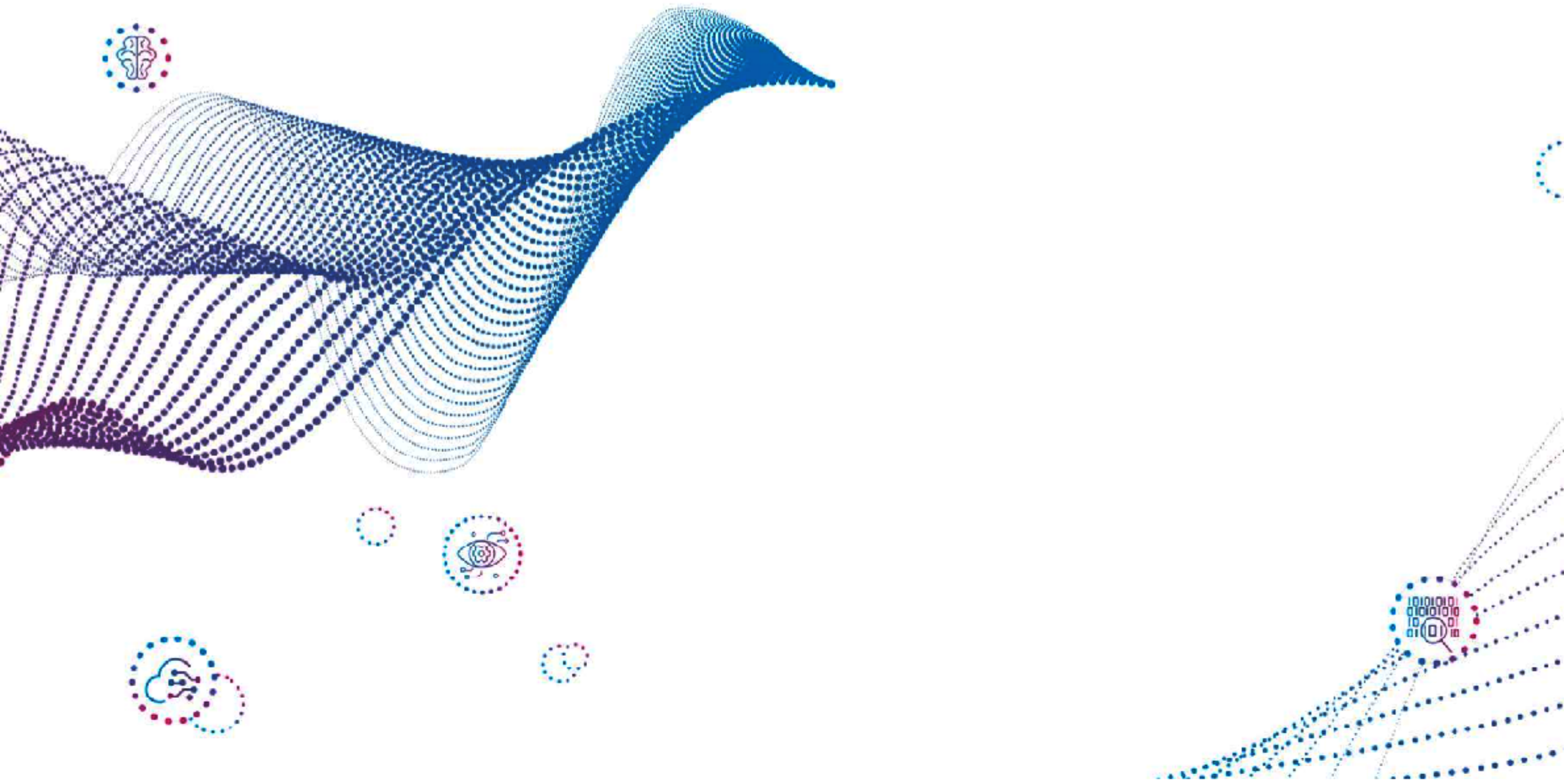


Email analysis intelligence decision-making platform

Product white paper

(V1.0 version in 2022)

Table of contents

1 Introduction

Email is a communication method that uses electronic means to provide information exchange. It is the most widely used service on the Internet. Through the Internet email system, users can send emails at a very low price (no matter where they are sent, they only need to pay the network fee) and in a very fast way (it can be sent to any designated destination in the world within a few seconds). Connect with Internet users anywhere in the world.

With the rapid development of Internet technology, people send emails through the Internet, making communication easier and faster. Email has also become one of the indispensable and important communication methods in modern society with its new, fast and economical characteristics. At the same time, various criminals have begun to widely use emails to engage in various illegal and criminal activities. Emails are involved in many computer crime cases as well as commercial and civil disputes.

E-mails contain a wealth of useful information, which is one of the important contents for computer analysis and evidence collection. It can provide powerful clues for case detection. In order to improve efficiency, people often use various email clients (such as Foxmail, Outlook Express, Microsoft Office Outlook, etc.) to process emails. Therefore, analyzing email data files saved by various email clients is also one of the important means of computer analysis and evidence collection. Extract the sender and receiver email addresses, sender and receiver names, sending time, subject, email body content, attachments and other information of the emails of interest from the massive email data obtained, and statistically classify the recipient and sender email addresses, using visualization , interpersonal network analysis and other technologies to draw time relationship diagrams and interpersonal network relationship diagrams of emails, thereby providing a good reference for analyzing and discovering hidden relationships between recipients and senders.

2 Requirements analysis

The construction of an "email analysis intelligence decision-making system" based on massive email data should be able to fully meet the all-round intelligent analysis of email data, mining email person personal information, email person relationship network, email attribute IP, email attachment extraction, etc. Through the construction of this system, the import, analysis and judgment of massive email data can be realized.

(1) Email data retrieval function

In order to ensure that key information can be quickly queried in massive email data, and key evidence can be locked to improve the efficiency of research and judgment, the construction of an "email analysis intelligence decision-making system" should be able to achieve rapid analysis and second-level retrieval of terabyte-level email data. Users can use key words to You can quickly search and query related emails.

(2) Email target analysis function

In order to ensure the analysis of a single individual's email in a massive amount of emails, the relevant information of the individual is highlighted. The construction of the analytical intelligence decision-making system should be able to meet the comprehensive analysis function of a single target email, and can realize the analysis and display of comprehensive data such as semantic habits, relationship networks, and personal information of the target email.

(3) Email analysis collaboration function

In order to establish email correspondence between groups or organizations, record the activities of the target group or organization in an organization chart. The construction of the "email analysis and intelligence decision-making system" should be able to realize collaborative analysis by multiple people. The system can realize various types of marking and remarks on different emails. and group management to facilitate internal collaborative analysis.

(4) Email data management function

In order to manage massive email data more effectively. The construction of the "Email Analysis Intelligence System" should be able to realize the comprehensive management function of email data, grasp the total amount of system email data, email language types, etc. in real time to ensure that the system can operate reliably and stably.

3 Product Introduction

3.1 Product introduction

The "Email Analysis Intelligence Decision System" is developed and designed based on text recognition big data technology. It supports the rapid identification and analysis of massive email data and extracts intelligence information such as keywords, sensitive words, personal relationships, and contact information.

3.2 Product composition

The "Email Analysis Intelligence Decision System" supports two forms: private cloud and public cloud. Private cloud means that users build a complete set of localized independent platforms themselves; public cloud adopts SaaS model, and users can log in directly through authorized accounts.

The product composition list of "Email Analysis Intelligence Decision System" (Private Cloud) is mainly as follows:

1. Platform server deployment software: 1 set

2. Platform client login software: 1 set

3. Platform authorized login account: 1

4. Platform user manual: 1 copy

The product composition list of "Email Analysis Intelligence Decision System" (Public Cloud) is mainly as follows:

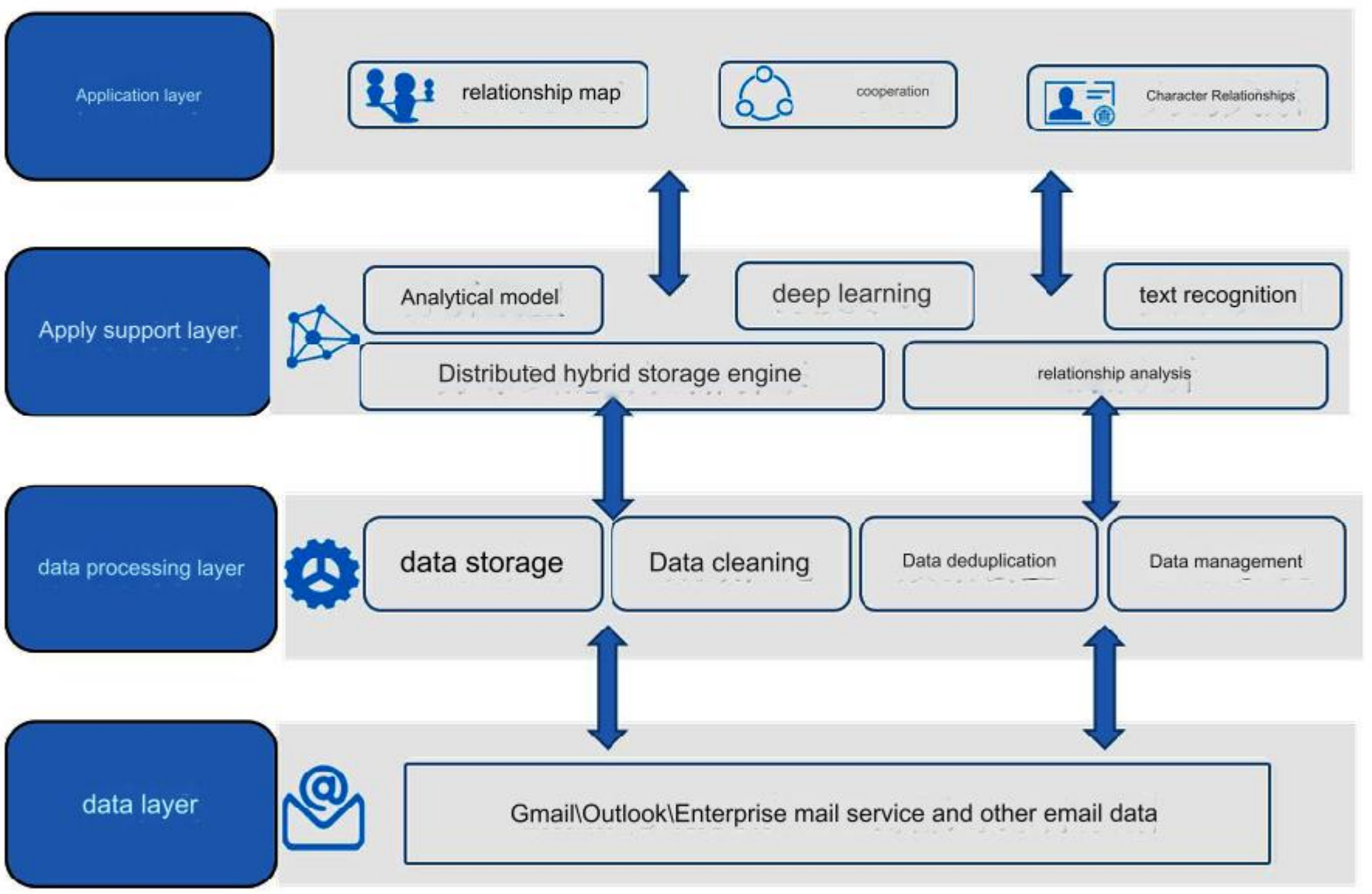
1. Platform client login software: 1 set

2. Platform authorized login account: 1

3. Platform authorized login dongle: 1

4. Platform user manual: 1 copy

3.3 System architecture



(Architecture diagram of "Email Analysis Intelligence Decision System")

The "Email Analysis Intelligence Decision System" is designed and developed around the data layer, data processing layer, application support layer and application layer. The system supports the import of various data sources. Users can import all email data in ".eml" format into the system for processing. Analysis and calculation; the data processing layer performs analysis, processing and intelligent extraction of imported various email data to realize the functions of cleaning, deduplication, storage and management of email data; the application support layer performs comprehensive and in-depth analysis of the cleaned data and processing, combined with built-in analytical models and deep

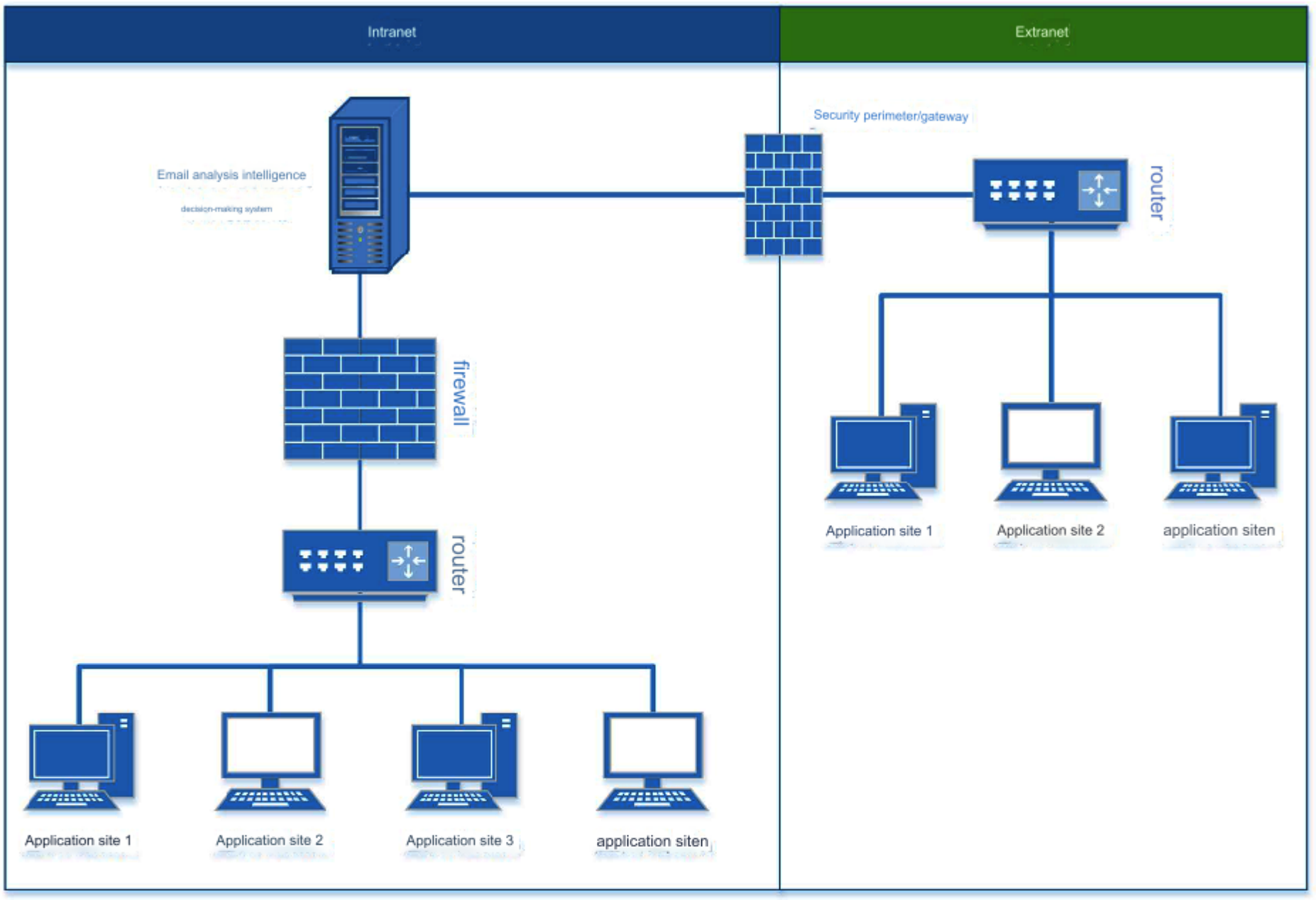
Learning algorithm, realizes the extraction and classification of email text and various types of data in emails, mines various relationship networks in emails,

and distributes the mined structured data and unstructured data to realize rapid call of data. and query; the application layer is to UI the system platform.

Users can log in to the system through the Web interface to achieve comprehensive research and analysis of email data, and can perform a series of functions

such as search, analysis, and research and judgment.

3.4 Network architecture



(Network architecture diagram of "Email Analysis Intelligence Decision System")

The "Email Analysis Intelligence Decision System" is designed using a B/S architecture to facilitate users to log in and use it at any time. In order to ensure the security of the

obtained email data, the "Email Analysis Intelligence Decision System" is deployed in principle on the user's intranet and is isolated from the external network as much as possible. If there

is a need for external network access and use, it can be accessed through a security gateway or gatekeeper. Map the intranet "email analysis intelligence decision-making system" to the

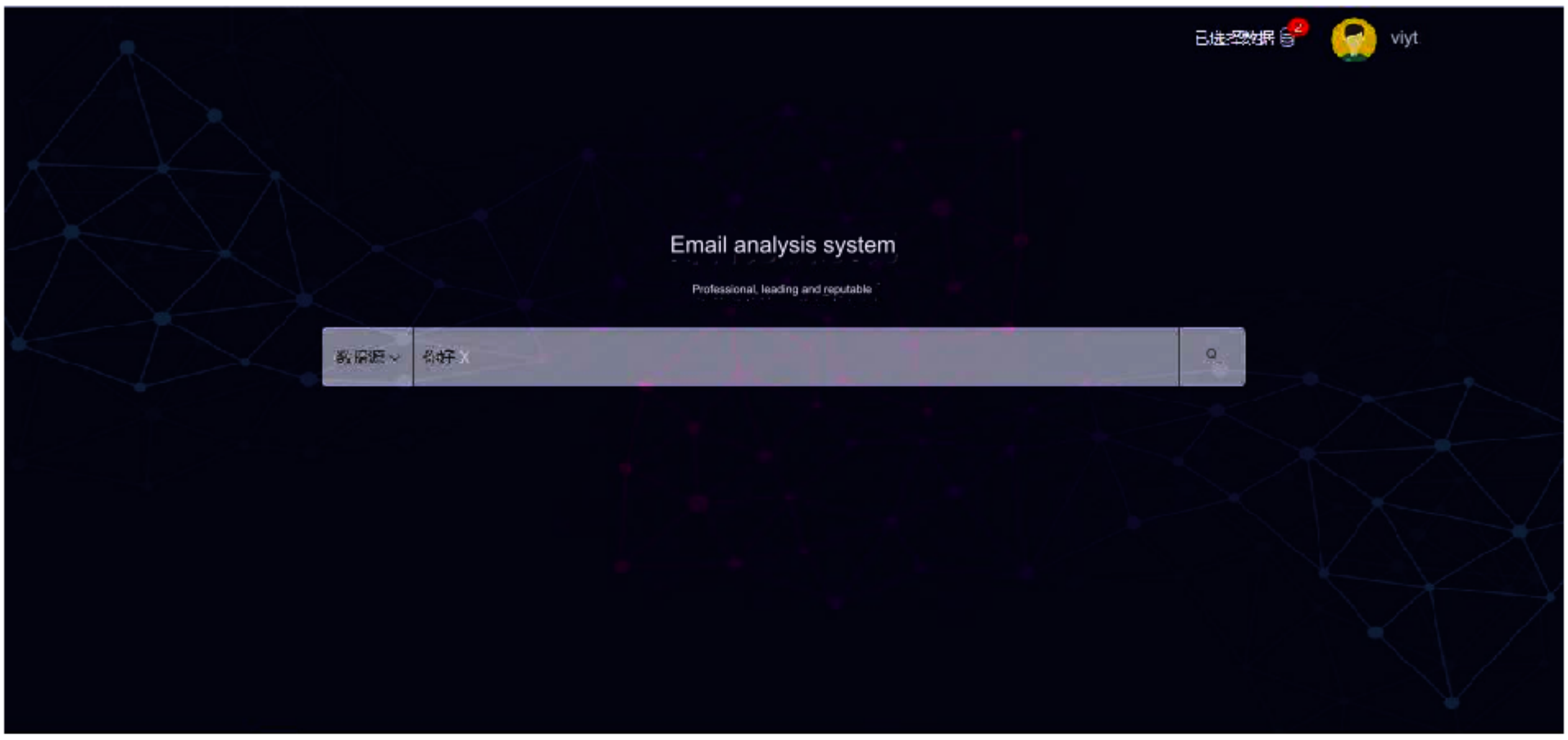
external network to meet external network access and usage requirements.

4 product features

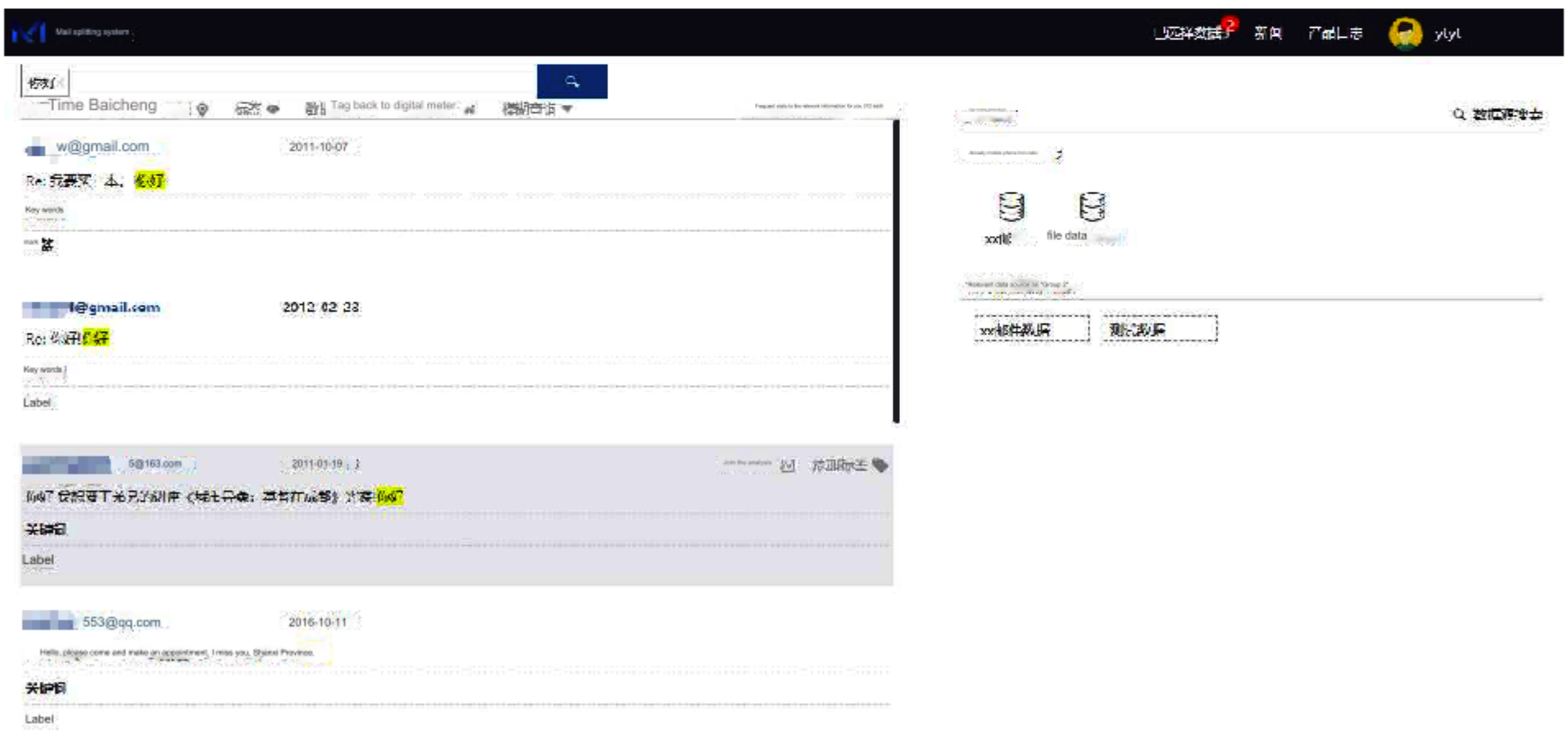
The "Email Analysis Intelligence Decision System" provides four functional modules: global search, email browsing, user backend and workbench.

4.1 Global search

After entering the "Email Analysis Intelligence Decision System", you can quickly retrieve and query global email data by entering key words in the search box according to the corresponding data source group.



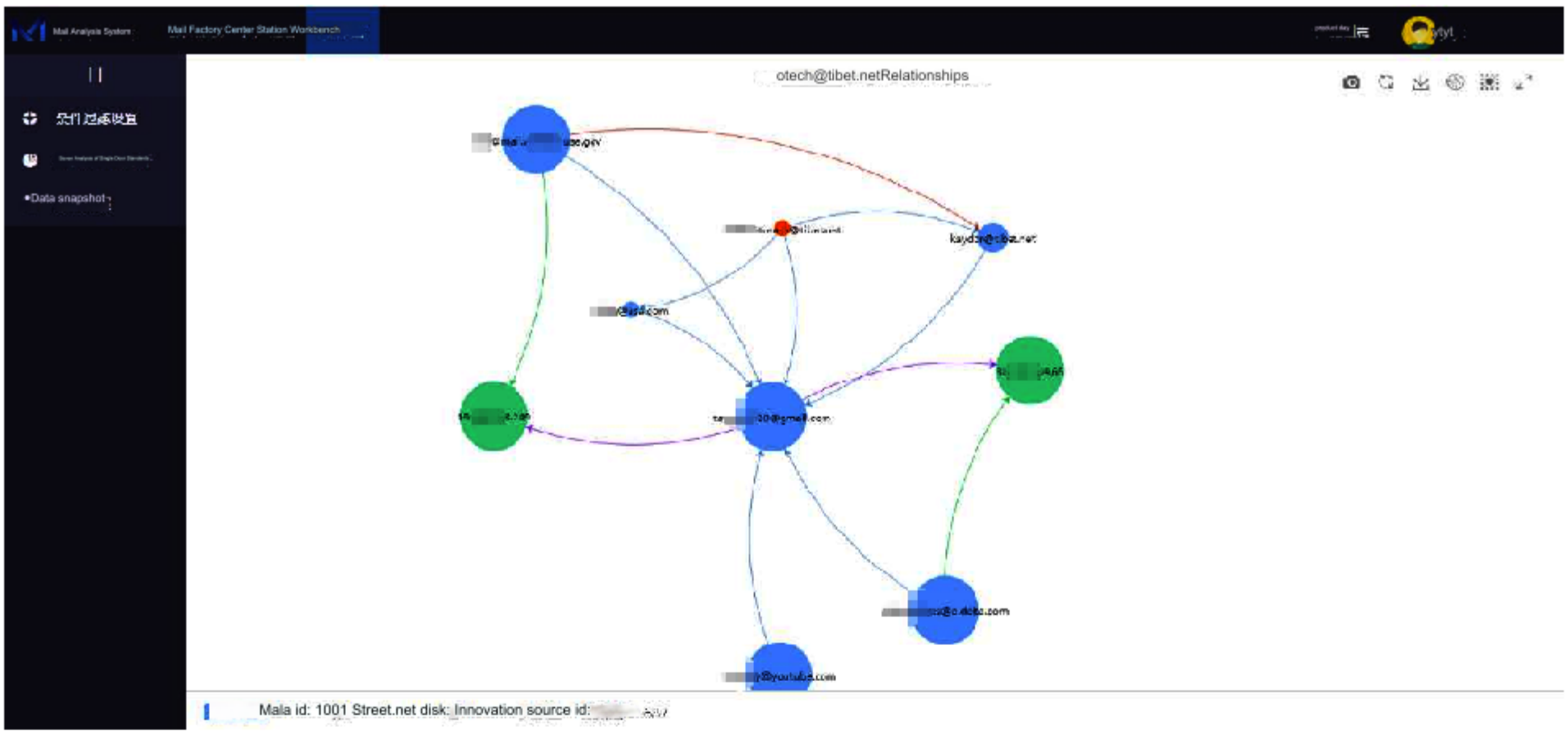
(Search interface)



(Search Results)

4.2 Workbench

The retrieved target email data can be added to the workbench for further analysis and judgment. Provides three major functions: conditional filtering settings, single target comparison analysis, and data snapshots.



(Workbench main interface)

4.2.1 Conditional filtering settings

Conditional filtering settings mainly include four function points: relationship calculation, filtering settings, relationship creation, and recovery and deletion. The main functions of each function point are as follows:

1) Relationship calculation: Mainly set and optimize parameters such as target type, relationship type, target list, relationship dimension, maximum number of points, etc., to display and present the target relationship network that meets the expectations.

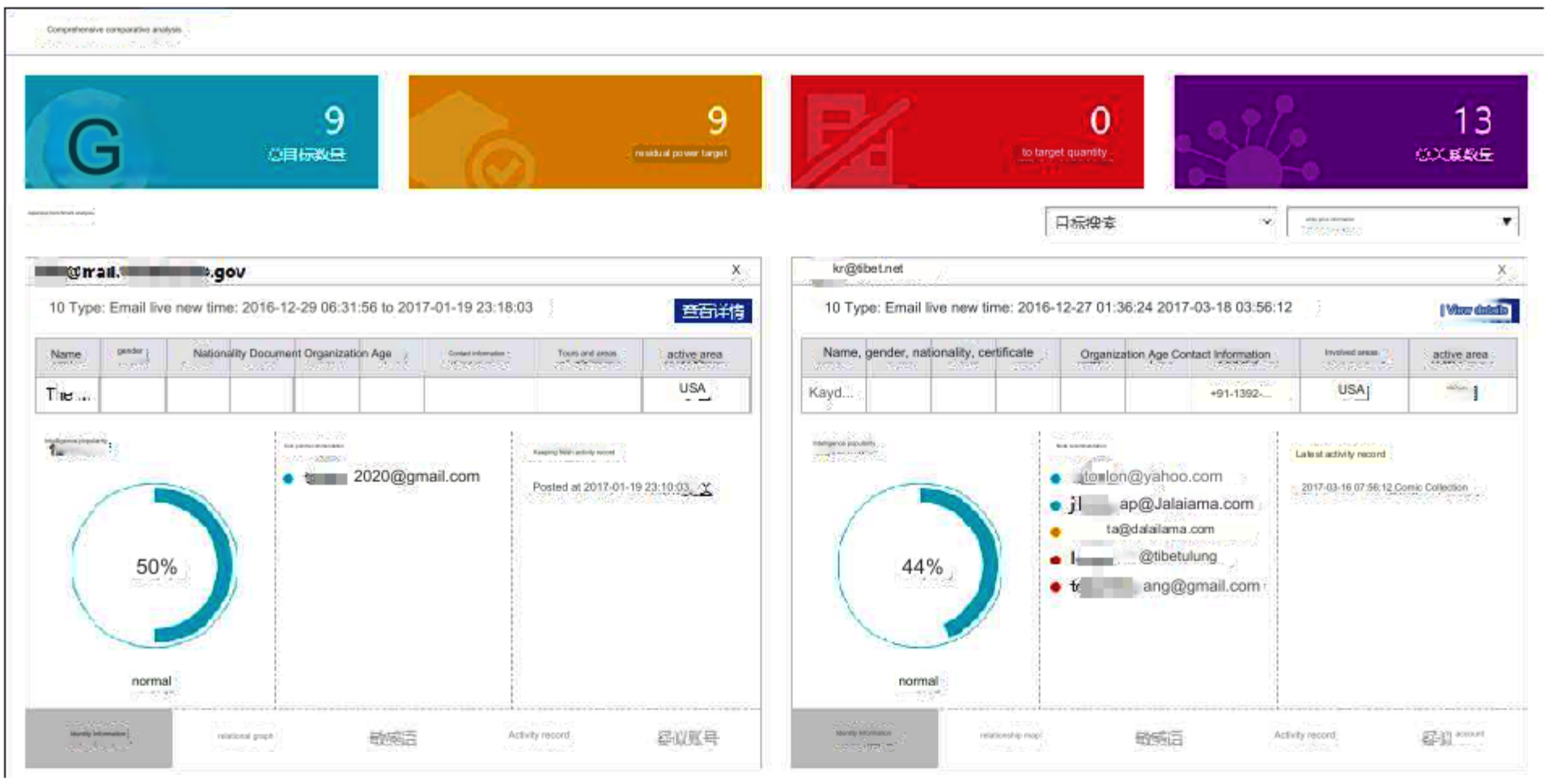
2) Filtering settings: You can filter the generated relationship network to filter out unnecessary relationship types and target types in the relationship network.

3) Create a relationship: Based on the target contact information that has been mastered, create a relationship independently to conduct correlation analysis with the massive data of the entire system.

4) Restore deletion: Unnecessary points can be deleted in the generated relationship network, and deleted relationship points can also be restored to the relationship network.

4.2.2 Single target comparison analysis

The system will compare and analyze all information such as identity information, relationship maps, sensitive words, activity records, suspected accounts, etc. after analyzing the target, and support the generation of report documents for printing and analysis.

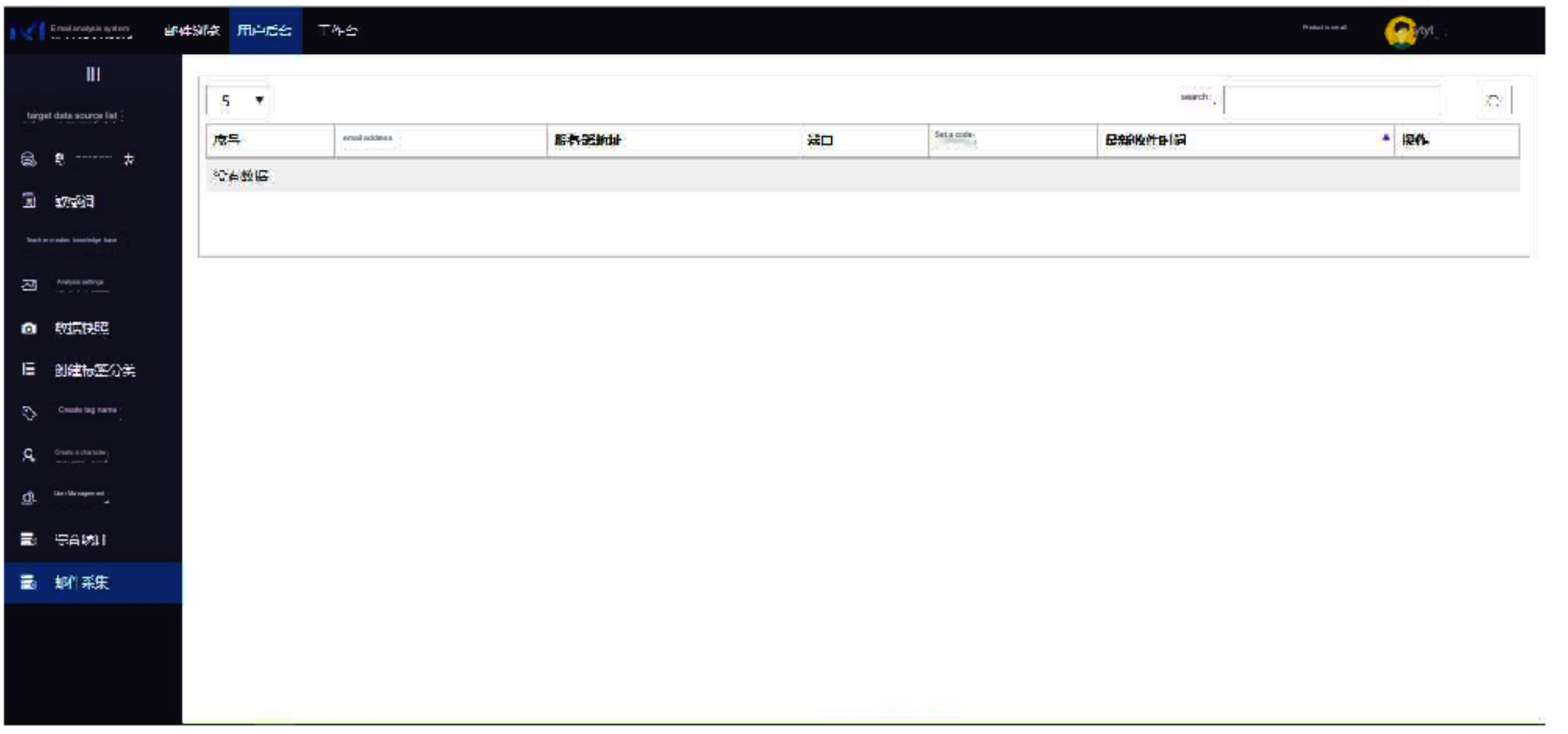


4.2.3 Data Snapshot

The data snapshot function is mainly used to take screenshots and save the relationship network when analyzing target relationships to facilitate subsequent decision-making and analysis. For the relationship network that clicked on the data snapshot, you can view and retrieve it in real time on the data snapshot toolbar.

4.3 User background

The user backend function is mainly the centralized management and maintenance of the entire "Email Analysis Intelligence Decision System" data, which mainly involves data source list, data source group list, sensitive words, creation of knowledge base, analysis settings, data snapshots, creation of label classification, It has twelve major functions: creating tag names, creating characters, user management, comprehensive statistics, and email collection.



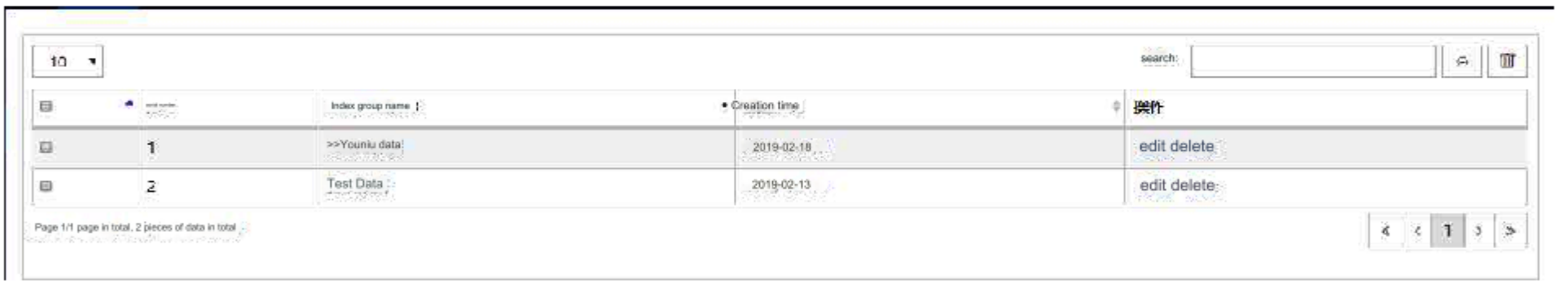
(User backend interface)

4.3.1 Data source list

The data source list is mainly used to add, manage and delete data sources. And any combination of various data sources can be imported into the system for analysis.

(data source list)

The data source group list is used to group and classify the imported data sources to facilitate unified notes, unified analysis, unified query, and unified management of the same or related data sources.



(data source group list)

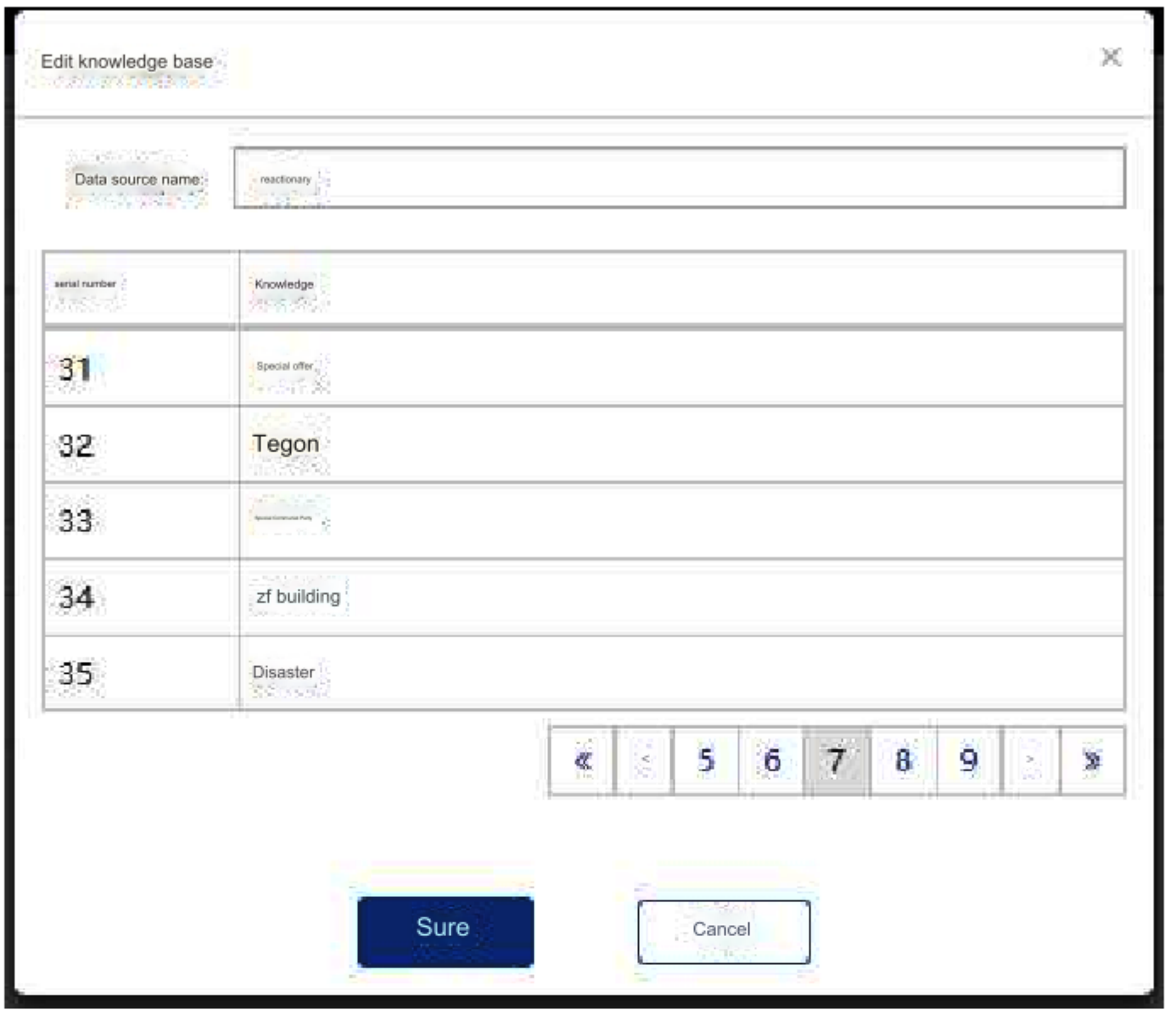
4.3.2 Sensitive word settings

The system supports setting the created knowledge base as sensitive words. The analysis system will automatically identify the sensitive words and display them when performing data analysis.

(sensitive word settings)

4.3.3 Create knowledge base

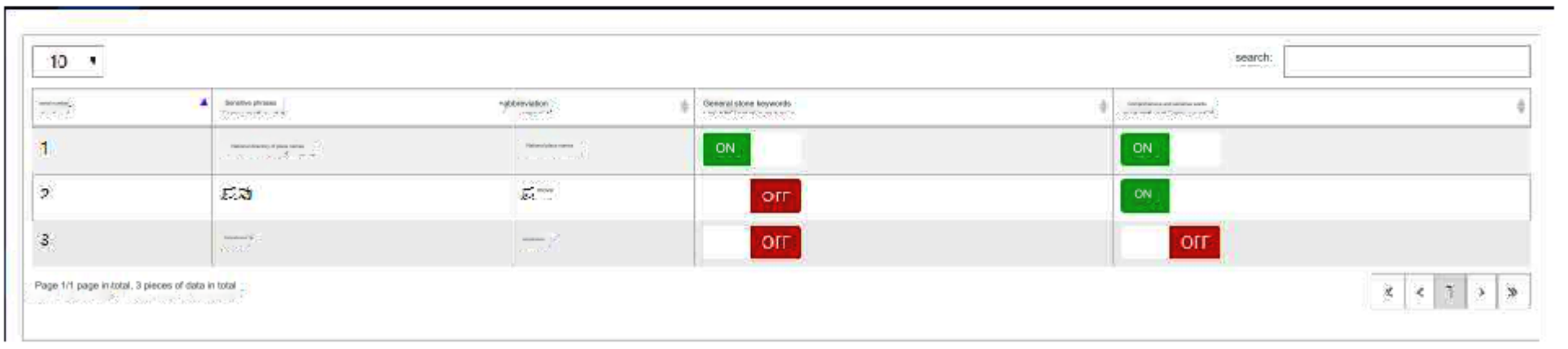
The system supports self-built knowledge bases, such as establishing knowledge bases for sensitive names, sensitive words, etc. After the knowledge base is created, the knowledge base can be set as sensitive words for call analysis.



(create knowledge base)

4.3.4 Analysis settings

Analysis settings are mainly for managing sensitive phrases. Different sensitive words can be set on and off, and search page keywords and comprehensive page sensitive words can be turned on or off as needed.



(Analysis settings)

4.3.5 Data Snapshot

The background management function of data snapshots is mainly for downloading and deleting management of executed data snapshots.

(data snapshot)

4.3.6 Create label classification

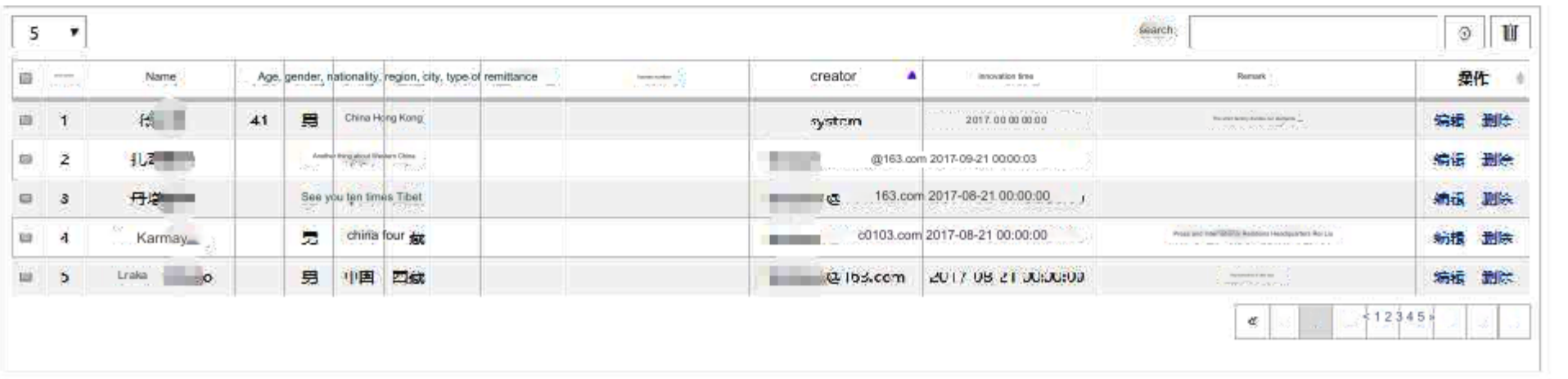
Tag classification mainly involves unified management and distribution of corresponding tag names, and classifying tags of a certain type into the same category to facilitate analysis, management and query.

(Key tag classification)

Users can add tag names as needed, and support functions such as tag name classification and adding notes.

(create tag name)

Supports character creation based on the acquired target person information, email information and other data. The created character attributes mainly include name, age, gender, nationality, region, certificate type, certificate number, creator, creation time, remarks, etc. Created characters can be edited and deleted.



(create character)

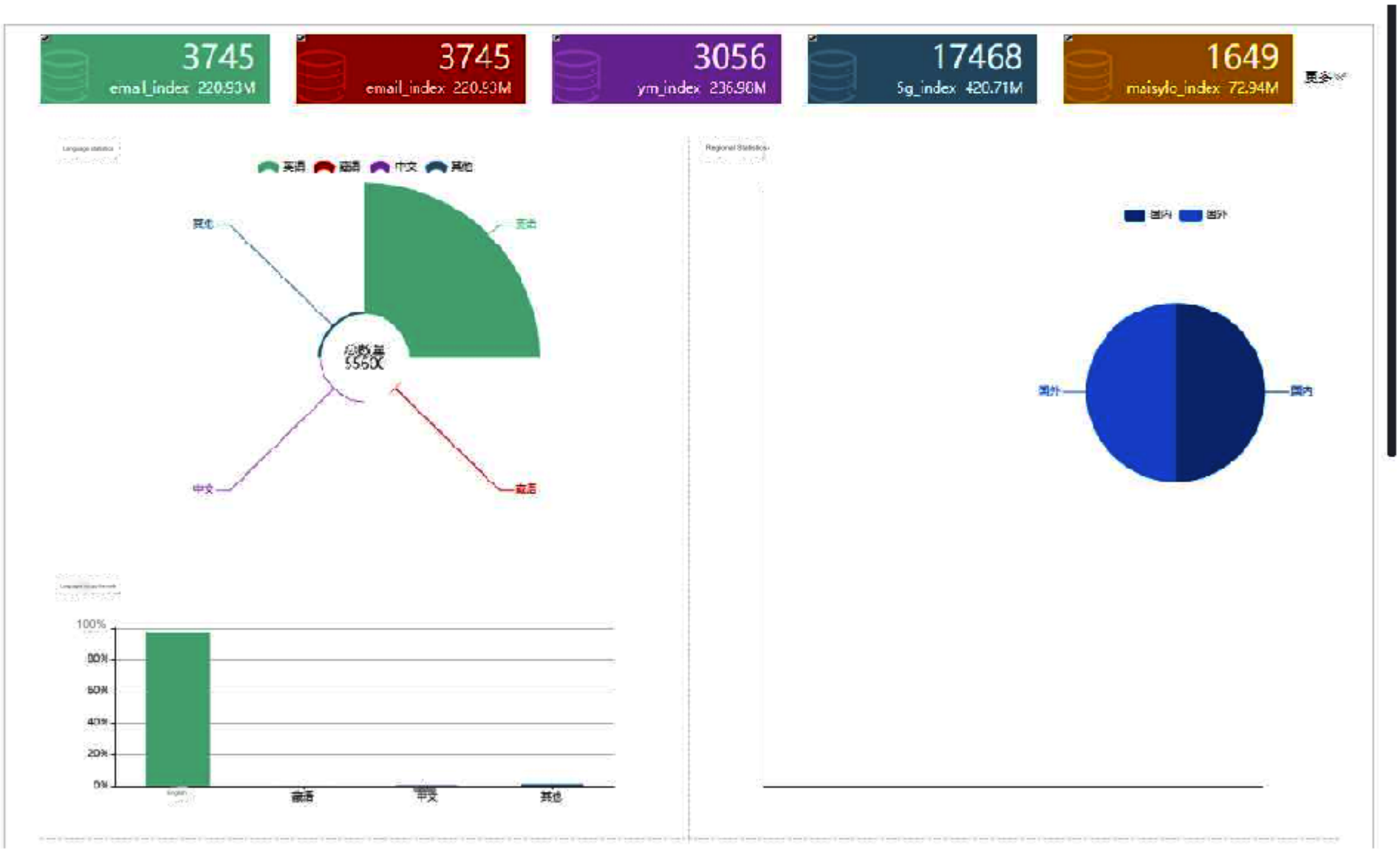
4.3.7 User management

The user management management function is mainly for the management and operation of relevant users of the system, including user creation, password modification, permission setting, binding secret keys and other functions.

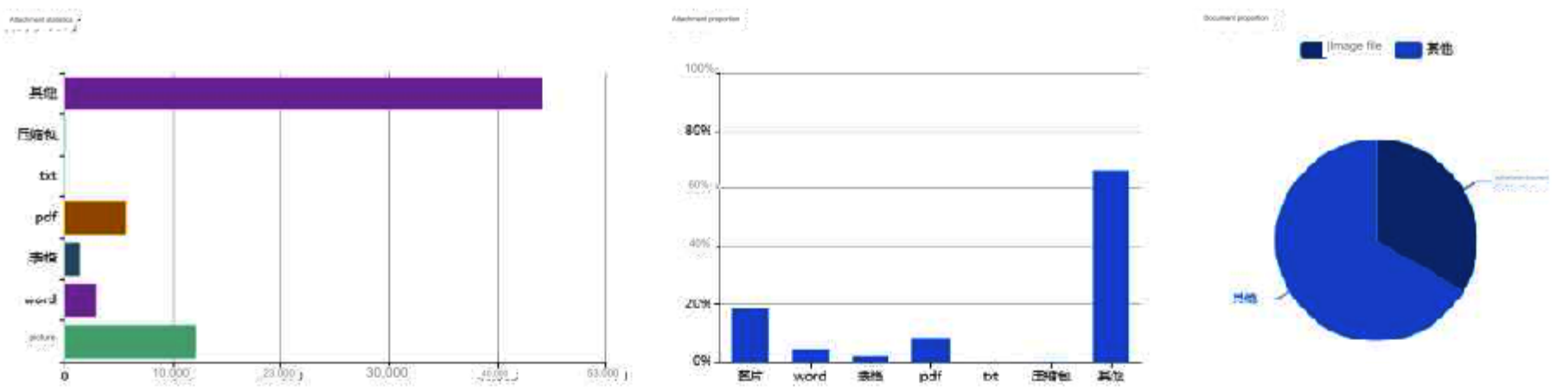
(User Management)

4.3.8 Comprehensive Statistics

Comprehensive statistics mainly implements comprehensive statistical analysis of all email data in the system, and masters comprehensive information data such as email data volume, language statistics, regional statistics, attachment statistics, etc.



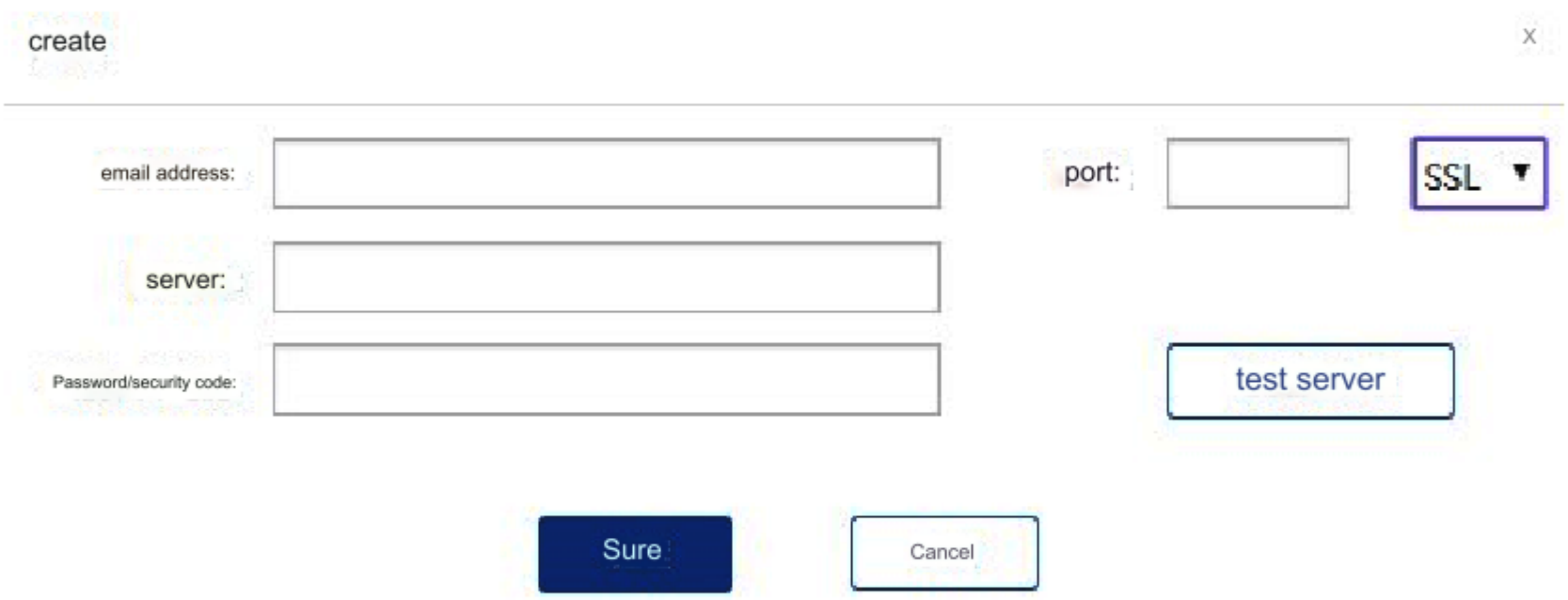
(Comprehensive Statistics 1)



(General Statistics 2)

4.3.9 Email collection

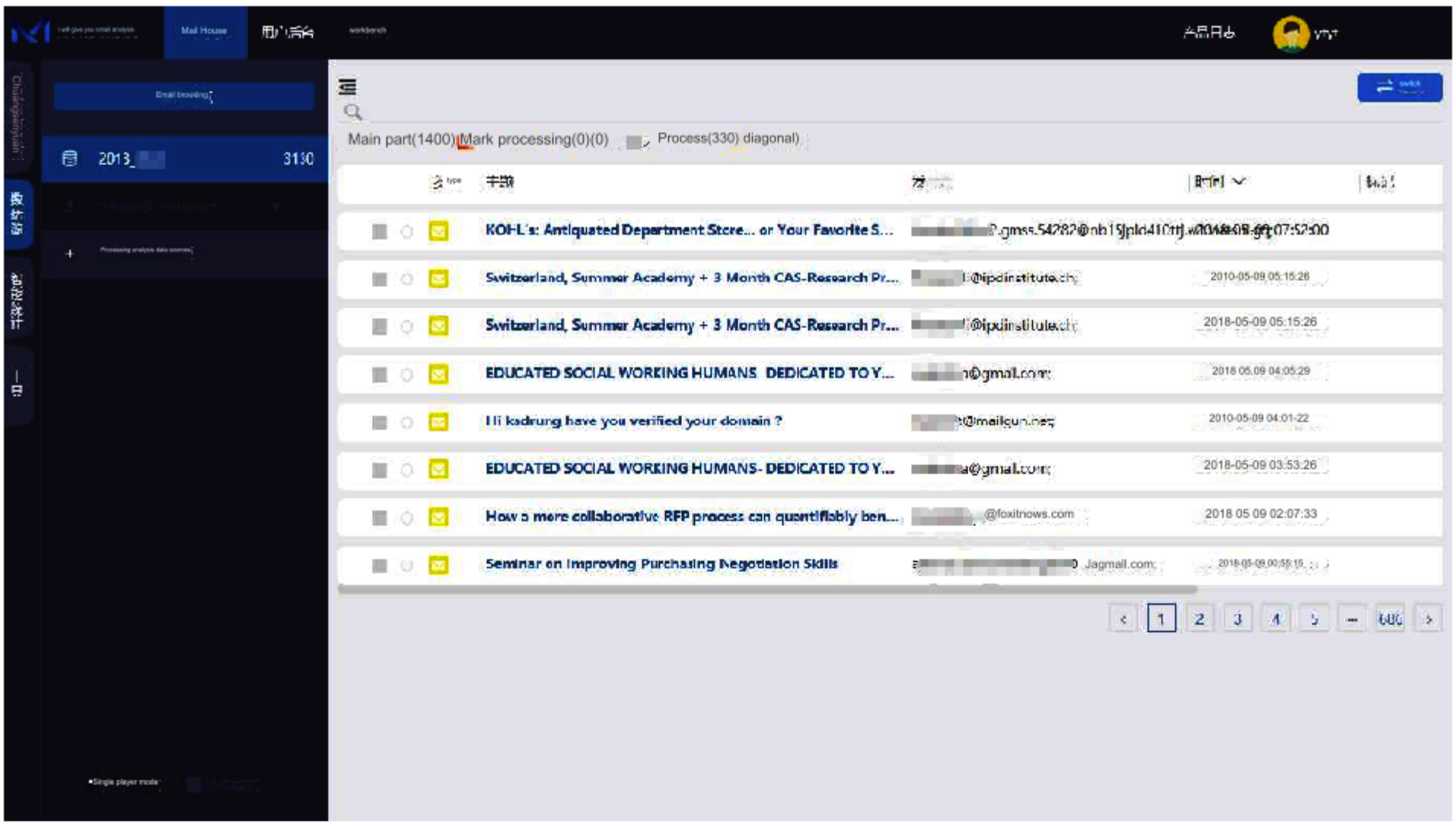
The "Mail Analysis Intelligence Decision System" provides automatic collection of mails. You only need to configure the target mailbox account, port, server IP address, password/security code and other information as required, and the system can automatically collect mails from the target mailbox account.



(email collection)

4.4 Email browsing

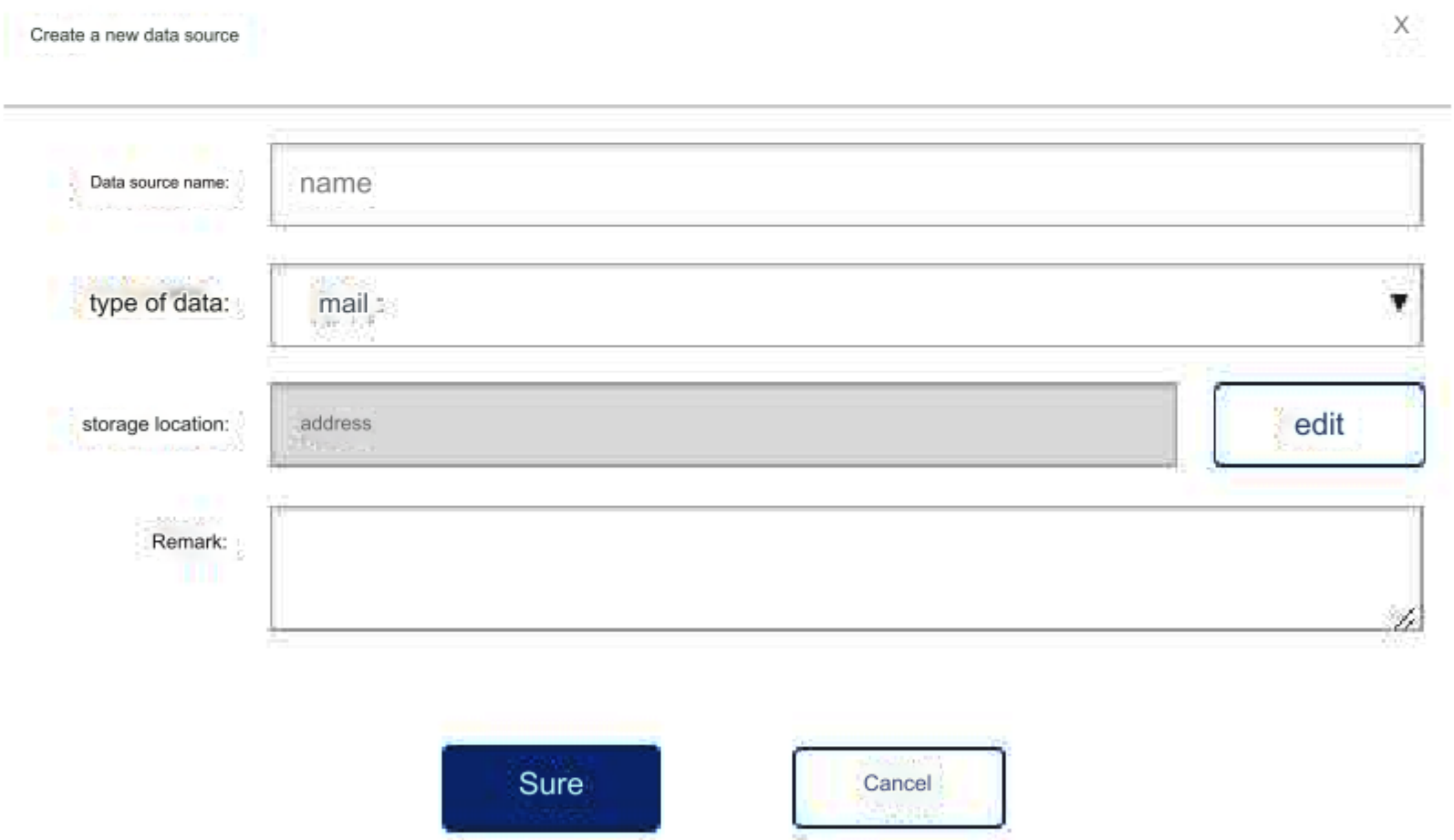
The email browsing function mainly implements the analysis and judgment function of original email data. It supports single-player mode or collaborative mode to achieve in-depth analysis and collaboration of emails. It can mark and comment on various email data to facilitate further tracking and analysis. It mainly supports the creation of four major modules: data source, data source, intelligent statistics and tools.



(Mail browsing interface)

4.4.1 Create data source

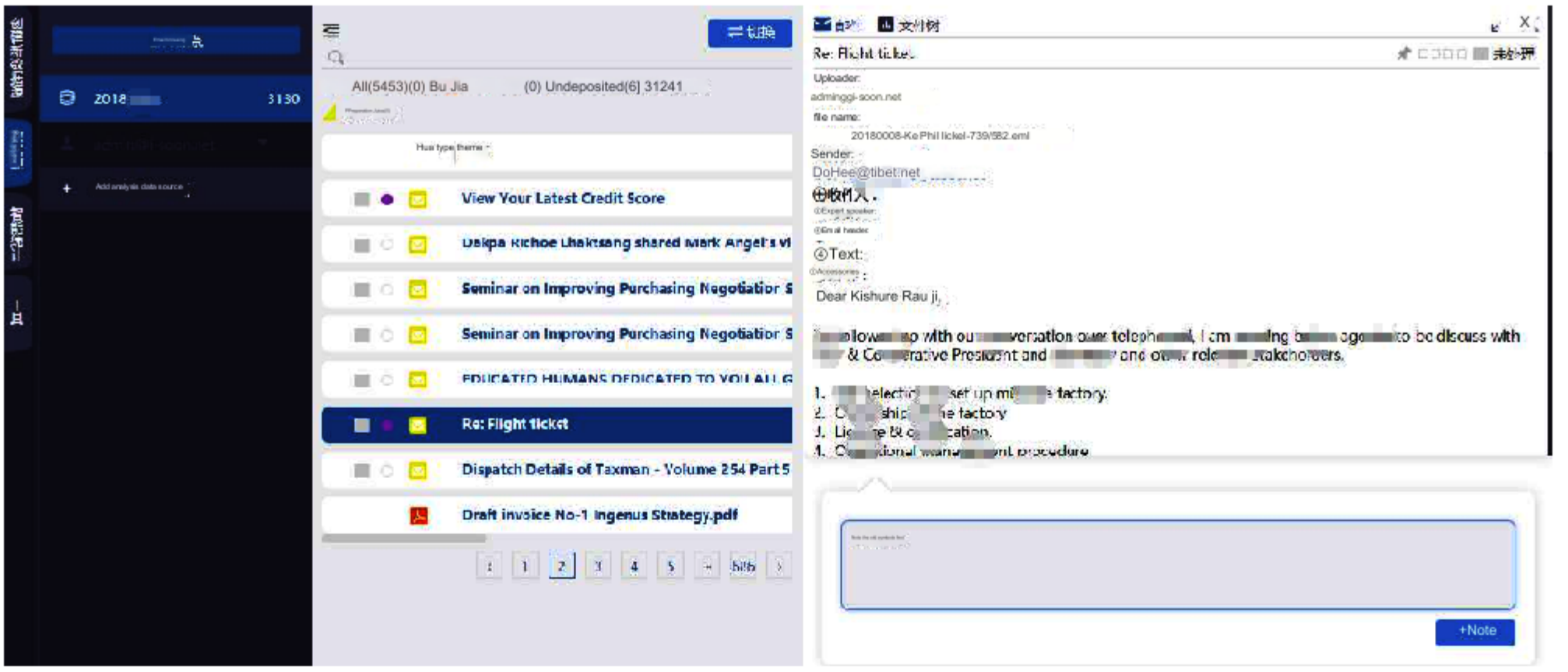
The email browsing function supports the direct creation of data sources and the direct import of relevant email data into the system.



(create data source)

4.4.2 Data source

The data source function is to view original email data, attachment data and other information from various data sources, and perform analysis operations such as annotation processing.



(data source analysis)

4.4.3 Intelligent statistics

Intelligent statistics is to conduct intelligent statistical analysis on the selected data source data, and master various statistical results such as file type, nickname, address, time, tag and other original data and post-analysis data under the data source.

File type: File type statistics include comprehensive statistics of emails, tables, documents, pictures, compressed files and unidentified types.

Nickname: It is the nickname of each email account in the data source for statistics.

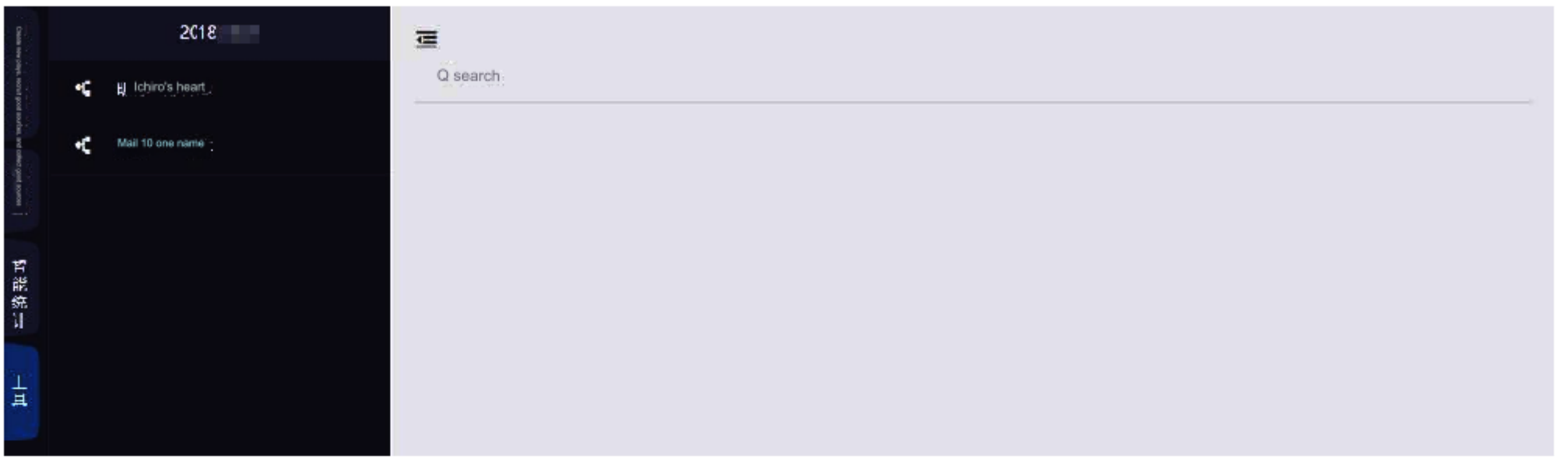
Address: Statistics on the sending email account and receiving email account in the data source.

Time: Automatically count the number of emails sent by the email data source under specific dates and time periods for each month and year.

Marks: Mark statistics for manual email analysis, such as processed, unprocessed, unprocessed, noted, and other types of mark information.

4.4.4 Tools

The tool module mainly provides two functions: retrieval of email IDs based on nicknames and retrieval of nicknames by email ID, which facilitates users to retrieve target emails based on nicknames and query target nicknames based on emails.



(Tool interface)

5 Product parameters

category	parameter
Architecture	C/S architecture
Email import format	eml format
Processing efficiency	100,000 emails/station (Under the condition of 1 server with standard parameter configuration)
Query response time	Second level response
Email relationship analysis	support
Conditional filter settings	support
Single target comparison analysis	support
Snapshot function	support
Comprehensive statistics	support
cluster analysis	support

6 Product Deployment

6.1 Applicable environment

As a professional email data processing and analysis platform, the "Email Analysis Intelligence Decision Platform" is suitable for users to conduct structured processing and analysis of massive email data, and quickly conduct analysis, judgment and collaborative analysis business scenarios. It is also suitable for users to establish their own massive email intelligence The application scenarios of the database can comprehensively improve the efficiency of business personnel's analysis and judgment of email data.

6.2 Deployment method

The "Email Analysis Intelligence Decision Platform" provides two application forms: public cloud and private cloud, so different deployment methods are adopted according to the different forms purchased by users.

> Deployment method of public cloud form of email analysis intelligence decision-making platform

The "Email Analysis Intelligence Decision Platform" public cloud provides users with authorized accounts to log in and use. Therefore, users do not need to purchase a server. They only need to provide a computer that meets the requirements and an available Internet network to access and use the platform.

> Deployment method of private cloud form of email analysis intelligence decision-making platform

"Email Analysis Intelligence Decision Platform" private cloud users need to deploy it locally to use it. Deploy the platform software on a server that meets standards, and then log in to the platform through a browser to import, analyze, and judge email data. The entire platform deployment environment requirements are as follows:

System parameters		
Server performance	operating system	Windows/Linux operating system
	CPU	CPU: 16 cores, main frequency ≥2.2GHz
	Memory	8G
	Hard drive capacity	2T
Cluster function		Supported, performance increases linearly after adopting cluster mode

7 product advantages

> High accuracy

The system uses big data architecture and intelligent text recognition technology to achieve rapid analysis, accurate extraction, and rapid comparison of massive emails.

> Powerful

The system supports various relationship analysis and value information extraction of target emails, including but not limited to email exchange information, geographical location information, communication information, activity information, etc.

> High availability

The system is stable and reliable, supporting 7*24 hours of uninterrupted operation. The system uses a graphical interface to operate, which is easy to use and has good usability.