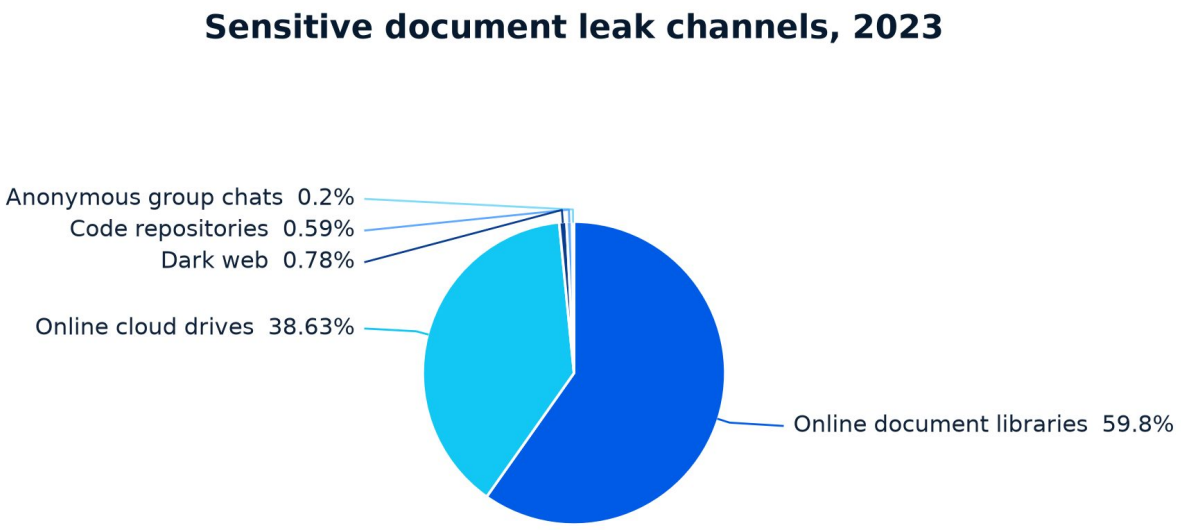
1.7.2 The main channels for sensitive data leakage are online libraries/cloud disks, which are mainly caused by internal security awareness issues within the company.

More than 510 corporate sensitive data leakage incidents were discovered, mainly caused by internal security awareness issues within the company, such as mistaken uploading to online cloud disks and libraries.



Main causes of sensitive document leaks in 2023: weak internal security awareness accounts for 93.95%

Source: Threat Hunter original report, page 18. Redrawn in English from the original figure.



Sensitive document leak channels in 2023: online document libraries 59.8% and cloud drives 38.63%

Source: Threat Hunter original report, page 18. Redrawn in English from the original figure.

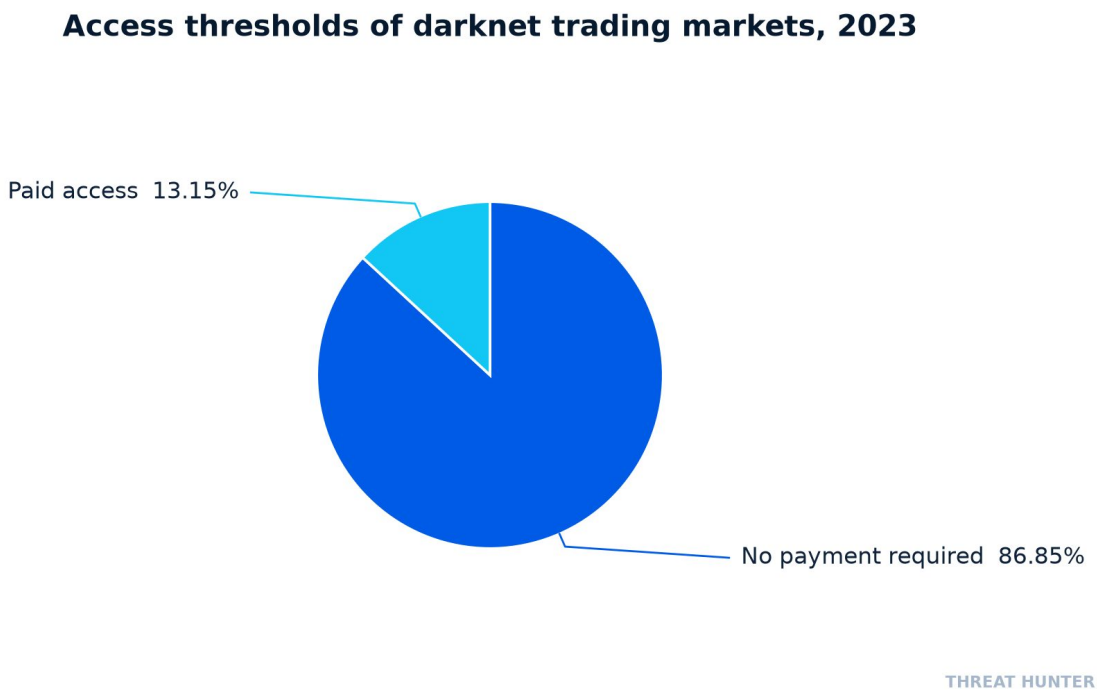
Research on threat-actor data trading market in 2023

2. Research on the threat-actor data trading market in 2023

2.1 More than 45 darknet platforms have raised the access threshold through “premium membership” and other forms

As illegal data trading markets and hacker forums gradually become public, in order to raise access thresholds and enhance platform security while achieving higher operating profits, more and more dark web forums and trading markets have begun to upgrade their membership systems, so that more valuable resources and interactions require paid membership to obtain and participate, and ordinary users can only enjoy limited resources and restricted interactions.

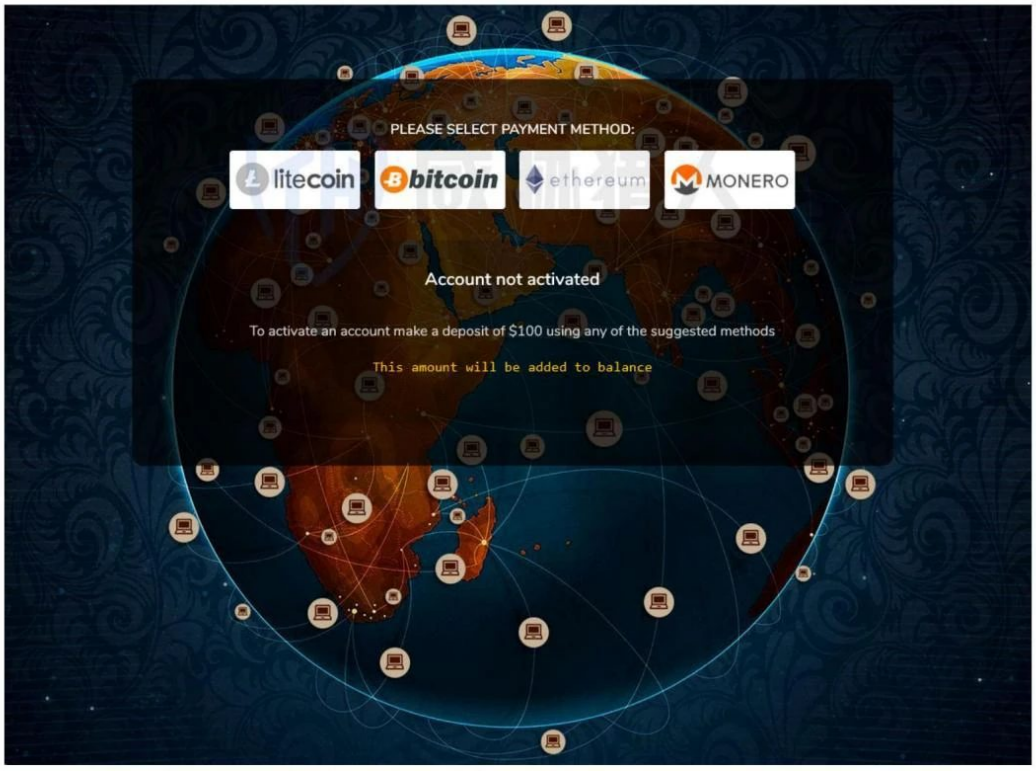
Threat Hunter conducted research on darknet trading markets and hacker forums and found that there are more than 45 darknet markets and forums that require paid registration, points unlocking, top-up membership, invitation codes, like replies, VIP upgrades, etc. to further view the specific data trading modules and related post content. The prices for top-up membership and paid viewing of information range from hundreds to thousands of dollars.



Access thresholds of darknet trading markets in 2023: 13.15% require payment

Source: Threat Hunter original report, page 20. Redrawn in English from the original figure.

Access to the website of the famous Russian data trading forum “russianmarket” requires a deposit of US$100 to register.



Registration page of a darknet data-trading forum that requires a deposit

Source: Threat Hunter original report, page 21. Original image retained.

2.2 More than 1,500 risk events were discovered in “private-domain” anonymous group chats

As mentioned above, 82.26% of the leaks of citizens’ personal information monitored by Threat Hunter in 2023 were concentrated in anonymous group chats.

10.01% occurs on the dark web. As the main channel for citizens’ personal information to be leaked, we will investigate anonymous group chats and the dark web.

Further analysis.

In order to avoid relevant legal and policy crackdowns, most threat actor groups have begun to turn to private domain groups for transactions, and data trading groups have also anonymously hidden their account information. The channel/group chat with the most risk events detected on encrypted messaging apps is the private group. Threat Hunter has discovered more than 1,500 risk events in private groups.

In addition, in order to strengthen the credibility of the channel and ensure the smooth progress of data transactions, some threat actor groups will form "public groups", which are equivalent to guarantee agencies, such as Huiwang Guarantee, Shenma Guarantee, etc. They will organize resource groups, supply and demand groups, guarantee groups, etc., undertake project needs and various resource docking, and provide guarantees during the data transaction process. The public group will also be set up as a separate private domain group, and traders can only enter through the invitation link/administrator approval.

Data trading channel/group chat top 5 data trading threat actors gang top 5

Channel / group name	Risk events
Public group ***	1,508
"Boss Crab"	234
"Fire Phoenix"	212
Big boss****	193
"Aipin"****	167

Top 5 data-trading channels and group chats by risk events, 2023 (names masked)

Source: Threat Hunter original report, page 22. Table re-typeset in English from the original.

Publisher	Risk events
Anonymous**	11,868
Bao***	208
Boss**	193
"Lonely"**	162
"Shake"**	141

Top 5 data-trading publishers by risk events, 2023 (names masked)

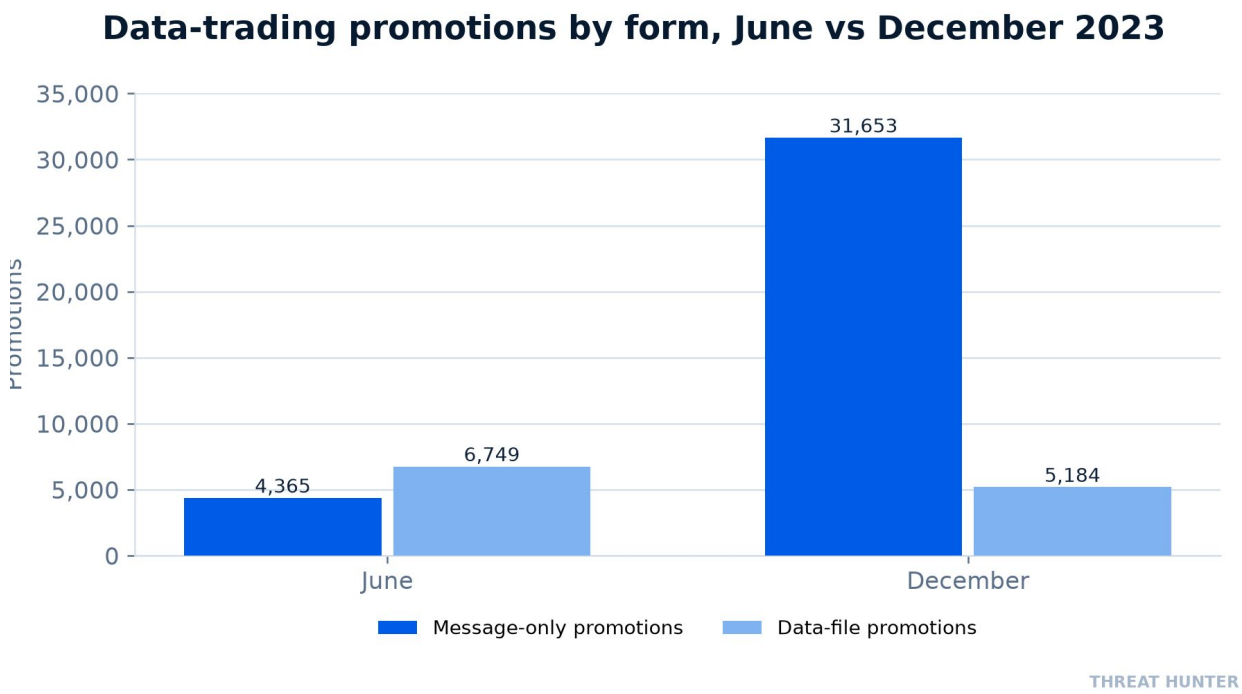
Source: Threat Hunter original report, page 22. Table re-typeset in English from the original.

2.3 The form of transaction data of threat actors is gradually transformed from "data file type" to more covert "pure message type" propaganda

Under market supervision and policy crackdowns, threat actors have become more cautious in their promotion of data transactions to better avoid fraud controls and protect their own interests.

Taking a certain financial industry as an example, Threat Hunter analyzed the changing trends in the data trading promotion forms of threat actors in the financial industry in 2023. From January to August 2023, threat actors' transactions were mainly based on "data file" promotion. threat actors traders attracted potential buyers to trade by publishing samples and descriptions of various data files.

With the intensification of supervision and the rise of market risks, in September 2023, the transaction promotion form of threat actors gradually changed to "pure news" promotion of data transactions. threat actors traders communicate with buyers through encrypted messages, secret codes and private chats, making them more covert and difficult to detect by regulatory agencies.



Data-trading promotions for one financial sector: message-only posts rose from 4,365 in June to 31,653 in December

Source: Threat Hunter original report, page 23. Redrawn in English from the original figure.

2.4 Among different threat actors transaction data formats, the SMS hijacking format has the highest data accuracy.

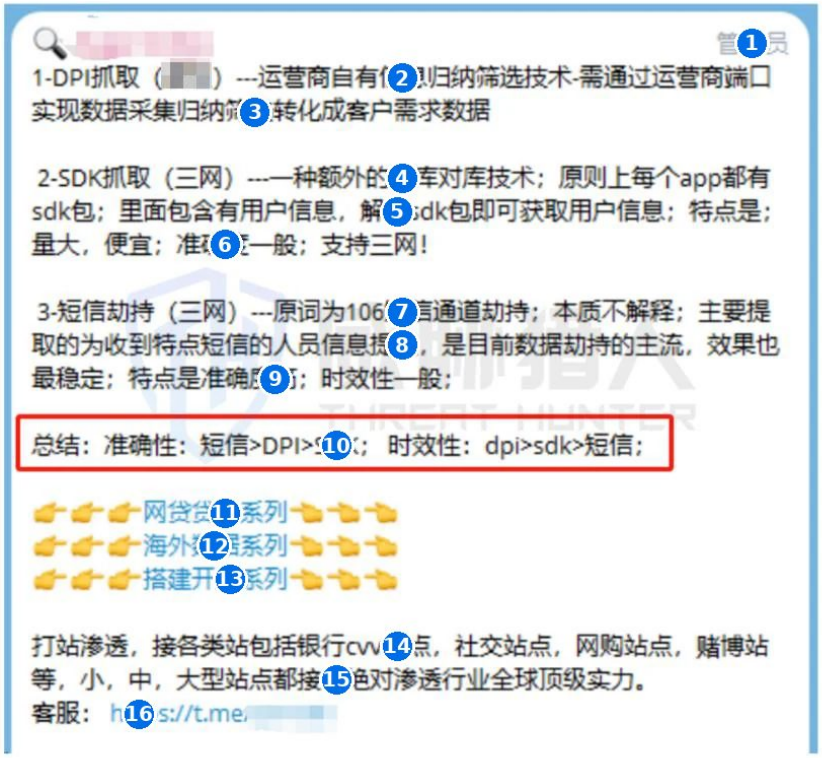
There are two main dimensions for threat actors to evaluate data. One is the timeliness of the data, and the other is the accuracy of the data.

Threat Hunter conducted in-depth research on data trading threat actors and found that the accuracy and timeliness of data in different formats are different.

It is mainly information data in SMS hijacking format, DPI format, and SDK format. It is understood that the amount of data that can be stably generated every day is as high as 10,000 pieces.

In terms of accuracy: data in SMS hijacking format > data in DPI format > data in SDK format;

In terms of timeliness: data in DPI format > data in SDK format > data in SMS hijacking format.



A seller explaining three data formats: DPI capture, SDK capture and SMS hijacking

Source: Threat Hunter original report, page 24. Original image retained; numbered translations added by Threat Hunter.

Figure guide

1	Admin
2	1. DPI capture: carriers' own information, sorted and filtered through carrier ports
3	to collect, sort and convert data into what clients need
4	2. SDK capture (all three carriers): a supplementary matching technique; in principle every app has
5	an SDK package containing user information, which can be parsed to obtain it. Features:
6	large volume, cheap, average accuracy, all three carriers
7	3. SMS hijacking (all three carriers): hijacking of bulk-SMS channels; extracts
8	details of people who receive specific SMS messages; currently the mainstream form of data hijacking and
9	the most stable: high accuracy, average timeliness
10	Summary: accuracy SMS > DPI > SDK; timeliness DPI > SDK > SMS
11	Online-lending series
12	International data series
13	Site building and development series
14	Site penetration of all kinds: banking, social, shopping and gambling sites
15	small, medium and large; claims to be a top global player
16	Customer service: messaging-app link

Threat Hunter verified the leaked data sample and found that the accuracy of data in different formats is indeed different. The results are consistent with the descriptions of data trading threat

actors. In terms of accuracy: data in SMS hijacking format > data in DPI format > data in SDK format. The specific reasons are as follows:

1. Data in SMS hijacking format: Contains the content of the SMS. The authenticity of the data is basically confirmed as the SMS notification has been truly received.

personal information of citizens;

2. Data in DPI format: mainly data for traffic analysis, mainly traffic data obtained through the operator side, so

The authenticity is low. According to Threat Hunter survey statistics, the authenticity of this type of data basically does not exceed 20%;

3. Data in SDK format: mainly by parsing the software package name information, and then providing it to the operator side to obtain download traffic, and storing it.

The software in the application market is traffic-manipulated with user information or user information that has not been registered, so the authenticity is the lowest.

2.5 Changes in the data trading situation of threat actors in 2023 and new downstream malicious methods

2.5.1 “Cancellation, change, or booking” fraud incidents occur frequently in the aviation industry and have a high success rate

Threat Hunter's investigation of malicious activity gangs in the downstream of the data trading market found that in 2023, a large number of threat actor gangs appeared in the aviation industry to carry out online fraud by changing air tickets, and the fraud success rate was high.

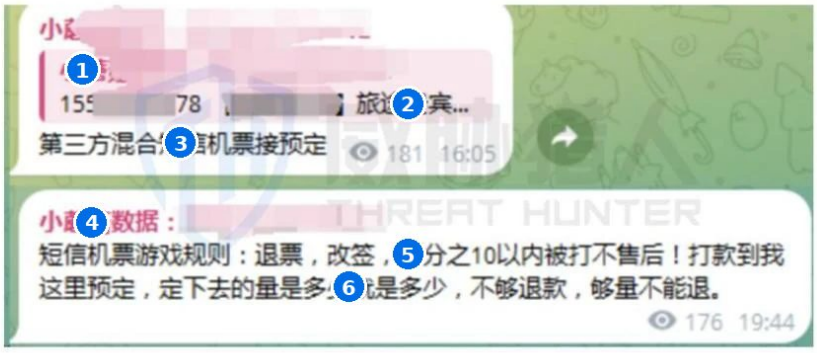
In 2023, as the COVID-19 epidemic improves, the tourism industry will pick up significantly, consumer demand will be strongly released, and passenger information such as air ticket information and hotel information will also increase significantly. Citizens' personal information on air travel is favored by downstream threat actors fraud gangs.

Case: The fraud gang contacted flight passengers by pretending to be airline staff, saying that the flight was delayed and that tickets could be changed or refunded. At the same time, in order to obtain "high compensation", they induced users to download video conferencing, remote sharing and other operations, and gradually fell into the trap of the fraud gang. After obtaining the user's bank card, mobile-payment account, verification code and other sensitive information, they transferred and stole the passenger's funds.

Step	What happens
1	User information data is leaked
2	Fraudsters pose as airline staff to win the user's trust
3	They report a flight change and lure the user into an unofficial refund or rebooking at a high price
4	They operate the refund remotely and obtain the user's bank card, verification codes and other details
5	They illegally take the money in the user's bank card

How "cancellation and rebooking" fraud uses leaked passenger data

Source: Threat Hunter original report, page 26. Table re-typeset in English from the original.



A seller of flight-booking SMS data sets out the "rules" for buyers

Source: Threat Hunter original report, page 26. Original image retained; numbered translations added by Threat Hunter.

Figure guide

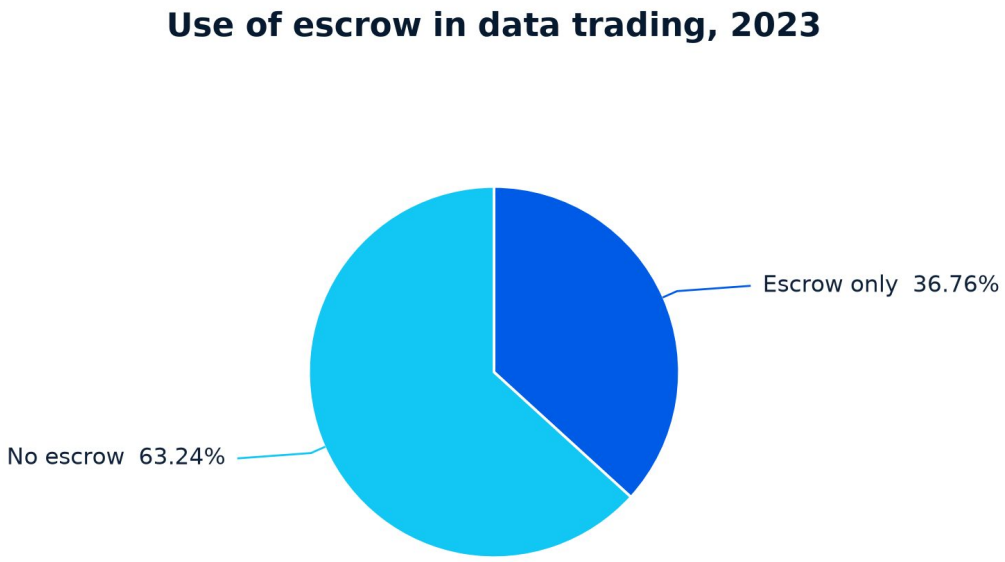
1	"Little Mushroom"
2	Travel VIP …
3	Third-party mixed SMS flight bookings taken
4	Little Mushroom Data:
5	SMS flight-ticket rules: refunds and changes; no after-sales if under 10% unreachable. Pay me to book
6	the quantity ordered is final; refund if short, none once the quantity is met

2.5.2 Third-party guarantees and blockchain payments add security to threat actors’ data transactions

Third party guarantee:

In the past, the two parties to the transaction usually adopted a one-to-one approach, that is, one party paid and the other party delivered the data. Due to the continuous outbreak of transaction fraud and transaction scams in recent years, the data trading market has become more chaotic, and the one-to-one transaction method is no longer feasible.

In order to avoid being cheated, data trading threat actors gradually changed their trading methods and began to rely on intermediary guarantee platforms and third-party guarantee platforms for transactions, which provided more reliable guarantees for transactions and made transactions more secure and credible.



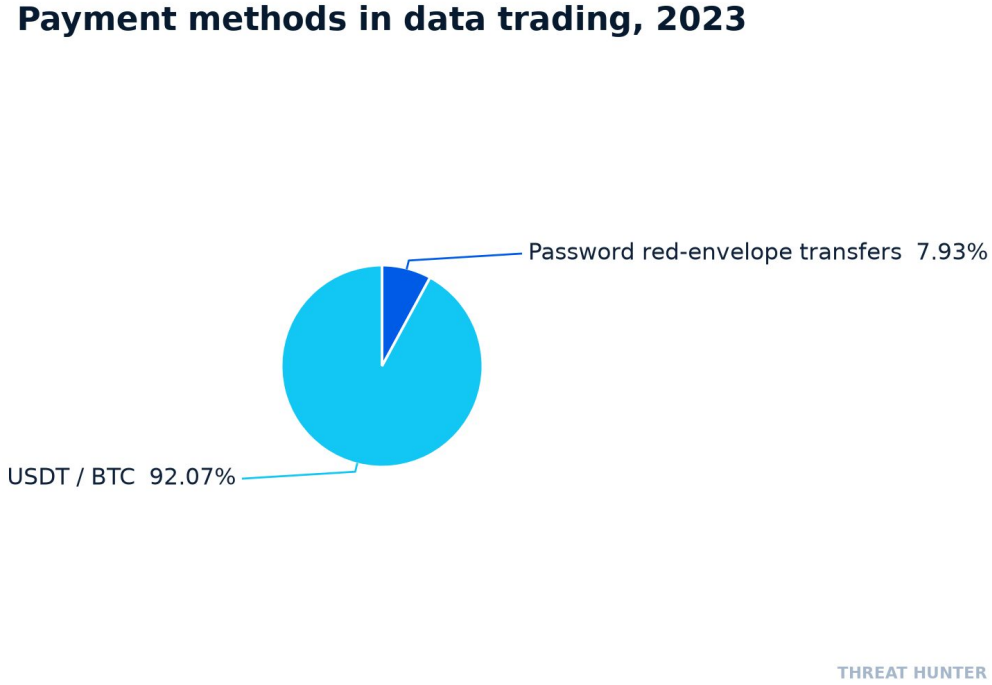
Use of escrow in data trading, 2023: 63.24% of deals skip escrow

Source: Threat Hunter original report, page 27. Redrawn in English from the original figure.

Blockchain payment:

In the past, threat actors often made transaction payments through password red envelopes, payment platform transfers, etc. Because these payment methods are easy to track and control, and the judicial authorities have severely cracked down on data trading threat actors in recent years, the data trading threat actors have had to change their transaction payment methods and gradually switch to using only cryptocurrency for payment.

Cryptocurrency has the characteristics of anonymity and decentralization, which allows the identities of both parties to the transaction to be kept confidential and is difficult to be tracked and controlled by regulatory agencies. Therefore, using cryptocurrency to pay has become a new trend in data transactions among threat actors. This change further increases the difficulty of cracking down on data trading by threat actors, and more powerful measures need to be taken to protect the security and stability of the data trading market.



Payment methods in data trading, 2023: USDT or BTC account for 92.07%

Source: Threat Hunter original report, page 28. Redrawn in English from the original figure.

2.5.3 The transaction price of “full format” information data in the financial industry is the highest, with a unit price of 25 yuan

Threat Hunter research and statistics found that, driven by interests, the types of data stolen by upstream threat actors have changed.

In the financial industry, the value of citizens' personal information data is closely related to the number of fields it contains. "Full-format" information usually includes ID card, name, phone number, bank card, loan information, address and other fields. Downstream threat actors can use complete citizen portrait information to achieve precise malicious activity. Therefore, "full-format" data has the highest price, with a single piece of data priced up to 25 yuan. In contrast, citizens' personal information data that only contains the phone number field has lower returns due to the restrictions on the methods of downstream malicious activity gangs, and the data price is generally around 30 cents.

In addition, there are differences in data prices in different industries. For example, the personal information of citizens in the financial industry is targeted at high-net-worth individuals who hold more funds, so the data price is more expensive; student information data is relatively cheap because the downstream income from malicious activity is low. This differential pricing reflects the needs and value perceptions of different industries for citizens’ personal information.

2.1

Improve data leakage risk monitoring and early warning capabilities from multiple dimensions

3. Improve data leakage risk monitoring and early warning capabilities from multiple dimensions

Judging from the current status of data breach risks in 2023, enterprises need to focus on the following issues:

1. There is a frequent phenomenon of threat actors re-splicing "historical personal data information" and selling it multiple times.

There are a large number of threat actors selling false data, spliced data, historical data, transaction fraud and other behaviors in the data trading market. It is difficult for enterprises to verify the authenticity of data.

2. More than 50% of citizens' personal information is traded at night, and more than 30% is traded on non-working days.

In the data transaction time distribution of citizen personal information leakage incidents in 2023, the number of incidents that occurred on non-working days (weekends and holidays) was as high as 31.21%, and the number of incidents that occurred at night accounted for as high as 51.88%, more than half.

3. The number of data breaches in the financial industry exceeds 8,700, ranking first in the distribution of data breaches in the industry

The financial industry is still the hardest hit area by personal information leaks, with more than 8,700 data leakage incidents involving information data of high-net-worth individuals in banking, insurance, securities and other industries. This is mainly due to the higher value of profits used by downstream threat actors for marketing promotion and fraud.

In response to the above situation, enterprises need to comprehensively improve their ability to identify and respond to risks, understand the details of risk events, including verifying the authenticity of risks, promptly tracing their sources, handling delisting and following up on potential risks, and enhancing the timeliness of data leakage risk monitoring and early warning, etc.

In this regard, Threat Hunter proposed a targeted solution:

1. Strengthen the verification of risk authenticity: Based on the whole network intelligence monitoring and in-depth mining, based on the number of monitored transactions of threat actors

According to the data, through the risk authenticity verification engine + manual data verification service, comprehensive credibility assessment results are provided to help enterprises accurately perceive risks and handle risks in a timely manner. Among them, the risk authenticity verification engine is based on the "source confidence factor, three factors

The three elements of "matching factor and historical coincidence factor" verify the authenticity of risks, help enterprises accurately perceive risks, and provide more reliable data security protection for enterprises and individuals.

2. "7×24×365" emergency response: In October 2023, Threat Hunter established the Digital Risk Emergency Response Center (DRRC).

Carry out round-the-clock monitoring, review and early warning of enterprise-related risk information, and provide services such as "7×24×365" sample acquisition, intelligence mining, assistance in traceability, disposal and removal.

Data leakage risk monitoring and early warning capabilities need to be continuously improved from multiple dimensions. Only by continuously deepening the comprehensive understanding of the enterprise itself and the industry's risk landscape can we lay a solid foundation for security for the efficient and sustainable development of enterprise digitalization.