Documentation for installing and executing the Rapid Supply Network Extractor (RASNEX) including a test run

You will need the provided files:

- RASNEX 1.0 KNIME Workflow (program code) .knwf
- RASFF test dataset.pdf

Installation

Step 1:

Download and install the free "Konstanz Information Miner" (KNIME) from www.knime.com

Dedicate **as much memory (RAM) as possible to KNIME**. This will accelerate the processing, which is especially relevant when using multiple RASFF notifications. You can dedicate memory during the setup when installing KNIME ("Specific Memory Settings" during the setup process).

Step 2:

Execute KNIME and create a workspace repository (directory).

Step 3:

Within KNIME, use Use "File" → "Install KNIME Extensions" to install:

- KNIME Textprocessing
- KNIME Labs Extensions

Detailed instructions on how to install KNIME extensions can be found at https://www.knime.com/downloads/update (or search the web for "knime install extension").

Step 4:

Install FoodChain-Lab following the instructions at https://foodrisklabs.bfr.bund.de/installation/

If further extensions are necessary, KNIME will inform you in the console on the bottom right corner. Please install any missing extensions.

Step 5: Import the workflow

Use "File" → "Import KNIME Workflow" to import the provided file "RASNEX 1.0 KNIME Workflow (program code) .knwf". Use the "Select file" field. Click "Finish" to import the RASNEX 1.0 workflow into your KNIME workspace. The process may take a few minutes (especially if RAM settings have not been adjusted and only 2048 MB of RAM are dedicated by default for KNIME), even if no progress bar is shown.

Step 6: Load the workflow

Loading the workflow: The KNIME Explorer on the left hand side of the program will contain the "RASNEX 1.0 KNIME Workflow (program code)" now. Double click on the workflow to open it. Loading may take some time. You will now see a graphical depiction of the RASNEX 1.0 workflow.

Execution

Step 1: Execute the workflow

To execute RASNEX 1.0, you need to import the RASFF notification(s) you want to extract data from or visualize. Inside the yellow box "Importing RASFF notifications to use RASNEX 1.0 locally in the KNIME Desktop version", configure the "Tika Parser" node by double clicking it and confirming with "Ok". When working locally on your device, make sure to not reconfigure the "Tika Parser" node within the yellow box "Importing and uploading RASFF notifications to use the KNIME server application". This part of the workflow is only for working on a KNIME server. Select the directory of the saved RASFF notification(s), in the test case the directory where you have saved "RASFF test dataset.pdf". Execute the Tika Parser (right click -> "Execute")

Step 2: Inside the yellow box "Extracting relevant data" execute the Processing node (right click -> "Execute"). This may take some time. Afterwards, the third outport of this Processing node will contain the information RASNEX 1.0 extracted from the imported RASFF notification.

Step 3: Obtain the information from Table 1 of the manuscript (Sheets 1-5) using the yellow box "Generating the resulting tables and uploading the result of the workflow to the KNIME server application". Each of the five "RASNEX Sheet…" nodes corresponds to a Sheet. To display the information, for each node individually Right-click -> Execute and then use right-click -> "Connected to: Transposed Table".

Step 4: Visualize the extracted supply chain actors on a map using the yellow box "Mapping the extracted supply chain actors using FoodChain-Lab". An active internet connection is required. Execute the "Tracing View" node (right click -> "Execute"). KNIME will automatically queue and execute all previous nodes ("Preparation", "Station_Prep", "Deliveries" and "Prep"). Right click the "Tracing view" node and click "3: GIS Image" to show the resulting map of the supply chain actors. Clicking "Configure" instead allows you to zoom in and out of the map and to make further adjustments.

Next steps: You have now successfully executed RASNEX 1.0 on the test dataset and can follow similar steps to process your own RASFF notification(s). The following boxes provide troubleshooting information as well as the steps necessary to use an alternative GIS provider.

Troubleshooting boxes:

If the Station_Prep node from Step 4 is not successfully executed (traffic light turns red):

- Double click the "Station_Prep" node to get access to the contained nodes and to see the two "Geocoding" nodes.
- Use right click → "Configure" to configure the upper "Geocoding" node. The setting "Service Provider" allows you to adjust the provider of the mapping data and service. Photon is used as the default provider in the provided RASNEX 1.0 workflow. Different service providers are available, e.g. MapQuest or Gisgraphy. Choose the provider you would like to use and close the dialog by applying the changes (see Step 5 of the installation).
- Adjust the second "Geocoding" node accordingly. The workflow executes the upper Geocoding node first. The RASNEX 1.0 workflow uses Photon as a first and also second provider (default in both "Geocoding" nodes). Within the "Station_prep" node all addresses that could not be mapped by the first service provider (upper "Geocoding" node) are forwarded to the second Geocoding node. Here a second

provider may be used, for example Bing. The second Geocoding node adds all geolocations that could not be resolved by the first Geocoding node.

- If Geocoding nodes are not executed successfully, a license may be missing. Make sure that you complete Step 5 of the installation.
- If the Geocoding node is still not working, make sure that your Internet connection is working. The proxy entries are normally set to manual by the provider for HTTP and HTTPS (following "File" → "Preferences" → "General" → "Network Connections"). These settings may vary within the workplace (ask your IT department). Your Geocoding nodes should be working now.
- The "Station_Prep" metanode by default only allows to geocode up to five stations at a time. This setting may be adjusted by changing the "Last row number" in the Row Filter just before the lower Geocoding node.
- Sometimes it may happen that the Geocoding nodes are not able to find a corresponding geolocation for the extracted data from the RASFF notification. Right click -> "Coordinates" on the "Geocoding" node will display the extracted geographic coordinates (scroll to the right). If one "Latitude" or "Longitude" field is not filled out, manual adjustments are needed.
- You will need to use a map service (e.g. Google Maps or Bing Maps) to identify the geographic coordinates (latitude and longitude) for the related station.
- You can use the provided Java Snippets to import longitude, latitude or both. Therefore, just connect the relevant Java Snippet with the Concatenate node and adjust the lower "Row Filter" (the "Last Row Number" needs to be "1" in order to just geocode the station for which you just searched the longitude and/or latitude). Execute the "Station_Prep" node. Now that location has been added to the table containing all geographic coordinates identified so far. Reset the "Station_Prep" node by right click -> reset. Go back to the "Station_Prep" node. If another manual integration of a location is necessary, adjust the "Java Snippet" node with the geographic coordinates of that new (second) location. If you want to continue processing more locations, readjust the "Row Filter" node in front of the lower "Geocoding" node by setting the "Last row number" back to five and reconnect the "Geocoding" node to the "Concatenate" node again. Make sure that you do not execute the "Java Snippet" node again! This would adjust a coordinate by accident that just got extracted if the corresponding latitude or longitude could not be extracted. Those cases need to be separated and adjusted manually by the aforementioned procedure.

If not all extracted supply chain actors are mapped on the map:

- If the "Tracing view" node was successfully executed but not all extracted supply chain actors are displayed, open the "Station_Prep" node. Right click -> "Connected to: Concatenated table". This table displays the output of the "Station_Prep" node. Look at the top left corner and check for the amount of "Table "default" – Rows:" (14 in case of the provided RASFF test dataset). Now compare this number to the input of the "Station_Prep" node. To do so, right click -> "Connected to: Type-casted columns". The numbers of "Table "default" –Rows" should be equal. If those numbers are not equal, you will need to reset (right click -> reset) and execute (right click -> execute) the "Station_Prep" node until those two numbers are equal.
- If the two numbers are not equal after several iterations, follow the instructions under "**If the Station_Prep node is not successfully executed (lights underneath turn red)"** section. You may need to add some geographic coordinates manually.

Using a geographical information (GIS) service other than Photon (e.g. Mapquest or Bing):

- In the provided RASNEX 1.0 workflow, Photon is used as the default provider**.** To obtain access to GIS to transform addresses into geographic coordinates, some service providers (besides Photon) demand the creation of identifying keys. In case you want to use the open source provider Mapquest, an API key is needed. Search the web for “Mapquest developer key” and obtain your free API key. Within KNIME enter this key under File -> Preferences -> KNIME -> Geocoding -> “MapQuest API Key”. Similar steps apply for other providers that may be used, such as Bing.
- After successfully obtaining and approving your API key, you need to adjust the RASNEX 1.0 workflow. Double-click the “Station_Prep” node to open it. Right-click the upper “Geocoding” nodes to choose the source provider you would like to use. Save and use the workflow as stated above.