Table of Contents

1. π-ProteomicInfo

π-ProteomicInfo is an automated working processing flow, mainly consisting of three modules: SearchServer, SearchClient, and sendInfo. Whole workflow including data transmission from the data acquisition computer to the database search computer, extracting file information based on database search results, quality control of mass spectrometry files, and sending information of quality control result.

2. SearchServer

The SearchServer module operates on a dedicated data processing workstation or server. It receives data and parameters from the SearchClient module and executes predefined data processing workflows. Upon completion of the data searches, a quality control information extraction program is automatically triggered using the comprehensive search results. The extracted quality control information is then relayed to the sentInfoUI module.
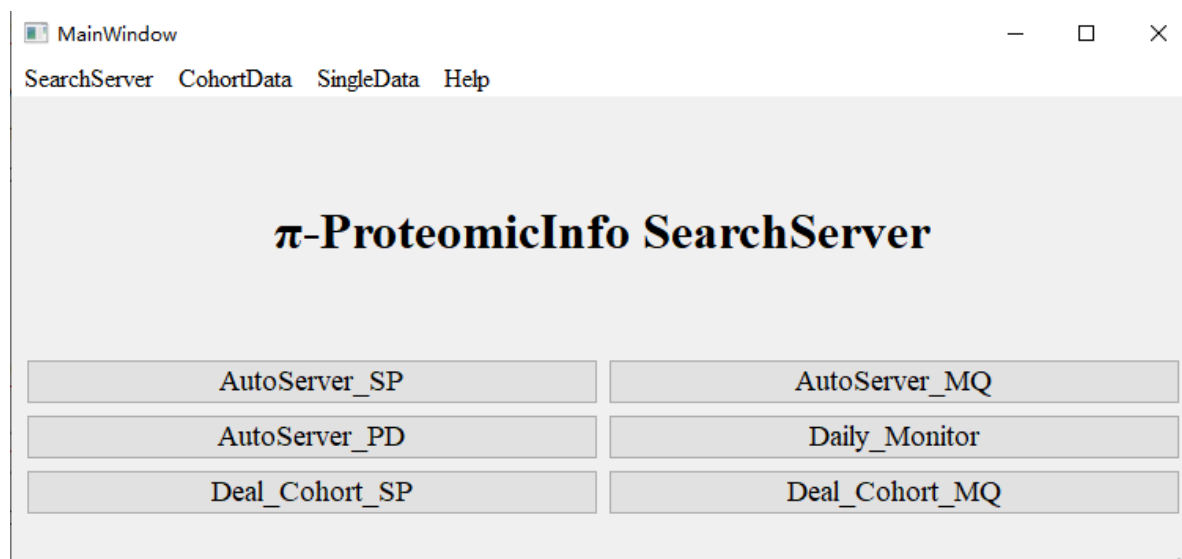
main window of SearchServer

2.1 Menu bar

The SearchServer menu bar comprises three sub-modules, each with a distinct function:

1. The **SearchServer** sub-module performs the automatic database search functionality of the software.
2. The **CohortData** sub-module handles the preprocessing of queue search results.
3. The **SingleData** sub-module organizes the results of single-needle automatic database searches.

The software is designed to support the concurrent operation of different sub-functions or multiple instances of the same sub-function.
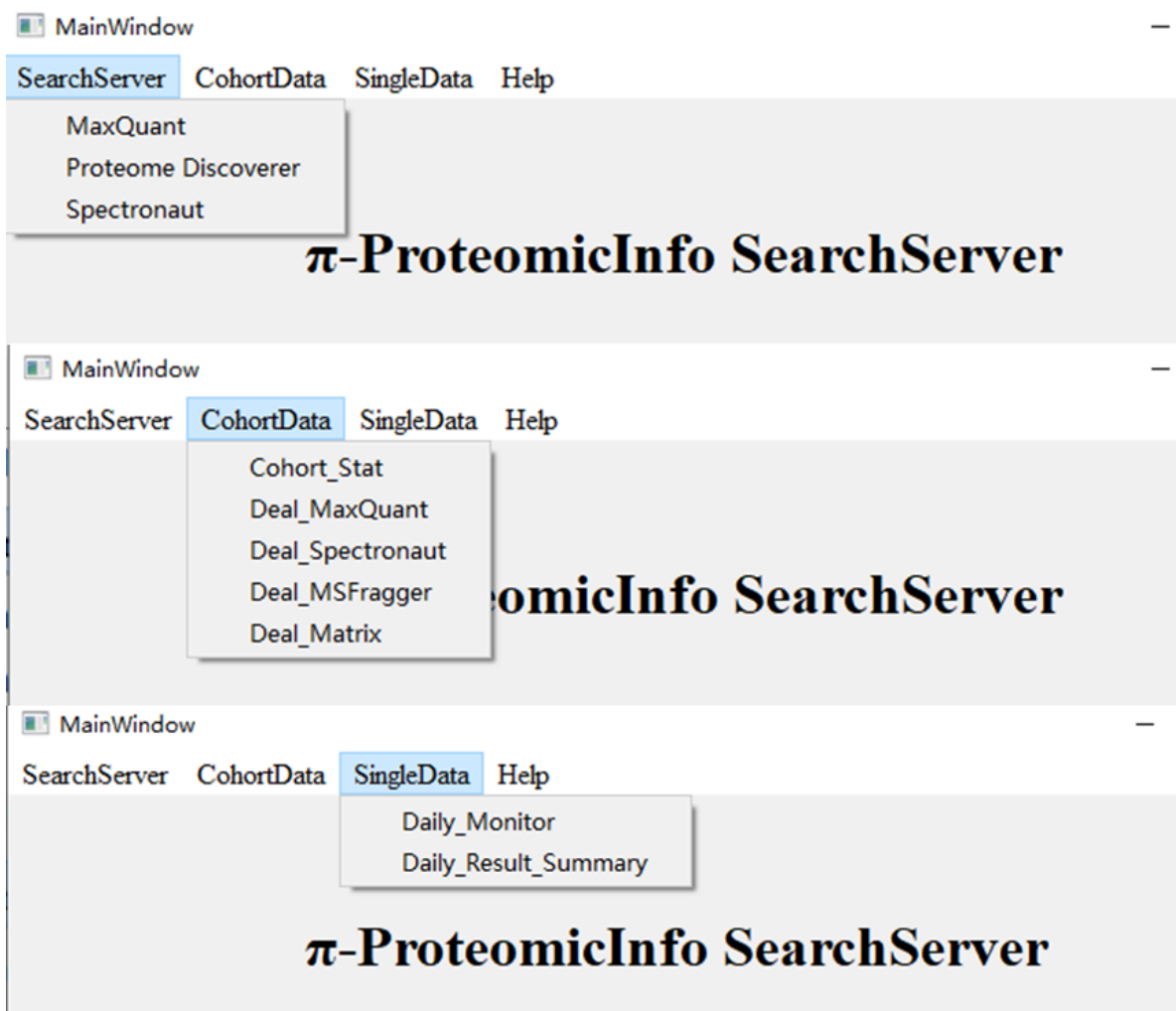
2.2 RunServer

The software assigns different running parameters based on distinct Port numbers. Data processed by a specific port on the mass spectrometry computer will be searched using the parameters associated with the same port on the database search computer. It is important to note that once the interface has successfully launched, any modifications to the parameters will not take effect for the current interface.

2.2.1 MaxQuant

Parameter File Generation

The parameter file can be generated through the user interface of MaxQuant software. After setting the parameters, save the file as mqpar.xml for subsequent use.

Data Transmission via Network Ports

After configuring the necessary parameters, users will find the newly generated mqpar.xml parameter file in the designated data reception path. This file should be copied to the data collection computer for use. For detailed instructions on how to use the file, refer to the corresponding client documentation.

Form — □ ×

MaxQuant

| Setting | Port | 13579 | **1** | (10000-60000) | **2** Add | **3** Query | ☐ Checked |

| MaxQauntCMD | **4** |
Received_Dir	**5**
ParamFile	**6**
Raw file	**7**
Experiment	**8**

Status | Statby | Run

1. The port set by the MaxQuant automatic database search operation program for data reception;
2. Parameter recording button. After clicking the button, the set port parameters will be recorded and used in conjunction with 3.
3. Parameter query button. After clicking the button, the parameters corresponding to the port number will be automatically displayed and used in conjunction with 2.
4. MaxQuantCMD.exe file path, the button can be used to search for the path;
5. Receive the data storage address, and the button can be used for path search.
6. The MaxQuant parameter document is the mqpar.xml file;
7. Automatically display according to the mqpar.xml file;
8. Automatically display according to the mqpar.xml file;

2.2.2 Proteome Discoverer (PD)

Parameter File Generation

The parameter files required for PD (Proteome Discoverer) automatic database search are identical to the corresponding study files, specifically the Consensus and Processing files. Users can configure the necessary modification and sequence information replacement steps directly within the software. It is crucial to include the "Result Exporter" node in the Consensus step to ensure the generation of subsequent output files. Once the settings are complete, export the corresponding files. For detailed usage instructions, refer to the description provided for the relevant client.
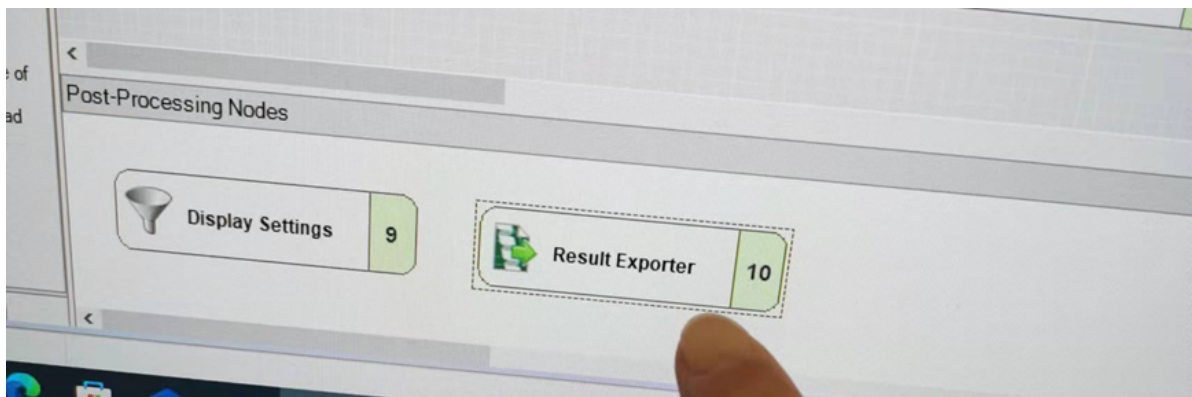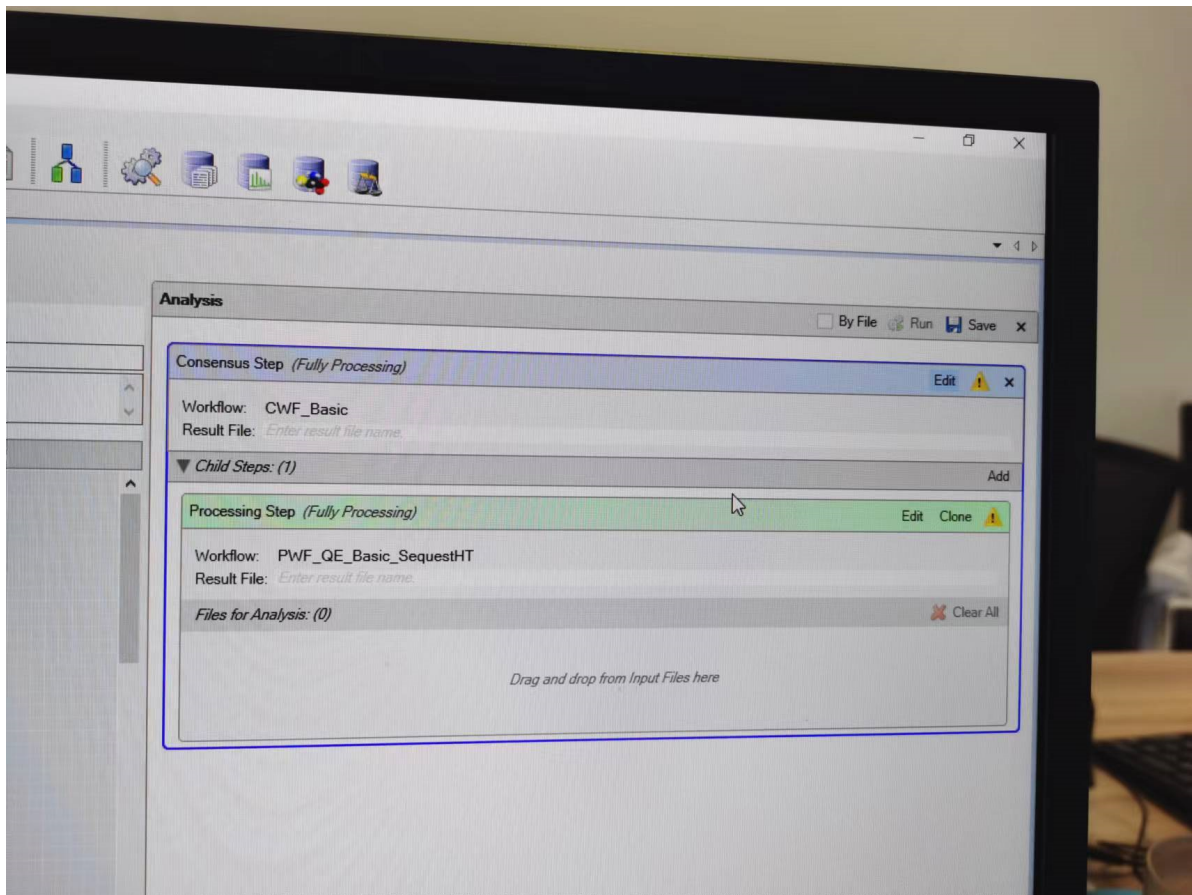
Add "Result Exporter"

Import the corresponding Consensus and Processing files in the Daemon, set the project name for file transfer, export the parameter files and copy them to the file collection computer for use.

Data Transmission via Network Ports

By setting the following parameters using the generated parameter file, the data processing of the preset parameters can be carried out.

1. The port set by the Proteome Discoverer automatic database search operation program for data reception;
2. Parameter recording button. After clicking the button, the set port parameters will be recorded and used in conjunction with 3.
3. Parameter query button. After clicking the button, the parameters corresponding to the port number will be automatically displayed and used in conjunction with 2.
4. Daemon.exe file path, the button can be used to search for the path;
5. Receive the data storage address, and the button can be used for path search.
6. Executable parameter files of the Proteome Discoverer software;
7. Computer username for database search;
8. Computer host name for database search;

Data Transmission via Daemon

When the file collection computer and the database search computer can be accessed through the Daemon software, users can skip the above Settings, enter the Proteome Discoverer Settings interface, and complete the Settings of the Daemon program, as shown in the following figure. Click administration in PD. In the Configuration at the lower left, select the Discoverer Daemon. In the interface on the right, choose the storage path for data reception. After completion, select Apply, then click Yes to save, and restart PD.
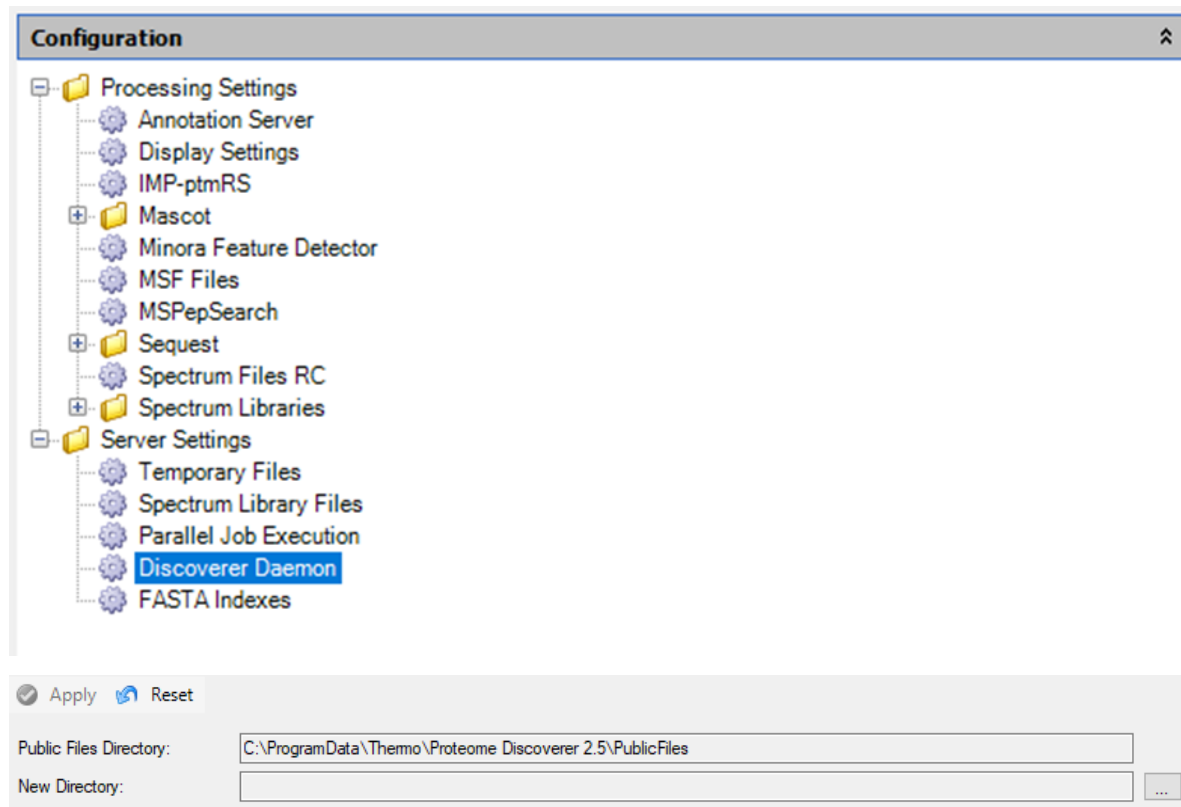
2.2.3 Spectronaut

Parameter File Generation

In the Spectronaut database search parameter file, a parameter setting file named "huiyan.rs" must be added to the **directDIA/Library** Settings section of the corresponding mode. Additionally, configure the export options to specify the content of the report to be generated. The saved report template (i.e., the *.prop file) will be utilized in the port settings. For detailed instructions on how to use this template, refer to the corresponding client documentation. The settings interface is illustrated in the figure below.
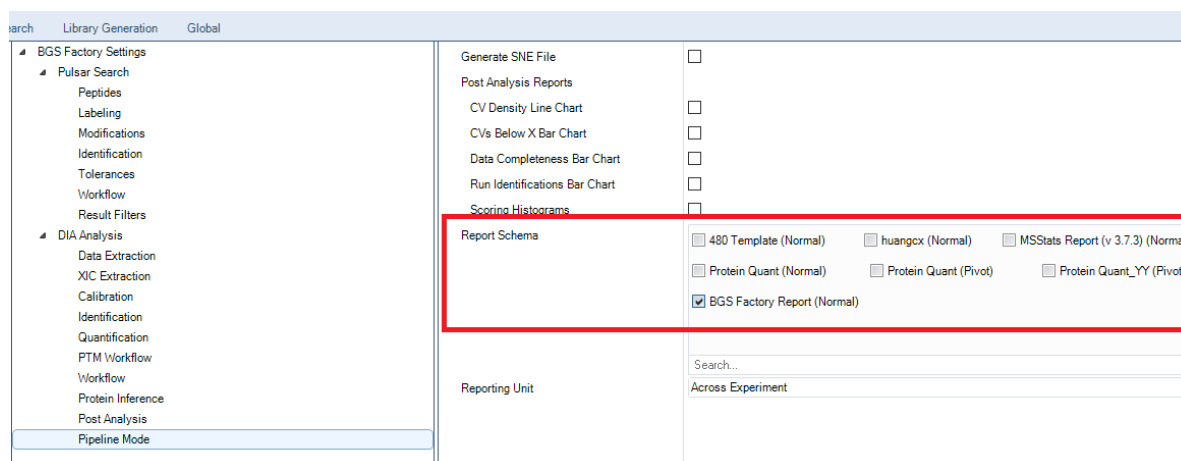
Data Transmission via Network Ports

By setting the following parameters using the generated parameter file, the data processing of the preset parameters can be carried out.


1. The port set by the Spectronaut automatic search program for data reception;2. Parameter recording button. After clicking the button, the set port parameters will be recorded and used in conjunction with 3.
2. Parameter query button. After clicking the button, the parameters corresponding to the port number will be automatically displayed and used in conjunction with 2.
3. Spectronaut.exe file path, the button can be used for path search;
4. Receive the data storage address, and the button can be used for path search.
5. Storage path for single-needle database search results;
6. The bgsfasta file used by Spectronaut for database search;
7. Optional, whether there are used library files;
8. Workflow files used by Spectronaut;

2.2.4 DIA-NN

DIA-NN is configured to run on a Linux server. The script is capable of performing single-needle database searches for most files during the collection period. Leveraging DIA-NN's ability to utilize single-needle "*.quant" files, the database search for cohort data will significantly reduces processing time. For raw files, the "ThermoRawFileParser" software must be installed in advance to convert the file format.

2.3 RunDaily

The module supports parallel operation. To launch it, click the "Daily_Monitor" button on the main interface or select it from the menu bar. The module's functionality is divided into two main parts:

1. **Database Search Result Extraction**: It collaborates with the automated database search program to extract information from database search results corresponding to single-needle data. Users can customize the result scanning software and set the frequency of result scanning.
2. **Mass Spectrometry Status Monitoring**: Based on a predefined reference template, it compares files that meet specified rules. By analyzing the extracted result information, it determines whether the mass spectrometry status is abnormal.

When the "Ignore Abnormal" option is selected, the function only performs the single-needle result information extraction function. When the "Report Abnormal" option is selected, users can achieve the evaluation of new mass spectrometry files based on reference data through steps such as adding and updating reference data.

2.4 RunStat

The data statistics module can be used for the original result statistics of single-needle and cohort data, and is compatible with the RunDaily and StatCohort modules.

2.5 ResultSummary

The sub-module performs the single-needle result merging and statistics function, which is used in conjunction with RunDaily results to facilitate users' statistics. The "Regular" option supports the simultaneous execution of multiple matching rules. Data that satisfy all rules will be counted, and different rules are marked with a semicolon **";"**. Separate, such as **"TOF2;" QC; "DIA"**.

2.6 StatCohort

The sub-module performs the statistics of the original database search results of the software.

1. For MaxQuant, Monitor_Dir selects the txt folder in the combine folder;
2. For MSFragger, Monitor_Dir selects the MSFragger result folder;
3. For Spectronaut, Monitor_Dir selects the software export file (the export module refers to the "huiyan_protein" template set by RunServer);

2.7 RunCohort

The sub-module executes data preprocessing functions for various inputs, including:

- The quantitative matrix derived from Spectronaut (specifically report, huiyan_ProteinQuant).
- The original database search results from MaxQuant.
- The original database search results from MSFragger.
- General quantitative matrices.

The preprocessing tasks encompass several key steps: **Quality Control/Normalization/Filling Missing Values/Basic Display of Results**: Providing a straightforward visualization of the processed data. The following figure illustrates the settings interfaces for the original export results from MaxQuant, MSFragger, and Spectronaut software, as well as the general quantitative

matrix (Maxtrix).

Example of MaxQuant Settings

MSFragger

Filter_Option
- ◉ Both_Filter
- ○ Sample_Number   Min_Identified_Number   `5`
- ○ SubGroup_Percent   SubGroup_Min_Percent   `0.7`

Unique peptides number (>) `1`   Protein Probability `0.99`

Data
- ◉ Intensity   ○ LFQ

Normalize
- ◉ Median   ○ Mean   ○ Quantile

Fill_NA
- ○ Minimum   ◉ KNN   ○ Quantile

Dim_Redution
- ◉ PCA   ○ UMAP   ○ T_SNE

Input   `I:\Example_Data\Res_MSFragger`

Name_Maping   `I:\Example_Data\Label.csv`

Output_Dir   `I:\Example_Data\PreDeal_Output`

☐ Checked   [Generate]

Example of MSFragger Settings

Spectronaut

Filter_Option	○ Sample_Number	Min_Identified Number	5
	○ SubGroup_Percent	SubGroup_Min_Percent	0.7
	◉ Both_Filter	Unique peptides number (>)	1
Normalize	◉ Median	○ Mean	○ Quantile
Fill_NA	○ Minimum	◉ KNN	○ Quantile
Dim_Redution	◉ PCA	○ UMAP	○ T_SNE
Input	I:\Example_Data\Result_SP18_Report.csv		
Name_Maping	I:\Example_Data\Label.csv		
Output_Dir	I:\Example_Data\PreDeal_Output		

☐ Checked   Generate

Example of Spectronaut Settings

Example of Matrix Settings

The "Name_Mapping" option is used to rename the original files. **The file must be in CSV format. It is recommended to save the CSV file using the "Save As" option in Excel**. Users can adjust the naming conventions for subsequent charts and specify the corresponding group information for each sample in this file, as illustrated in the table below:

#RawName	ChangeName	Group	Other
xxxx_A_xxxx_XXX_XX_1	A_1	A	XX
xxxx_A_xxxx_XXX_XX_2	A_2	A	XX
xxxx_B_xxxx_XXX_XX_1	B_1	B	XX
xxxx_B_xxxx_XXX_XX_2	B_2	B	XX
xxxx_C_xxxx_XXX_XX_1	C_1	C	XX
xxxx_C_xxxx_XXX_XX_2	C_2	C	XX

#RawName: The corresponding column name in the quantitative file. In MaxQuant, it is the set Experiment column name. The first column name is fixed and it is not recommended to modify it.

ChangeName: The content of the sample after renaming. Duplicate values are not allowed.

Group: Sample grouping information. The grouping situation will be displayed in some pictures.

Other: Other clinical information documents, not used;

3. SearchClient

SearchClient module operates on the mass spectrometry data acquisition computer, assessing the integrity of MS files at regular intervals to determine if they are ready for submission to the Server module for database search processing. This module also supports parallel execution with different parameter templates. Data transfer between modules is facilitated using either Socket tools or Daemon tools. In the sample template setting interface, clicking on the drop-down menu bar can display the preset examples.
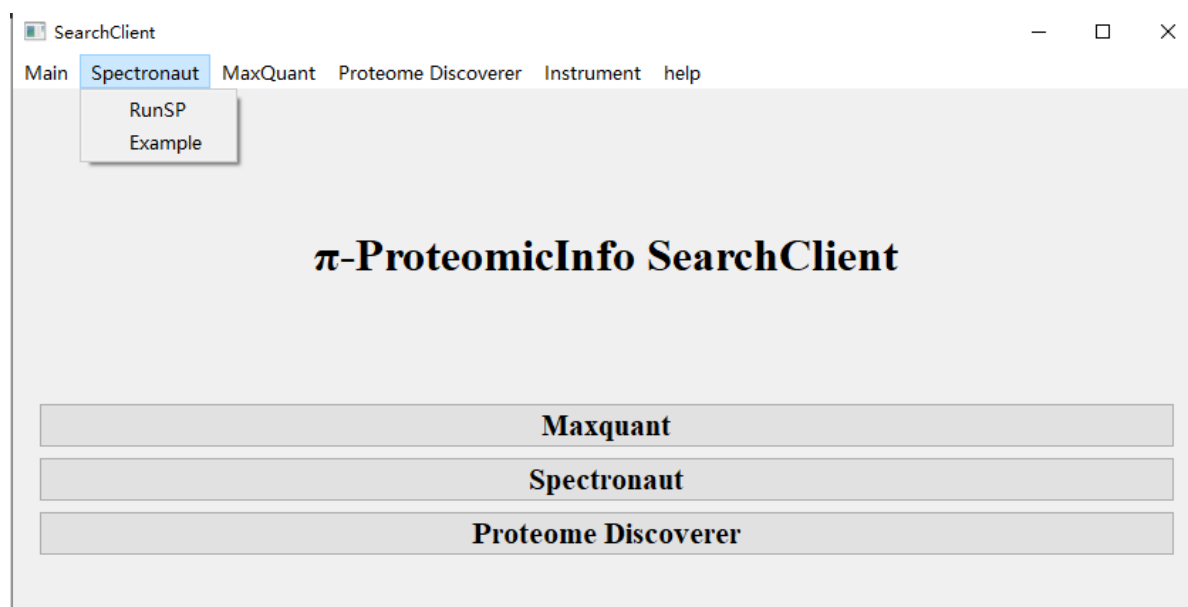
3.1 MaxQuant & Spectronaut & DIA-NN

The interface consists of the following six parts. After the parameter Settings are completed, the data monitoring for a specific folder and its subfolders will start to run. Old files will not be processed by default. When the collection of new MS files is completed, data transmission will start and the database will be automatically searched on the computer end of the database search.

DIA-NN does not need a specific client, and can complete data transfer with any client of MaxQuant or Spectronaut. Parameter files can be selected from any file :

Spectronaut

| Add_Computer | HPC 1 | Update_Computer | Run_Port | 3 |

| Data Storge | 4 |

| Acquisition Dir | 5 |

| Run_Param | 6 |

Status   Standby   Run

1. Data transmission purpose: computer;

1. Data transmission computer configuration: Users can update it based on IP addresses.
2. Data transmission port: This port must be consistent with the port set on the RunServer side to execute the parameter document set on the Server side.
3. Temporary storage location for data. After successful data transfer, the transferred files will be cleared. **It is strongly recommended to distinguish it from the storage location of the collected files.**
4. Data collection folder: Determine whether to start data transmission by monitoring the status of new collected files.
5. Parameter file, generated by the RunServer client;

3.2 Proteome Discoverer

The Proteome Discoverer (PD) interface has specific configurations. When the PD program on the database search computer can be accessed via Daemon software on the client side, data transfer can be facilitated using the Daemon program on the mass spectrometry computer. The path to the corresponding **daemon.exe** program is specified in **Daemon_Exe**.

If access through Daemon is not possible, data transmission via the Socket port must be executed. In this case, the **Socket** button must be clicked, and the corresponding parameters can be referenced from the settings described above for MaxQuant and Spectronaut.

Proteome Discoverer

| Add_Computer | HPC | Update_Computer | Run_Port | ⦿ Daemon ○ Socket |

| Acquisition Dir | |

| Run_Param | |

| Daemon_Exe | |

Status   Standby   Run

4. sentInfoUI

4.1 Overview

This program is designed to:

1. Receive mass spectrometry result files (Search Result) and instrument warning files (Warning) in real time.
2. Forward the extracted summary or warning information via **WeChat** or **Email** to the corresponding engineers.

Application Scenario:
 Daily operation monitoring of mass spectrometry platforms, enabling timely awareness of data processing results and instrument statuses.

4.2 Installation and Startup

1. **No Installation Required**: Simply download the ZIP package and extract it to any local folder.
2. Double-click `sentInfoUI_en.exe` (or `sentInfoUI.exe`) to open the main program interface. No additional dependencies are needed.

4.3 Interface and Parameter Explanation

The main interface consists of two sections: **Parameter Settings** and **Run Control**.

4.3.1 Parameter Settings

- **Send Method**
  Choose the method for sending notifications:
    - **WeChat**: Simulate mouse and keyboard operations to send messages through the locally logged-in WeChat client.
    - **Email**: Send messages through a configured SMTP server.
- **Daily Sending Time Period**
  Set the time range during which the program is allowed to send messages:
    - **Start Time**: Start time (e.g., 08:00)
    - **End Time**: End time (e.g., 18:00)
      Outside of this period, no messages will be sent.
- **Port Number**
  Specify the local port number the program will listen on to receive incoming result files. Ensure the port is available and not occupied.

- **Administrator**
  Enter the WeChat remark name of the administrator who will receive system status updates or error notifications. It must exactly match the remark name in WeChat.

- **Storage Location for Search Results**
  Select a local folder where received search result files or warning information will be saved.

- **File Extension**
  Define file extensions to distinguish different software outputs, for example:
    - `MQ_Stat_Summary.txt` indicates MaxQuant results
    - `PD_Stat_Summary.txt` indicates Proteome Discoverer results
    - `SP_Stat_Summary.txt` indicates Spectronaut results
    - `warn.txt` indicates warning files
- **Email Information**
  If Email mode is selected, fill in the following fields:
    - Sender email address
    - Authorization code or email password
    - SMTP server address (e.g., smtp.qq.com)
    - SMTP port number (typically 465 or 587)
- **Mass Spectrometer Information**
  Add the full name and abbreviation for each mass spectrometer. Abbreviations support multiple values separated by `|`. The program will identify the file source based on these abbreviations.
    - **Name**: Full name of the mass spectrometer
    - **Abbreviation**: Abbreviations found in file names, separated by `|`
- **Engineer Settings**
  Configure the mapping of engineers to their responsible mass spectrometers and projects. Specify their WeChat remark name or email address as needed.
    - **Engineer**: WeChat remark name
    - **Mass Spectrometer**: Associated mass spectrometer
    - **Project**: Project identifier (optional)
    - **Email**: Required only for Email mode
- **Software Field Settings**
  Choose the fields to extract from the search result files.
    - The left panel displays available fields (Options).
    - The right panel displays selected fields to be sent (Content to Send).
    - Use the `>>>` or `<<<` buttons to add or remove fields.
- **Save**
  Click this button to save all current settings into a local configuration file. They will be automatically loaded the next time the program starts.

4.3.2 Run Control

- **Run**
  Start the receiving and sending service. The program enters a "listening" state and will automatically send messages within the specified time period.

Functions:

- After startup, the program listens on the specified port for incoming files and automatically matches them to the corresponding mass spectrometer and engineer based on file extensions and the configuration tables.
- Within the designated "Daily Sending Time Period", the program will send generated summaries or warning messages to the appropriate recipients.

4.4 Typical Workflow

1. Extract and double-click `sentInfoUI_en.exe`.

2. In **Parameter Settings**:

   - Select **Send Method** → "Email".
   - Fill in **Email Information** (Sender Email, Password/Token, SMTP Server, Port).
   - Set the **Daily Sending Time Period** (e.g., 09:00–17:00).
   - Specify **Port Number** (e.g., 5000) and **Storage Location** (e.g., `D:\Results`).
   - Configure the **Mass Spectrometer** and **Engineer** tables, ensuring that project identifiers are correct.
   - In **Software**, select the result fields to be pushed and click **Save**.

3. Switch to **Run Control** and click **Run**.

4. When a file arrives, the program will automatically classify, summarize, and send out notifications.

4.5 Important Notes

- For WeChat mode, ensure the local machine is logged in to the WeChat client, and that the WeChat remark names exactly match those set in the "Engineer" table.
- For Email mode, use an SMTP authorization code or token instead of the regular login password.
- To modify configurations, first click **Stop** to halt the program, make adjustments, then **Save** and **Run** again.
- Make sure the firewall allows the program to listen on the specified port or add it as an exception.

5. Frequently Asked Questions (FAQ)

5.1 SearchServer

1. **How to confirm whether the automatic database search program is running normally?**

   - Check if there are received files in the data reception folder.
   - Verify if result files are being generated in the database search result folder. For MaxQuant, the result folder is the same as the data storage location. For Spectronaut, the result file is located at the set path (generated upon completion of the search). For PD, results can be viewed in the corresponding interface.

2. **How long does it take to see the search results after the program runs normally?**

   - Under normal operation, data will be automatically sent to the search computer to start the search approximately 1-2 hours after the default settings. The search time depends on the configuration of the search computer and the search software (the same as manual operation time).
   - The file scanning interval is set in the SearchClient. The path is the "Monitor_Time.txt" file in the Config directory where the SearchClient installation files are located.

3. **How to view the search results?**

   - Result files are stored in the corresponding software result folder. PD and Spectronaut results are set by the user, while MaxQuant result files are stored at the same address.
   - When the sendInfoUI module is running, the results extracted by RunDaily will be sent to the designated engineer via the sendInfoUI module.

4. **RunCohort module reports an error?**

   - The NameMaping file (sample column name renaming file) is used to modify the column names of subsequent processing files. Ensure that the selected file is not corrupted and is in CSV format. Regenerate the file if necessary.
   - Check if the Input options meet the requirements: MaxQuant and MSFragger inputs are result folders, while Spectronaut and Matrix inputs are individual files.

5.2 SearchClient

1. **How to add a search computer?**

   - Ensure that the search computer and the mass spectrometry computer are on the same network. Add a new computer through the PC_Setting option on the main interface. Alternatively, you can directly modify the "IP_Config.txt" file in the Config directory.

2. **Why does the interface close automatically?**

   - Check if the network connection to the search computer is stable.
   - Verify if multiple programs are monitoring the same folder or if another program is monitoring a subdirectory of the current folder.

3. **How to restart the program after an abnormal shutdown?**

   - For short-term abnormal shutdowns, users can simply repeat the previous operations to resume the transfer of samples during the interruption.
   - For long-term shutdowns, it is recommended to delete the "Complete_FileName_List.txt" file in the monitoring directory and restart the program. The program will then process the newly generated files for database search.

4. **How to perform automatic database search for files that have already been collected?**

   - Modify the "Complete_FileName_List.txt" file in the target folder. Files already recorded in this file will not be transferred to the search computer. By deleting the corresponding file names, the software will treat these files as new and execute the automatic search process.
   - When deleting multiple files at once, the transfer order is random. Only after all files in the current batch have successfully completed the process will their names be recorded in the "Complete_FileName_List.txt" file. If a failure occurs midway, adjust the content of the "Complete_FileName_List.txt" file to avoid repeatedly transferring files that have already been searched.

5.3 sendInfoUI

1. **Why did I not receive a WeChat message?**

   - Ensure WeChat is logged in and the program has sufficient permissions.
   - Verify that the WeChat remark name matches exactly.
   - Check if the current time falls within the configured "Daily Sending Time Period".

2. **Email sending failed with an authentication error?**

   - Obtain an SMTP authorization code from your email settings and update the **Email Password or Authorization Code** field accordingly.

3. **The file was not recognized?**

   - Verify that the **File Extension** configuration matches the actual file suffix.

- Ensure that the **Mass Spectrometer Abbreviation** matches the abbreviation used in the file name.