

KOREAN NOVEL GLOSSARY MAKER

BY: FOODIEMONSTER007



WHAT DOES GLOSSARY MAKER DO?

- This tool helps you create a translation glossary spreadsheet for Korean novels, to improve the quality of your machine translations. Here is an example of the output:

File Home Insert Draw Page Layout Formulas Data Review View Help

Calibri 11 A A

Wrap Text

General

Conditional Formatting

Format as Table

Normal Bad Good

Neutral Calculation Check Cell

Insert Delete Format

Cells

AutoSum

Find & Select

Find

Filter

Clear

Editing

Add-ins

Share

GB

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
1	hangul	hanja	chinese	english	frequency												
2	혁련일휘	赫連日輝	赫連日輝	Hyeokryeon Ilhwi	1												
3	제갈공명	諸葛孔明	諸葛孔明	Jegal Gongmyeong	1												
4	진하연	陳振連	陳振連	Jin Hayeon	54												
5	서지훈	徐志勛	徐志勛	Seo Jihun	7												
6	이지연	李智淵	李智淵	Lee Jiyeon	2												
7	김창석	金昌錫	金昌錫	Kim Changseok	4												
8	위진학	魏真虛	魏真虛	Wi Jinhak	2												
9	최제우	崔濟愚	崔濟愚	Choi JEU	1												
10	서완평	徐完平	徐完平	Seo Wanpyeong	1												
11	종리추	鍾離推	鍾離推	Jongri Chu	1												
12	일목	一木	一木	Ilmok	44												
13	태현	太賢	太賢	Taehyeon	5												
14	미륵	彌勒	彌勒	Mireuk	4												
15	순자	純子	純子	Sunja	3												
16	천양	千陽	千陽	Cheonyang	1												
17	예수	耶穌	耶穌	Jesus	1												
18																	
19																	
20																	
21																	
22																	
23																	
24																	

character names skills and techniques character titles locations and organizations item names misc

Ready Accessibility Good to go Display Settings

REQUIREMENTS

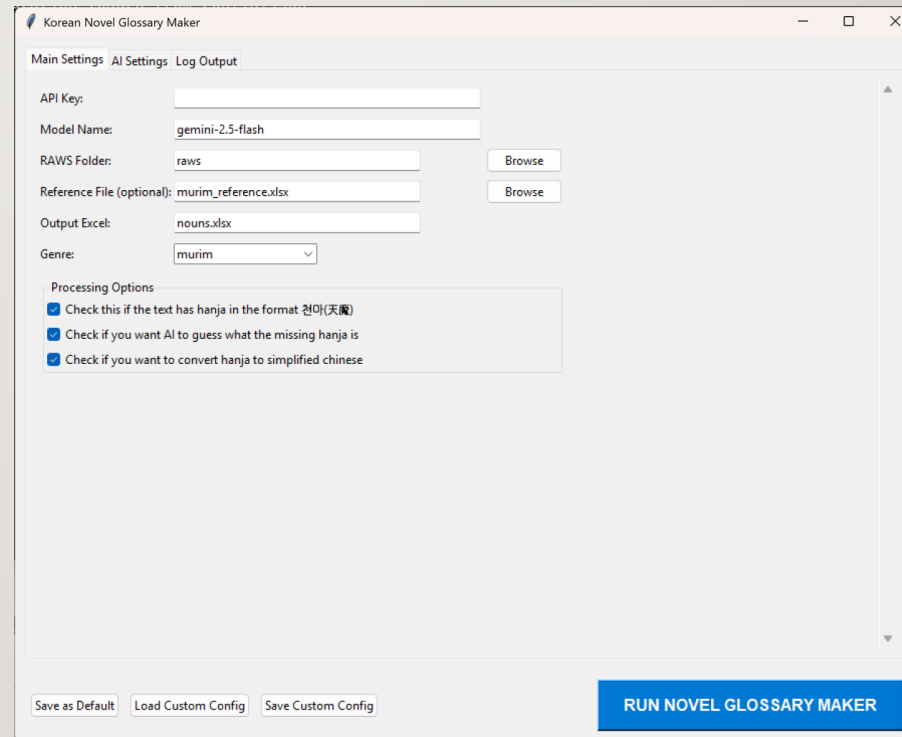
- 1) **Python:** This is a programming language that makes the tool work. You'll need to install it from the official website (<https://www.python.org/downloads/>). Make sure to check the 2 boxes that say "Add Python to PATH" and "pip" during installation. This step lets your computer easily find and run the program.
- 2) **Google API Key:** This is a special password that connects the program to Google Gemini AI. You can get a free one from the Google Cloud Console (<https://console.cloud.google.com/>). Here are the instructions on how to do it: <https://www.youtube.com/watch?v=brCkpzAD0gc>
- 3) **Run "setup.py"** by double-clicking it before you proceed to ensure everything is properly set up. You only ever need to run it ONCE. You should see the message "SETUP SUCCESSFUL!" in the command prompt window that pops up.

INSTRUCTIONS

- Put your novel chapter Korean raws in .txt file format inside a folder and copy it into the main directory. I have provided a folder named "raws" as an example.
- You can name the folder anything you want, but it must not have spaces. E.g. "Absolute_Regression" or "Solo-Leveling" are fine but not "Return of the Blossoming Blade"

INSTRUCTIONS

- Start the Program: Find the program file called **GLOSSARY_MAKER.py** and double click it. You should see the following screen:



The screenshot shows the 'Korean Novel Glossary Maker' application window. It has a title bar with standard window controls. Below the title bar are three tabs: 'Main Settings', 'AI Settings', and 'Log Output'. The 'Main Settings' tab is active and contains the following fields and options:

- API Key:** An empty text input field.
- Model Name:** A text input field containing 'gemini-2.5-flash'.
- RAWS Folder:** A text input field containing 'raws', with a 'Browse' button to its right.
- Reference File (optional):** A text input field containing 'murim_reference.xlsx', with a 'Browse' button to its right.
- Output Excel:** A text input field containing 'nouns.xlsx'.
- Genre:** A dropdown menu with 'murim' selected.
- Processing Options:** A section containing three checked checkboxes:
 - ☒ Check this if the text has hanja in the format 천마(天馬)
 - ☒ Check if you want AI to guess what the missing hanja is
 - ☒ Check if you want to convert hanja to simplified chinese

At the bottom of the window, there are three buttons on the left: 'Save as Default', 'Load Custom Config', and 'Save Custom Config'. On the right, there is a large blue button labeled 'RUN NOVEL GLOSSARY MAKER'.

INSTRUCTIONS – MAIN SETTINGS TAB

Korean Novel Glossary Maker

Main Settings Settings Log Output

API Key:

Model Name:

RAWs Folder: Browse

Reference File (optional): Browse

Output Excel:

Genre:

Processing Options

- ☒ Check this if the text has hanja in the format 천마(天魔)
- ☒ Check if you want AI to guess what the missing hanja is
- ☒ Check if you want to convert hanja to simplified chinese

Save as Default Load Custom Config Save Custom Config

RUN NOVEL GLOSSARY MAKER

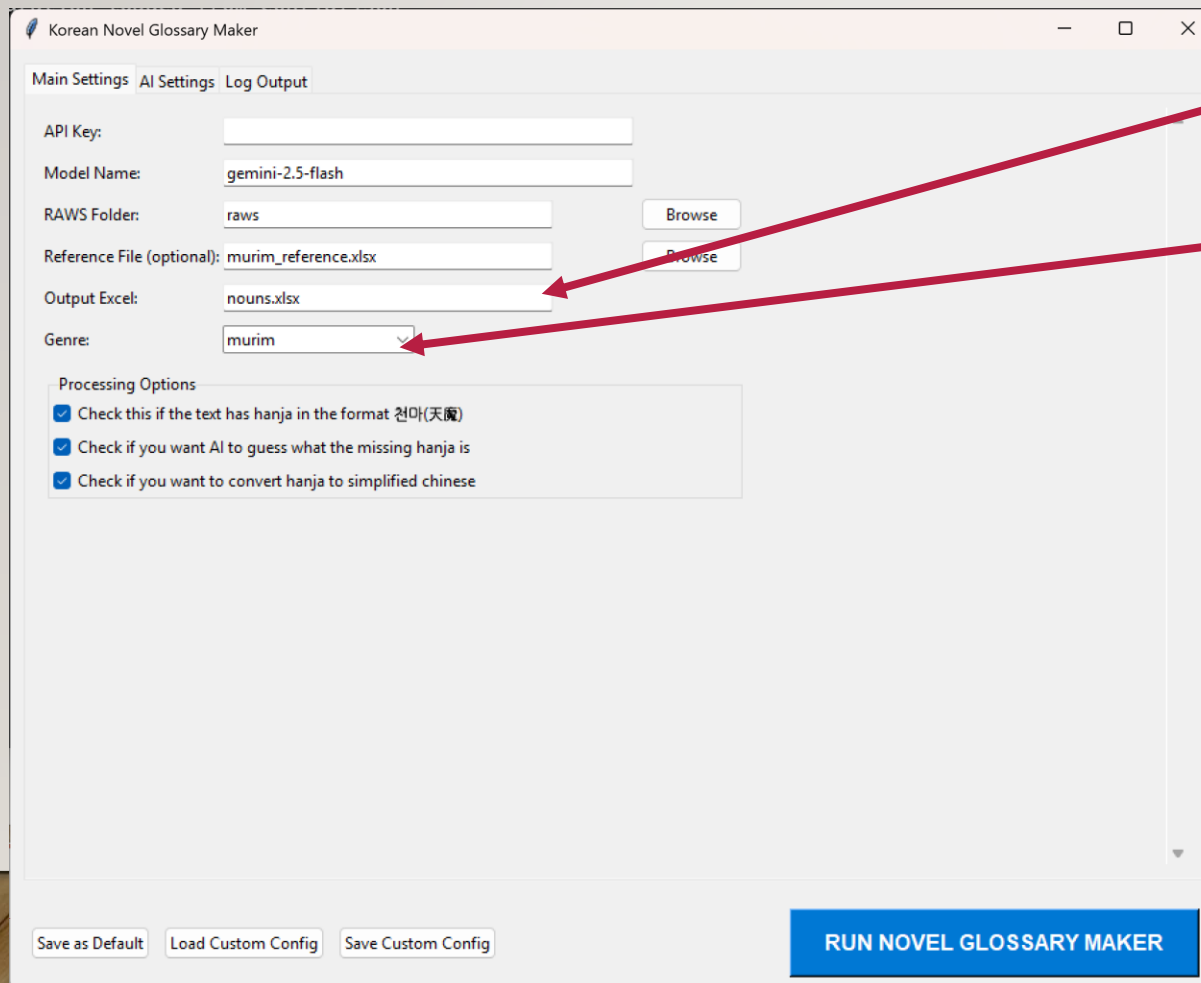
Paste the Google API key you got earlier here

Choose your favorite Gemini AI model.
Note not all models are available for free usage.

Type the folder name or browse to the folder where
you saved your Korean novel .txt files

If you are making a glossary for a murim novel, leave
this as is. If you have made a glossary manually
before, you can use your own reference excel
spreadsheet, but note that it must follow the same
format as murim_reference.xlsx. If the novel does
not belong to the murim genre and you don't have
your own glossary yet, delete the text and leave this
blank

INSTRUCTIONS – MAIN SETTINGS TAB



The screenshot shows the 'Korean Novel Glossary Maker' application window. The 'Main Settings' tab is active. The settings are as follows:

- API Key: (empty text field)
- Model Name: `gemini-2.5-flash`
- RAWS Folder: `raws` (with a 'Browse' button)
- Reference File (optional): `murim_reference.xlsx` (with a 'Browse' button)
- Output Excel: `nouns.xlsx`
- Genre: `murim` (selected from a dropdown menu)
- Processing Options:
 - ☒ Check this if the text has hanja in the format 천마(天魔)
 - ☒ Check if you want AI to guess what the missing hanja is
 - ☒ Check if you want to convert hanja to simplified chinese

At the bottom, there are three buttons: 'Save as Default', 'Load Custom Config', and 'Save Custom Config'. A large blue button labeled 'RUN NOVEL GLOSSARY MAKER' is at the bottom right.

Type your desired glossary file name here. Make sure it ends with `.xlsx`

Select the genre of your novel. We have:

murim: Chinese martial arts novel

rofan: romance fantasy novel

modern: modern day novel

game: MMO/VRMMO novel or system novel

westfan: western fantasy novel

dungeon: modern day novel with dungeons or towers

INSTRUCTIONS – MAIN SETTINGS TAB

Korean Novel Glossary Maker

Main Settings | AI Settings | Log Output

API Key:

Model Name:

RAWs Folder:

Reference File (optional):

Output Excel:

Genre:

Processing Options

- ☒ Check this if the text has hanja in the format 천마(天馬)
- ☒ Check if you want AI to guess what the missing hanja is
- ☒ Check if you want to convert hanja to simplified chinese

Check this only if your novel has hanja in brackets next to hangul

Not all glossary terms identified will have hanja. Check this if you want AI to guess the hanja

Check this if your novel has hanja, you asked AI to guess the hanja, AND you can read Chinese

Click this to save your current settings as default

Click these to save your current settings as a new custom config file, or load a custom config file that you saved previously

INSTRUCTIONS – AI SETTINGS TAB

The screenshot shows the 'AI Settings' tab of the 'Korean Novel Glossary Maker' application. The 'AI Settings' tab is highlighted with a red circle. The settings are as follows:

Setting	Value
Number of Chapters Analyzed Simultaneously:	5
Categorization Batch Size:	20
Translation Batch Size:	15
Hanja Guessing Batch Size:	15
Max Retries:	10
Retry Delay (seconds):	30

Annotations with red arrows and brackets provide additional context:

- A bracket groups the first four settings (Number of Chapters Analyzed Simultaneously, Categorization Batch Size, Translation Batch Size, and Hanja Guessing Batch Size) with the text: "Increase these numbers if you wish to use fewer AI credits and save money. Decrease these numbers if you keep seeing the MAX_TOKEN error in the output, or your output file looks strangely truncated".
- An arrow points from the 'Max Retries' setting to the text: "How many times to retry when you run into an AI error".
- An arrow points from the 'Retry Delay (seconds)' setting to the text: "If you are on a billing account (i.e. paying money for the gemini api), you can decrease this to 1 second. If you are on a free account, leave as is."

At the bottom of the window, there are three buttons: "Save as Default", "Load Custom Config", and "Save Custom Config". A large blue button labeled "RUN NOVEL GLOSSARY MAKER" is located at the bottom right.

Increase these numbers if you wish to use fewer AI credits and save money

Decrease these numbers if you keep seeing the MAX_TOKEN error in the output, or your output file looks strangely truncated

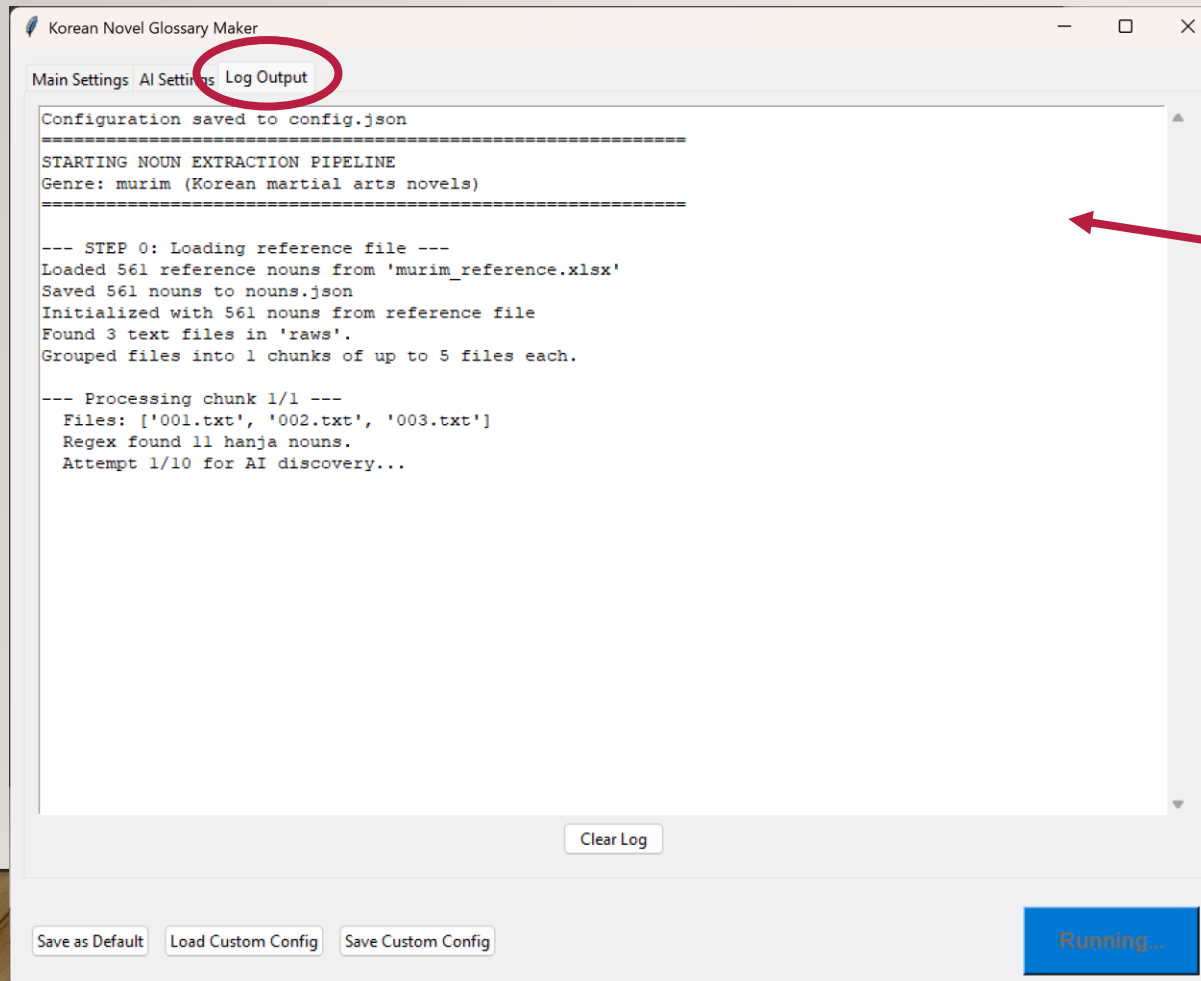
Number of Chapters Analyzed Simultaneously: How many novel chapters are analyzed at the same time. Decrease this if the novel chapters are long.

Categorization Batch Size: How many nouns are put into categories (names, locations, items) at the same time

Translation Batch Size: How many nouns are translated at the same time

Hanja Guessing Batch Size: How many nouns are submitted for hanja guessing at the same time (if enabled)

INSTRUCTIONS – LOG OUTPUT TAB



Observe your glossary making progress

Watch out for errors

PIPELINE COMPLETED SUCCESSFULLY!

- You should see a file named **nouns.json** in the folder. If you wish to continue adding nouns to the same novel (e.g. you made a glossary for chapters 1-10, and now you're updating it for chapters 11-20), this file will be necessary.
- nouns.json will be **overwritten** whenever you run the code, so if you are working on multiple novels, remember to **BACKUP** this file.
- You should also see **two new excel spreadsheets**. One of them has the terms in different categories in separate worksheets for easier visualization, while the other has every term on ONE worksheet for easier MTL-ing.
- You probably want to **check** the generated glossary spreadsheets manually and change any mis-translated terms and names. AI is stupid, it happens.

HAPPY MTL-ING!

CREDITS: FOODIEMONSTER007