

Memorandum

To: Users of Red Ink  
From: David Rosenthal, VISCHER AG  
Date: 6. August 2025 (updated continuously)  
Topic: Guide to Red Ink – Generation 2

---

**I. OVERVIEW ..... 2**

**II. USING RED INK..... 4**

A. Fundamentals..... 4

B. Basic functions of Red Ink in Word ..... 7

C. The Freestyle function within Word.....17

D. Search by topic: Context Search .....28

E. Access to other data sources and services: Special Services .....31

F. Transcriptor .....34

G. Creating Podcasts and Audiobooks .....40

H. Integrated anonymization function .....44

I. Word helpers – practical everyday helpers (almost) without AI .....48

J. Integrated chatbot "Inky" .....51

K. Further tips for using Red Ink in Word .....53

L. Red Ink functions in Excel .....55

M. Red Ink functions in Outlook.....61

N. Browser extension.....64

O. Using Red Ink in other programs.....66

**III. INSTALLATION..... 69**

A. For those who can't wait: The one-click installation .....69

B. The installation in more detail.....70

C. Preparation: API access .....72

D. Step 1: Download the installer/installations package .....73

E. Step 2: Run the installer .....75

F. Step 3: Initial configuration using the wizard .....76

G. Step 4: Install helper (optional, can be done later).....77

H. Step 5: Using add-ins .....78

I. Step 6: Making further adjustments as necessary .....79

J.	Installation of the browser extension .....	79
<b>IV.</b>	<b>CONFIGURATION (FOR ADVANCED USERS).....</b>	<b>80</b>
A.	Configuration file "redink.ini" .....	80
B.	Prompt library .....	99
C.	Other alternative language models .....	100
D.	OAuth2.0 (e.g. Google Vertex API).....	101
E.	Configuration of Advanced API Calls .....	102
F.	Security features .....	104
<b>V.</b>	<b>FAQS .....</b>	<b>107</b>
<b>VI.</b>	<b>RELEASE NOTES .....</b>	<b>118</b>
<b>VII.</b>	<b>ROADMAP .....</b>	<b>124</b>
	<b>ANNEX 1: IDEAS TO GET TO KNOW RED INK.....</b>	<b>125</b>
	<b>ANNEX 2: PROGRAMMING RESPONSE TEMPLATES .....</b>	<b>127</b>

## I. OVERVIEW

- Red Ink is an **AI tool** that we originally developed for our internal needs as a commercial law firm. It is integrated directly into **Word, Excel** and **Outlook** in the form of Office add-ins, where it allows a wide range of AI functions to be carried out with your own texts, spreadsheets and emails. Prompts can be entered directly, but it is also possible to have a selected text, cells in a worksheet or email chains translated, revised, summarised, commented on and searched by the AI. There is even a chatbot available, the tool can do transcriptions, create audiobooks from text and it can also be accessed in the browser.
- A special feature of the tool is that, unlike with "ChatGPT" or "Copilot", for example, each organisation can determine which language model from which manufacturer should be used. Not only common language models, such as those from OpenAI, Microsoft and Google, are supported in the cloud, but also open-source models operated on your own servers. This makes it possible to **control what happens to the data** and whether the data leaves your own premises; it is also possible to configure the tool in great detail to suit your own needs. We have no access to the data of other companies. The software's source code is also open-access and, in addition to Microsoft tools, only uses open-source libraries. This means that what the add-in does with the data is completely transparent. Each organisation has full control over



its content and can still allow all employees to use an "intelligent" agent as an everyday helper here and there.

- 3 The add-ins are programmed as COM add-ins for Word, Excel and (the classic, not the new) Outlook (in VB.net) and therefore only exist for **Windows**. The ready-to-install add-ins are digitally signed by VISCHER for security reasons (see below for further security features). This also applies to the two (optional) auxiliary add-ins for Word and Excel, which are programmed in VBA (i.e. in the macro language of Microsoft Office) and enable certain things that would otherwise not be possible (e.g. that the AI interfaces of Red Ink can also be accessed from your own Excel). Two additional optional extensions for the Edge and Chrome browsers has been developed in Javascript.
- 4 The add-ins are currently being beta tested and may be used free of charge during this period. Even after that, they should remain **free of charge for personal use**, in accordance with the licence conditions that are yet to be published. Free licences are also planned for commercial use, as are licences with only a moderate fee for larger setups.
- 5 The source code is available on GitHub (<https://github.com/VISCHER-LIL/redink>). The website <https://vischer.com/redink> contains further information and a download link, as well as a **demo video**.
- 6 However, the performance of Red Ink is ultimately **only as good and fast as the language model used**, because that's where the results are produced. We have found significant differences between the models. There are also some system-related hurdles, such as the fact that language models are not designed to process texts with formatting, and Word and Outlook do not make it easy to find workarounds for this. We have implemented a variety of tricks and methods to deal with this as best as possible – and we also hear that Red Ink apparently does this better than some other well-known tools.
- 7 That is why we recommend that everyone try it out for themselves to see how and where the tool can best help them. There are some **ideas to try out** in the appendix at the end, and the demo video is also helpful. Of course, the results provided by the AI should be checked for accuracy. Any shortcomings and, of course, **suggestions for additional features** can be reported directly to the author at [david.rosenthal@vischer.com](mailto:david.rosenthal@vischer.com).
- 8 The functions (para. 11 et seq.), the installation (para. 196 et seq.), the configuration (para. 246 et seq.) and other aspects such as the security features are described below. These instructions are available in German and English. We do not yet offer any guaranteed **support** – we are still a law firm and the author of Red Ink is the full-time head of the Data & Privacy Team at VISCHER. We are happy to advise on questions of AI law, or if a contract is to be concluded with one of the language model providers in a manner that is data protection compliant and, as in our case, even with professional confidentiality.

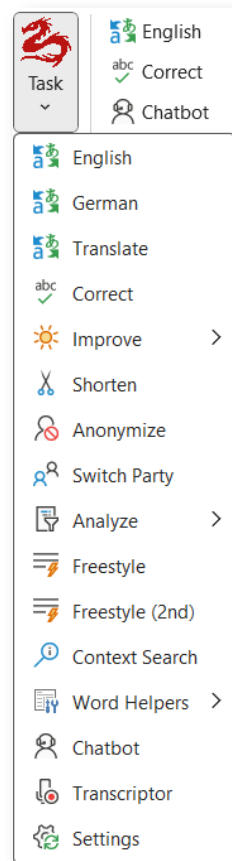
- 9 If, like most people, you don't run a language model on your own server, you need one to use Red Ink, because it has to be configured to a corresponding so-called API. Well known providers are OpenAI, Microsoft and Google; for our purposes, we mainly use Google. Although there is a charge for this, in our experience the costs are very low – much lower than if each employee were to have a subscription to some of the well-known services. It should also be mentioned that using Red Ink in an organisation does not require a login and is not recorded by Red Ink itself.
- 10 These instructions only apply to Red Ink **Generation 2**. This is the version from December 31, 2024.

## **II. USING RED INK**

### **A. Fundamentals**

- 11 A **demo video**, showing how to use Red Ink is available on the website <https://vischer.com/redink>.
- 12 At their core, the add-ins work by selecting text or cells and then asking the AI to do something with them. Exactly what that is depends on whether the add-in is being used in Word, Excel or Outlook. Red Ink is available via two tiles. In Word, Excel and Outlook, they appear in the main display, and in Outlook also when an email is open for writing (i.e. when composing, replying to or forwarding an e-mail or opening a draft):

VISCHER

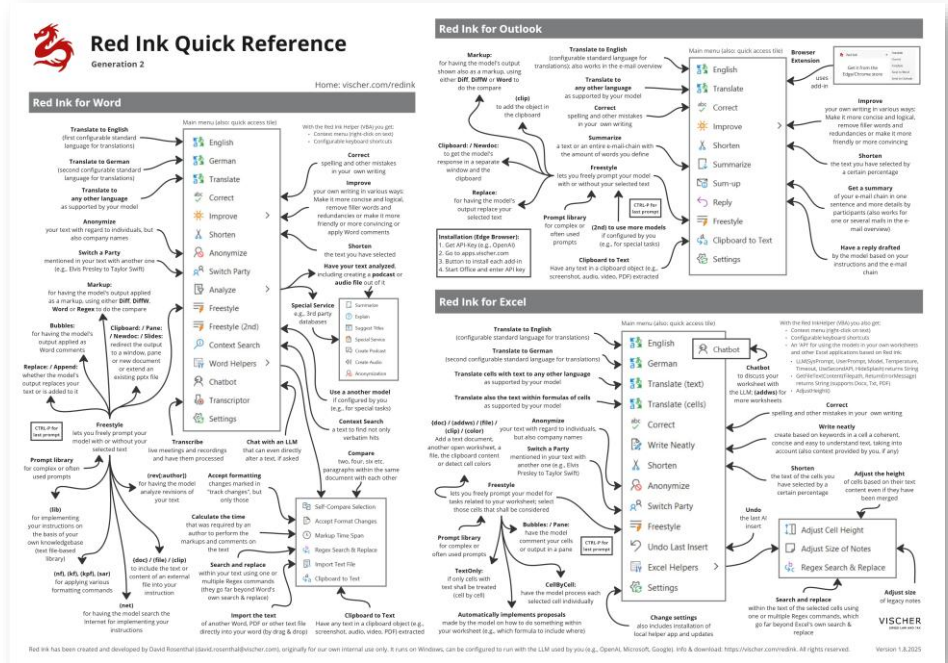


- 13 One of the tiles is used for the three most frequently used functions (for each Office application) for quick access (the language on the button for quick access to the translation function can be changed). The other gives access to the functions as soon as the logo is clicked (if you leave the mouse on it, the current version and the currently configured language model are displayed). The tiles can be repositioned in the respective applications (to the extent allowed by the system administrator).
- 14 In Word and Excel, the functions can also be selected using the context menu and certain keyboard shortcuts. The context menu appears when you right-click on selected text or cells:



- 15 The keyboard shortcuts can be programmed in the configuration settings. However, the context menus and keyboard configurations require two system-related utility programs (VBA add-ins) to be installed during installation (see para. 228 et seq. below). Red Ink will run without them, but if they are missing, the context menu simply will not appear. If keyboard shortcuts are defined, they will be displayed when you move the mouse over the menu item.
- 16 However, depending on your needs, you can also access the AI directly without selecting a text, for example to enter a prompt that is independent of the text or to have an entire email summarised. This way, you no longer require to switch to a separate chat while writing, and the result can be used immediately.
- 17 The output is normally displayed in the current document, worksheet or email. Certain functions (Freestyle) also allow output in a separate window and to the clipboard. The add-in for Word also has an integrated chatbot with its own window.
- 18 All three add-ins offer a range of functions for predefined tasks (e.g. translating, correcting), but can also be used with your own instructions via the Freestyle function. The Freestyle function in particular has numerous options in Word (such as commenting on markups or including external documents) that some other tools do not offer. Furthermore, a custom prompt library can be used in Red Ink. Incidentally, all prompts used by Red Ink can be changed and customised to your needs via the configuration file.
- 19 For Chromium-based web browsers (e.g., Edge, Chrome), there is also an extension that allows you to send text directly to the add-in in Outlook and process it there (e.g. for translation or correction) from anywhere in the browser where text can be edited or selected.
- 20 Red Ink itself is only available in English, but it can process texts in any language that the respective language model supports. The built-in prompts are also written in English, but they can be changed (normally this is not necessary).

- 21 If you don't want to read the whole manual, this short overview may help (it is included in the installation package for printing and can be downloaded from <https://vischerlnk.com/redink-gr>):



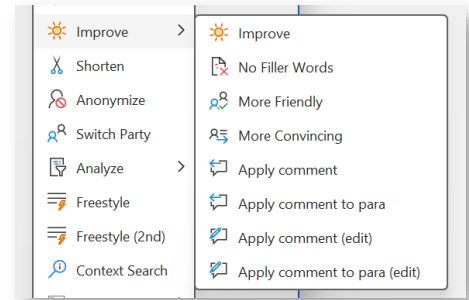
## B. Basic functions of Red Ink in Word

22 The predefined AI functions are:

- To English, To German, To Other:** The selected text is translated and replaced with the translation. To Other can be used for a translation into numerous other languages. It just needs to be specified (in English), e.g. "French" rather than "Français". The two languages "English" and "German" are preconfigured in the menus but can be changed. Two remarks: The AI is instructed not to insert double spaces after punctuation marks (as was previously common in certain languages). It is also instructed to maintain the style of the original text when translating, and to use a formal style in cases of doubt. The English "you" is translated to "Sie" on this basis and not to "Du", unless the text contains indications of informal language such as addressing or greeting by first name only.
- Correct:** The selected text is linguistically corrected, i.e. not only spelling mistakes but also other errors, such as unsuitable words or incorrect punctuation marks.
- Improve:** Here, the text is also proofread and improved in terms of content, but without adding new information. For example, suggestions are made to make it easier to understand. Besides Improve, the same menu also offers the options **No Filler Words** to remove filler words and redundancies, **More Friendly**

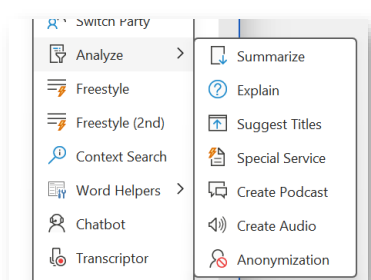
to make the selected text friendlier, and **More Convincing** to make it more persuasive.

A special function is **Apply Comment**. It can implement the content of a Word comment (i.e. a "bubble" or "balloon") in the document with the help of AI, e.g., make a correction or implement an addition described in the comment. For this, the comment bubble must be selected. All four variants implement the entire comment (or, if selected, the selected text within) either to the location



marked by the comment or the entire paragraph (or paragraphs) in which the comment is located ("to para"). The "(edit)" variant allows the user to edit the prompt for insertion beforehand. This function can be combined with the "Bubbles:" prefix of Freestyle, letting the AI comment on your document and then have the comments implemented with this function.

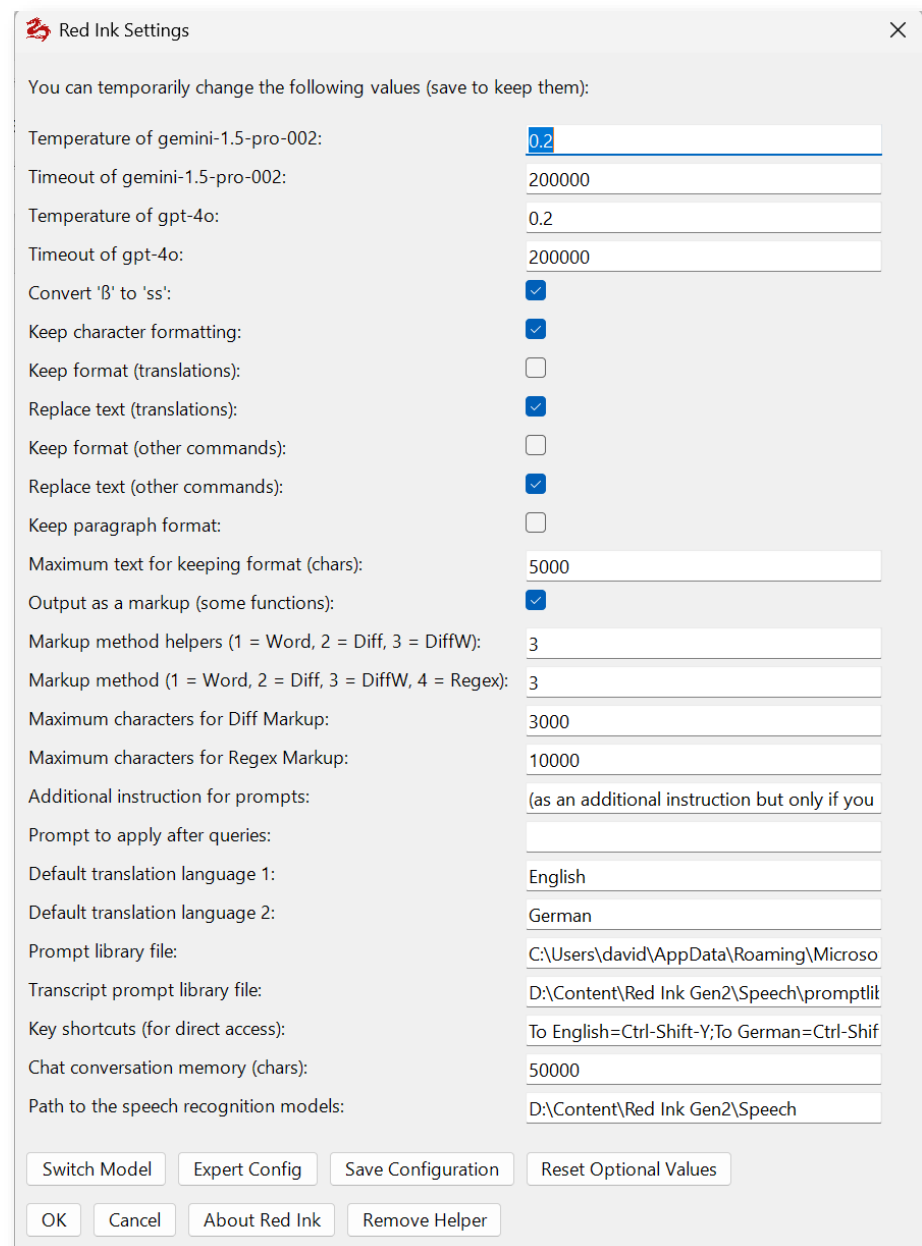
- **Shorten:** This function can be used to shorten a text, with the aim of losing as little information as possible, or only the information that is considered less important. The user can specify the percentage by which the text should be shortened; the AI, however, may not actually adhere strictly to such length specifications.
- **Anonymize:** The selected text is anonymised with regard to natural persons, but also organisations. The placeholder "[Redacted]" is inserted at the relevant places. Red Ink will propose the Regex markup method (see below), which may be better than others for larger texts.
- **Switch Parties:** This function allows for the "intelligent" replacement of references to specific persons in contracts, legal documents and other texts. For example, "the Provider" can be substituted for "the supplier of goods", whereby the entire formulation is taken into account and adapted. This is not possible with a normal "search and replace". When using this function, it may make sense to try the Regex markup method (see below). Red Ink will propose this method, should it not already be configured.
- **Analyze:** Here is a summary of various commands that analyze and process the selected text in one way or another. **Summarize** summarizes the text. This allows for a quick overview. The summary





is inserted at the end of the text. It is possible to specify how many words the summary should have. If **Explain** is selected, the tool also provides a short summary but goes into more detail about the things to be done according to the text, the arguments and logic of the author, and provides explanations of the technical terms and topics that the model itself knows. **Suggest Titles** suggests titles for the selected text, three different titles for different use cases (memo, blog, informal text, humorous text, food for thought). In these last two functions, the text is displayed separately, not inserted (but can be edited and thus copied to the clipboard). **Podcasts** and **audio recordings** of texts can also be generated via the menu. More on this is in para. 92 et seq. below. Any configured **Special Service** can also be accessed. More on this is described in para. 70 et seq. below. Finally, **Anonymization** can be used to test the built-in anonymization function, which can be used for the transparent anonymization of texts that are transmitted to the language models (see para. 110 et seq.).

- 23 If you want to **Undo** an insertion or change made by Red Ink, you can use the Undo function of Word, but you will have to do multiple Undos because Red Ink replaces and inserts text in several steps.
- 24 How these functions handle the text, i.e. whether they replace it or whether the output is appended, whether a markup (comparison version) is created and with which method, and whether and how an attempt should be made to preserve the existing formatting, can be controlled via the configuration file or the **Settings** function. It can be accessed via the tile menu and can be used to make the changes temporarily or to save them in the configuration file for future sessions (an explanation of each option is displayed when you move the mouse over the text):



**Red Ink Settings**

You can temporarily change the following values (save to keep them):

Temperature of gemini-1.5-pro-002:	0.2
Timeout of gemini-1.5-pro-002:	200000
Temperature of gpt-4o:	0.2
Timeout of gpt-4o:	200000
Convert 'ß' to 'ss':	<input checked="" type="checkbox"/>
Keep character formatting:	<input checked="" type="checkbox"/>
Keep format (translations):	<input type="checkbox"/>
Replace text (translations):	<input checked="" type="checkbox"/>
Keep format (other commands):	<input type="checkbox"/>
Replace text (other commands):	<input checked="" type="checkbox"/>
Keep paragraph format:	<input type="checkbox"/>
Maximum text for keeping format (chars):	5000
Output as a markup (some functions):	<input checked="" type="checkbox"/>
Markup method helpers (1 = Word, 2 = Diff, 3 = DiffW):	3
Markup method (1 = Word, 2 = Diff, 3 = DiffW, 4 = Regex):	3
Maximum characters for Diff Markup:	3000
Maximum characters for Regex Markup:	10000
Additional instruction for prompts:	(as an additional instruction but only if you
Prompt to apply after queries:	
Default translation language 1:	English
Default translation language 2:	German
Prompt library file:	C:\Users\david\AppData\Roaming\Microso
Transcript prompt library file:	D:\Content\Red Ink Gen2\Speech\promptlit
Key shortcuts (for direct access):	To English=Ctrl-Shift-Y;To German=Ctrl-Shif
Chat conversation memory (chars):	50000
Path to the speech recognition models:	D:\Content\Red Ink Gen2\Speech

25 In detail:

- **Keep character formatting** tells the add-in to provide the AI with the most common word formatting, such as bold, italic, and underlined, in a format it understands (Markdown), so that the output is also formatted accordingly. However, this does not always work, and in combination with markups, this function is only active with DiffW. This function is enabled by default. This function is also limited by the number of characters for the diff markup, otherwise the program takes too long. Caution: Formatting can still disappear if it is overwritten by the current paragraph formatting, which Red Ink also tries to preserve.
- **Keep format** tells the add-in that it should not only send the selected text to the AI, but also the basic formatting (such as bold

or a list) stored in it (temporarily in HTML). If the text to be translated is "Wir haben *viel* Spass", the AI will deliver "We are having a *lot* of fun" if it follows the instructions. However, it should be noted that this functionality is associated with a lot of additional data and is therefore not suitable for large texts (the add-in may warn you if necessary) because the AI can be overwhelmed, and it takes a lot of time. To keep the effort (and thus the waiting time) within limits, not all formatting is retained. In the Freestyle function, Keep format can be activated using the inserted abbreviation "(kf)".

- **Keep paragraph format** does not go quite as far than keep format, but it can still help preserve the formatting of the existing text. Format information is also stored temporarily in the text here, but only the paragraph formatting, which is much less. That is why this is also faster than the previous option and often sufficient for texts in which a lot of work is done with templates. If this is not selected, the add-in will still try to remember the paragraph formatting for relatively short texts, but this is less reliable because the output may not have the same paragraphs (even one additional paragraph mark will confuse the concept). If keep format is selected, this takes precedence.
- In addition to saving paragraph formatting, the Word add-in will also try to save footnotes, endnotes and dynamic fields (e.g. cross-references, date fields) in the text sent to the AI and insert them again afterwards (e.g. as a translated footnote). However, certain other information such as tables, images or comments will be lost. They are not "cached" and must therefore be saved beforehand. This also happens if "Keep paragraph format" is not selected, but it only happens when the existing text is replaced, but not where it is appended (can be overturned in Freestyle using "(sar)").

In the Freestyle function, keep paragraph format can be activated using the inserted abbreviation "(kpf)".

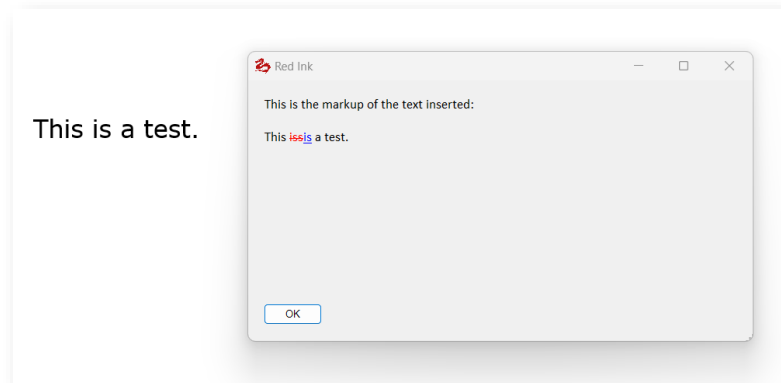
We recommend trying it first without "Keep paragraph format". In most cases, this is sufficient.

- With **maximum text for keeping format**, a value (number of characters, e.g. 10,000) can be set from which the add-in no longer takes into account the commands for keeping the format in order to avoid long waiting times. If the value is set to 0, this safety function is deactivated. This safety function also offers a degree of convenience, especially when using Freestyle (see para. 30 et seq. below), since experience shows that Red Ink is not frequently used for direct adjustments to larger texts, but rather for other functions that take less time. In such cases, it makes sense for the add-in to automatically disable the time-consuming

functions for storing and processing formatting information, especially when a text is only being queried or serves as the starting point for a query in which its original formatting is not important. In Freestyle, the trigger "(noformat)" or "(nf)" can be used for this purpose. If you want to switch off formatting functions temporarily or permanently, you can also set the value to 1.

- **Replace text** tells the add-in to insert the output of the AI, e.g. the translated text, in place of the selected text. This can be configured separately for translations and other functions. Replace is usually used for translations, and for corrections or other functions it depends on the user's preference (we use replace also for corrections).
- **Do markup** tells the add-in to create a markup for the selected text in functions where it makes sense (e.g. for corrections, but not for translations). If this function is activated, it will be visible in the context menu (as shown in para. 14 above). Creating markups is not a trivial matter from a technical point of view within Office. We therefore provide four different **markup methods** (specify the value 1, 2, 3 or 4):
  - If **Word** (value 1) is selected, then the Word-internal comparison function is used. This works by copying the selected text and the new text of the add-in into two temporary documents, and Word then creates a third temporary document from them. The content of this third document is then inserted into the main document with the markups. This is all done automatically, but it can be seen on the screen, which can be confusing. This cannot be technically suppressed and it can lead to disruptions when using third-party add-ins. The document management system "iManage" (which we use), for example, does not follow the Office defaults and blocks the automatic closing of temporary documents and asks the user whether the files should be saved (which is not the case); unfortunately, we have not yet been able to persuade the manufacturer to correct this incorrect behaviour of their add-in.
  - If **Diff** (value 2) is selected, the markup is created using a simple diff algorithm that compares the texts word for word. This doesn't have the shortcomings of the word comparison function, but it is less reliable and is too slow for long texts. Therefore, a maximum number of characters can be configured, after which the add-in asks whether the method should really be applied. In addition, if the output takes too long, it can be cancelled by pressing the "**Esc**" key.

- If you choose **Diffw** (value 3), the markup is created using the same diff algorithm that compares the texts word for word but unlike Diff (value 2), the comparison version is displayed in a window (W for "Windows"), which is much faster than the normal diff. The window remains open until it is closed with the OK button. This means that you can work on the new text at the same time. This is the default setting.



- Finally, we came up with the **Regex** technique (value 4), which works by having the AI compare the modified text with the original text first. It then writes a description of all changes (with some context if necessary), which is then implemented using a search-and-replace function (originally, we used a method known as Regex for "Regular Expressions"). How well this works depends heavily on the language model used. Again, a maximum number of characters can be configured, but this is typically higher than for the Diff method. This method is suitable for making selective adjustments to texts and can also cope with larger texts. However, the method requires a high-performance model and takes more time because after each query, a second query is made to prepare the compare.
- For users in Switzerland and others who do not want to use the "sharp S" ("ß"), which some language models provide for in German texts, the add-in can be configured to automatically replace it with a double S.

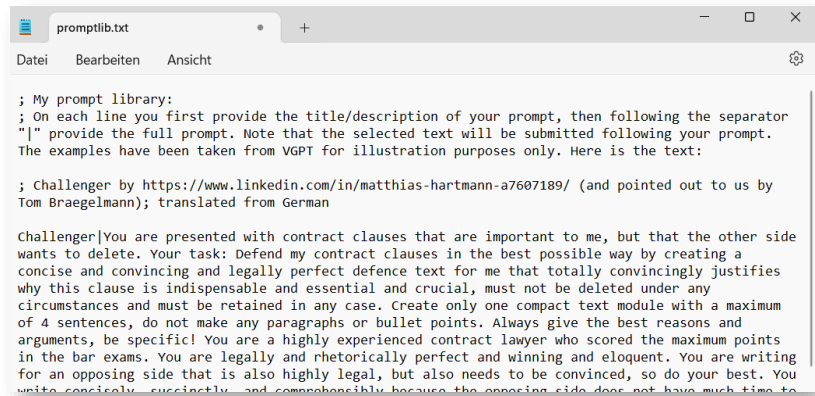
We recommend using Replace text in particular for translations, but also for the other functions. To start with, we also recommend not using Keep format and Keep paragraph format, and check-out whether the always enabled function for retaining paragraph formatting is sufficient. We do recommend to use Keep character formatting. If this is not sufficient, you can try adding Keep paragraph format. For long texts, we recommend a step-by-step approach (i.e., processing only a few paragraphs at a time); for this purpose, the "(iterate)" trigger can be used in Freestyle. Alternatively, the comment function of Freestyle

("Bubbles:") may be helpful to process larger texts (see para. 33 et seq. below).

26 Further configurations that can be made via the Settings menu include, in particular:

- The **temperature** specifies how creative the language model should be when providing answers. For tasks such as translations or corrections, we recommend a low value (e.g. 0.2) and for freer tasks such as finding arguments, a higher value (e.g. 0.8).
- The **timeout** value determines how long the system will wait for a response from the language model. Particularly complex tasks sometimes take a while to complete, and the system can give this time to the model. The timeout value is specified in milliseconds. If a timeout error occurs during normal use, you can try increasing the timeout value.
- The add-ins can be configured to permanently use **several language models** at the same time, a primary, a secondary and, as the case may be, as many additional models as desired (they can't be defined via Settings, see para. 267 below). The values for both are specified here. Switch Model can be used to switch between the two (if defined). Otherwise, the secondary model can be accessed directly via Freestyle. This can be used, for example, to store models with better problem-solving ability but that are slower, and only access them when needed.
- It is also possible to configure **two additional prompts**. The first is added to the language model for each predefined function (e.g. translate or abbreviate), except for the Transcriptor and Chat. This can be used to solve certain linguistic problems that arise regularly, e.g. if the language model does not adhere to the language or if certain terms should be spelled differently each time. If the second additional prompt is also filled in, the result of the query with this prompt is post-processed separately in each case. This can be more effective but takes much more time. We recommend that the second additional prompt should only be used in exceptional cases, and even the first only if the need arises.
- Two **default languages** can be specified, for which separate quick-dial buttons appear in the menu. These are English and German by default. The name of the language should be entered in English.
- You can specify the path and name of your own **prompt library**. This is available in Freestyle (see para. 30 et seq. below). It must be a text file and each prompt must be entered in a specific format (an abbreviation or title, then "|" without a space and then the prompt) on a separate line. Blank lines are not a problem, comments preceded by ";" are ignored. A sample prompt library

is provided in the installation package (get the most up-to-date version at <https://apps.vischer.com>). It can be changed in the add-ins (see para. 33 below for the layout, further details in para. 261 et seq. below).



- You can define the **keyboard shortcuts** that are used to access the functions in Word. To do this, enter the menu item (exactly as it appears in the context menu, e.g. "Correct"), then an "=" and then the key combination (e.g. "Ctrl-Alt-C") without a space. Multiple shortcuts should be separated by a ";". They also apply to the Excel add-in. The addition of "(Markup)" is not necessary. However, certain keyboard shortcuts are already in use and therefore do not work. The function also requires that an additional helper file is installed in Word (with VBA code, which is blocked in certain environments, see para. 228 et seq. below) and that the context menu is activated (which can also be configured). The keyboard shortcuts can also be edited directly in Word; they remain stored there. Technically, they work by launching a macro in the additionally installed file, which is loaded each time Word is started, and this in turn launches the code in the add-in. Unfortunately, there is no other way to do this, because Microsoft considers shortcuts to be an "outdated" technology that are not fully supported in the modern interface of Office products. However, since they can be very useful, we support them in this way.
- The parameter **Chat conversation memory** specifies how many characters of the previous dialogue the chatbot should remember. The text edited in Word, however, is not saved in this chat memory, unless the chatbot quotes such parts in its dialogue (which it is encouraged to do if it is supposed to remember something, e.g. when switching back and forth between different documents). The default value here is 50,000. If this value is too high, the language model will not be able to process some of the transmitted content, especially the user's document.
- It can also be specified where the local models required for **speech recognition** (and in the case of Whisper the additional

runtime libraries) are stored. More information about this is can be found in para. 70 et seq. below.

- 27 You can view or change further configuration values via **Expert configuration**. However, we do not recommend this. If specific configurations are required, this is easier and more reliable when the configuration file is edited manually (see below). It can, however, be updated from the add-ins. If the encryption of the API key or private key is activated, only the encrypted key is displayed there. When you close the expert configuration dialogue with OK, the (local) configuration file is updated or rewritten.
- 28 The changed configuration can be saved or added to a (local) configuration file by clicking **Save Configuration**. Otherwise, the changed configuration will be lost when you close Word.
- 29 The **Reset Optional Values** function is used to reset the respective add-in to the default values, whereby the values required for minimal operation (e.g. API key for the API) are not reset. The function can therefore be used without risk. Depending on the local configuration, an alternative function may appear instead of Reset Optional Values. This is used to "switch back" to a central configuration, i.e. to discard your own customisations and revert to the values that your organisation provides by default.
- 30 The **Install Helper** or **Remove Helper** buttons are used to install or remove the additional helper program; they enable the context menu and keyboard shortcuts in Word and Excel, and the API in Excel. If the helper is installed, it can be removed; if it is not, the install button appears. Clicking Install Helper downloads the latest helper file from the website <https://apps.vischer.com/> and stores it in the directory where Word (or Excel) stores and automatically loads VBA add-ins on startup. Installing or running such add-ins may be blocked by security features in the respective operating system; the function should only be used if internal guidelines permit the use of such VBA add-ins. Alternatively, manual installation is also possible (see para. 228 et seq. below). If the helper is to be removed, Red Ink attempts to deactivate it and delete the file; however, this is not always successful; in this case, the respective program must first be terminated and the specified file deleted manually.
- 31 In the **Settings** menu, buttons for checking for updates may also appear. Whether and how they work depends on whether the application was installed via the Internet (i.e. <https://apps.vischer.com/>) or from a local source.
- 32 When Settings is closed with OK, it takes 1-2 seconds for Red Ink to reconfigure itself.



### C. The Freestyle function within Word

33 The Freestyle function allows **free prompting** with a speech model and offers numerous other functions that can help with studying, capturing and editing documents. It is therefore a very powerful and versatile tool that requires a certain amount of practice to get the most out of it. **Example applications:**

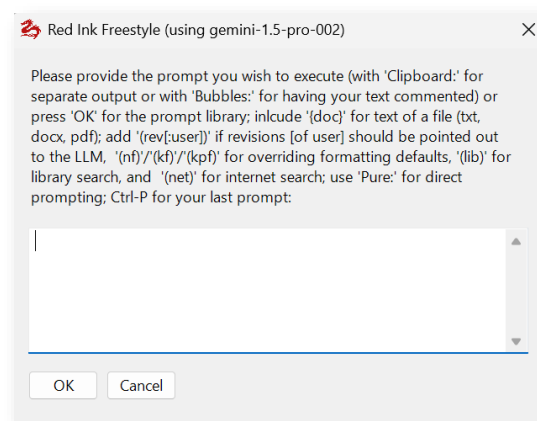
- Texts can be annotated with comments, e.g. indicating where improvements could be made;
- Texts can be queried, e.g. to find certain content that cannot be found using traditional search tools, or to determine which provisions a contract contains on a specific topic;
- Texts can be reworded where the predefined functions are insufficient, e.g. according to certain stylistic specifications (e.g. making a text more friendly or more specific or rephrasing a text to make it gender-neutral);
- Texts with specific content can be added, e.g. a contract can be supplemented with a specific clause that is given in keywords ("Write me a clause on the duration of the contract with a minimum term and monthly cancellation and take into account the existing regulations." – in which case the existing contract must be selected, otherwise the add-in will not access it) or based on a template from a clause library;
- Markups made by another person can be summarised and evaluated by the AI to provide a quicker overview;
- Extracts can be created from a text that can be used in another application;
- You can ask the AI to insert information from other documents into your own text depending on the situation, or a new text can be created based on information from another document (including PDF);
- Information can be extracted from a text and presented in a special form, e.g. in a table showing developments over time;
- The AI can formulate ideas for a text, e.g. how a contractual clause can be better defended in negotiations;
- The AI can critically assess a text, e.g. a legal document;
- The AI can compare the content of two texts.

In our experience, how well these examples work depends on the capabilities of the language model, the size and formatting of the text – and, of course, the prompt, which is either typed or taken from the prompt library.

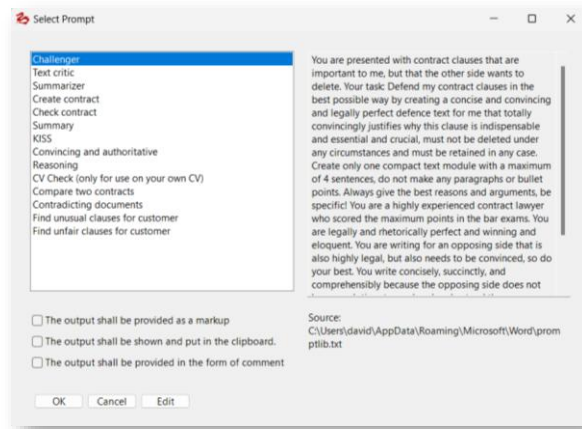
34 In principle, the function can be used to give the language model any command, **with and without selected text**, as is also possible in AI

chat programs like "ChatGPT" or "Copilot". But there are two important differences: Freestyle deliberately does not remember the previous interaction, i.e. it takes the text as it is and the current command; what has been discussed before does not influence the execution. The second important difference is that the AI's results can be processed directly in Word and applied to the existing text; there is no need to copy and paste. If you need a chat, the tool offers one through another command.

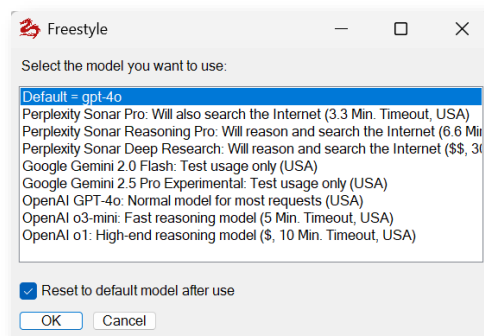
- 35 When Freestyle is called up, a window can be used to enter a **multi-line prompt** if required. The prompt is then transmitted to the language model together with the text to which it is applied (with some accompanying instructions in the background, which can, however, also be viewed and changed).



- 36 Red Ink remembers the last prompt entered in Freestyle (even if Word is closed in the meantime). It can then be inserted using **Ctrl-P**. Anyone who wants to run the prompt again unchanged with the current selection can also choose **Freestyle (redo)**.
- 37 If no prompt is entered and the process is not cancelled, **the prompt library** – if available – appears, and the desired prompt can be retrieved from it. It is suitable for complex prompts that are used repeatedly. We have stored some templates in it, including some from third parties. The prompt library can also be edited in Red Ink. We therefore recommend that you use your own local copy of the prompt library (even if it is possible to store it centrally). The prompt library is reloaded each time it is accessed. It can be used separately or together with Excel and Outlook. It can also be stored centrally in an organisation, but this carries the risk that an incautious user inadvertently changes it for everyone. Further details are in para. 261 et seq. below.



- 38 Freestyle is available in Word for both the **primary language model** and the **secondary language model**, if one has been configured. If additional models are configured (see parameter "AlternateModelPath" and para. 267), you can select in advance which of these models should be used (if the checkbox is unchecked, the newly selected model remains active as a secondary language model until Word is restarted and can be used, for example, in the chatbot):



The secondary and further models are accessed via **Freestyle (2nd)**.

- 39 Various output formats can be accessed using prefixes in the prompt (they are mentioned in the text of the prompt window as a reminder):

Please provide the prompt you wish to execute (with 'Clipboard:' for separate output or with 'Bubbles:' for having your text commented) or press 'OK' for the prompt library; include '{doc}' for text of a file (txt, docx, pdf); add '(rev[:user])' if revisions [of user] should be pointed out to the LLM, '(nf)'/'(kf)'/'(kpf)' for overriding formatting defaults, '(lib)' for library search, and '(net)' for internet search; use 'Pure:' for direct prompting; Ctrl-P for your last prompt:

- Freestyle can be asked to output the result as a markup for the selected text. This makes it easier to see what has changed. To do this, the text string "**Markup:**" must be added to the prompt. Alternatively, "**MarkupWord:**", "**MarkupDiff:**", "**MarkupDiffW:**" and "**MarkupRegex:**" can be used to apply a specific markup

method (see para. 25 above); otherwise the default set for the other functions will be used.

- If "**Replace:**" is used, the selected text is simply replaced by the output of the language model (without markup) (e.g. "Replace: Rephrase this sentence to make it sound more flattering."). If, however, "**Append:**" or "**Add:**" is used, the opposite happens: The output of the language model is inserted after the selected text, even if the default setting is different. Field, reference and formatting preservation will be turned off by default (can be overturned in Freestyle using "(sar)").
- If you don't want the AI output in the document, precede your command with the word "**Clipboard:**" (or "**Clip:**"). The output is displayed in a box at the end and can be edited there. The finished text (or the original text) can then be copied to the clipboard (without formatting). Clipboard and Markup cannot be combined, and formatting is not supported in the Clipboard variant (with the exception that in the box, places marked as bold by the AI are also displayed as such). However, it is possible to insert the text provided by the AI (i.e., without user edits), including formatting, into Word. A dedicated button is provided for this purpose. Clipboard is very useful when an answer is desired from the AI but should not be processed further in the text. However, further processing is still possible (via the clipboard, which is automatically operated).
- Instead of a window, the output can also be inserted into a new Word document. In this case, the prefix "**Newdoc:**" is to be used.
- If you want the output displayed so that you can continue working on the document, you should prefix your prompt with the word "**Pane:**" (or click the "**Transfer to Pane**" button when outputting via "Clipboard:"). The output is then transferred to a pane that opens to the right of the document. This can be enlarged and made smaller or even detached. The result can be edited in the pane itself. The pane has buttons to insert the selected text into the clipboard or to intelligently merge it with the selected text in the active document ("**Merge Selection**"). If Merge Selection is chosen, a window opens and a prompt is displayed, which can be modified, and which takes care of the merging. If you want to insert the original AI response with formatting into your document (or a new document), you can also choose this (the pane will then be closed) or the pane can simply be closed. The pane will remember the width and will reopen with the same width the next time.
- It is possible to output the AI's response neither in the form of a text in the document nor as its markup but rather using the

comment function (aka "bubbles"). In this case, the word "**Bubbles:**" must be added before the command (e.g. "Bubbles: Give me all sentences that I could correct and explain how"). The add-in will then transfer the selected text to the AI and display the answers in corresponding Word comments at the appropriate points in the text. This has the advantage that the comments can simply be deleted at the end. If the add-in cannot assign a response to the AI or otherwise evaluate it, it will output it at the end of the selected text. The comments can be recognised by the initials "RI: "; the current user name is deliberately used for the comments so that the comment written by the AI can be used immediately as your own comment if required. The reliability of this function depends on the performance of the language model; if it does not follow the instructions correctly, this function will not work well either. The comment function also requires that the text passages indicated by the AI are actually found in the document, which does not always work because invisible characters or formatting in the document can prevent a match; markups/revisions can also interfere with the mechanism. Despite countermeasures in the add-in, the commented sections may be shifted; they should therefore be checked in each case. If an assignment is not possible, the add-in will display in a separate window after completing the comments. The window will show what was delivered to the AI in the wrong format or could not be assigned to the existing text; unless you cancel the process, this text will be added in a final bubble at the end of the text. Bubbles can be applied to parts of the text or to the entire text (i.e. without selecting beforehand).

- Red Ink can also add new slides with AI-generated content to an existing PowerPoint file. For example, the AI can be asked to present a memorandum or a contract on a few slides. The prerequisite is an existing PowerPoint file in .pptx format, which also contains the slide templates or a slide with the desired design, so that Red Ink or the AI can use it as a reference. The information about the existing presentation (including its pre-existing text content) is passed to the AI, so that it can also be referenced in the instruction. The instruction must be preceded by the prefix "**Slides:**" (e.g., "Slides: Add further explanations about the contract to the existing presentation starting from slide 3, taking into account what is already in the presentation."). Red Ink will then ask for the PowerPoint file and adapt it (it must be closed for this, otherwise an error will occur). It may be necessary to make some adjustments or reset the slide layout if the AI did not apply the formatting correctly. When instructed accordingly, the AI will also incorporate graphic elements and icons ("Also add illustrative icons to the presentation where appropriate."). However, it cannot create images using "Slides:". Existing slides are also not

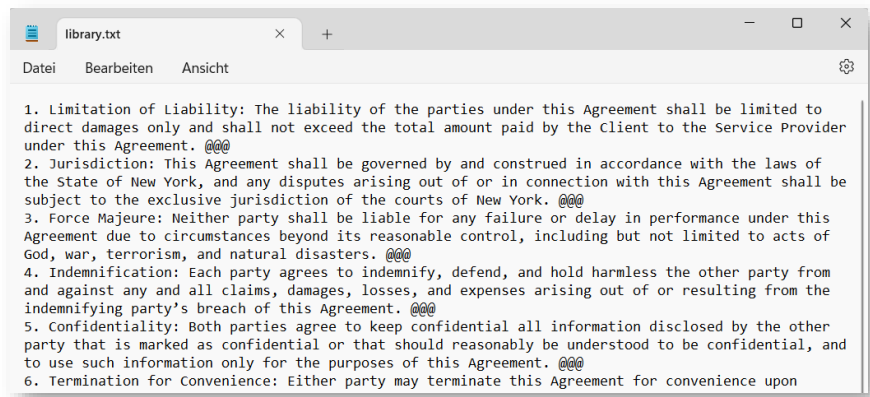
adapted (but the content of the inserted slides will be coordinated with them if requested). Such further adjustments must be made manually. Caution: The existing pptx file will be modified. It may therefore be necessary to create a backup copy beforehand. After generation, a spoken version can be created via "Create Audio" (para. 95), i.e. the speaker notes can be converted into an audio clip that is automatically inserted into the presentation.

- With "**Pure:**" Red Ink can be instructed to pass the entered instruction to the language model without additional instructions from Red Ink (except for the instruction to preserve existing formatting, if the corresponding option has been selected). This can be used to pass direct prompts to the AI (functionally as a system prompt, where a distinction is made). Normally, however, this is not needed. No further trigger codes are executed in this case. A marked text, however, is passed as a user prompt (embedded in a <TEXTTOPROCESS> tag).

40 The prompt itself may contain further functional trigger codes that trigger a function:

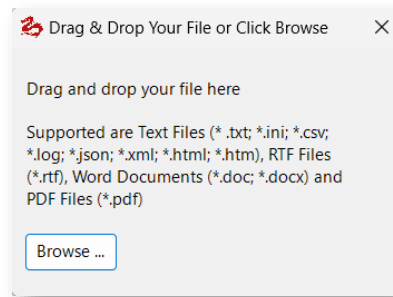
- In certain cases, it can be useful to retrieve additional **information from the internet** to supplement a text or answer a question. If a search engine is configured, the add-in in Word can be asked before executing the command to perform an internet search for information that is missing but is necessary for the execution of the command and then to use the information from the first hits for the command as well. To do this, "**(net)**" must be appended to the command. Whether this functionality is available is displayed in the help-text of the prompt box that appears when calling the Freestyle function (this can be configured, as can the search engine). If it is used, Red Ink will, if configured to do so (parameter "ISearch\_Approve", see below), display the search commands to be used to perform the search and prompt for confirmation. This ensures that no confidential information is passed on to the search engine. This function only works where the search hits actually display the information at the address found; for more complex websites this is not necessarily the case, because the found address itself refers to numerous other nested sub-pages. The add-in can be configured to determine how deeply it enters a website and for how long.
- The add-in can be asked to primarily access a suitable piece of information from a library instead of its own knowledge to execute a command. This is referred to in technical jargon as **Retrieval Augmented Generation**, or **RAG** for short. An example of its use is a text database of contract clauses. To do this, "**(lib)**" must be added to the command (cannot be combined with "(net)"), provided that the library search has been configured and the library is retrievable. In this case, Red Ink will first search the

library according to the configured settings and the entered Freestyle command, and then apply the command with the found content to the selected text. How it does this must be specified using the corresponding prompt. If markup is not used, the output is appended to the selected text. The library is a simple file in TXT or Word format (Word is slightly slower) with the corresponding entries separated by a character that is to be taken into account in the prompt (e.g. "@@@"). The "(lib)" function is available with both the primary language model and any secondary model. However, the model selection only has an effect on the second command, i.e. the application of the content extracted from the library, not the extraction of the term from the library. The prompts can be configured individually. This also allows for a possible special structure or other separators in the library file:



- If the placeholder "{doc}" is inserted in the prompt, Freestyle will insert the text of an **external text document** there. This can be useful, for example, to apply the content of that document to an existing text or to have it compared with it by the AI, which would not be possible with any other markup function. Supported file formats include plain text formats (such as ".txt", ".html" or ".csv"), Word documents (.docx" and ".doc") and PDF (whereby only the text portions of PDFs are processed, i.e. those that can be searched for; if additional text recognition is required because the text is only available as an image, e.g. in the case of scanned PDFs, this must be done beforehand using a separate PDF programme or you can use the trigger "(file)" or "(clip)", if your model supports this and has been configured accordingly, see right hereafter). This function is also available as a Word helper (para. 129 et seq. below).





If you include "{doc}" in the prompt, it is a good idea to tag it so that the language model can better distinguish it from the instruction, for example, with "Here is the external text: <TEXT>{doc}</TEXT>".

- The trigger "**(file)**" also allows external files to be read in, but in a different way. Here, the add-in is instructed to transmit the content of the file directly to the language model. However, this only works with file formats that the language model supports and if the language model has been configured accordingly ("APICall\_Object"). Modern language models support the common formats of images, audio files and videos as well as PDF documents. If the trigger is set, a window opens into which the respective file can be dragged. An sample prompt for this trigger would be *"What can be seen in this image? (file)"* or – if the model also supports image generation and has been configured accordingly – *"Color the object in this image blue and create a new image of it. (file)"*. Other applications are transcribing audio files and analyzing or converting content from PDF documents. However, the add-in does not check whether the file format is supported or whether the file is too large. The larger the file, the longer the response of the language model takes.
- The same also works with the contents of the clipboard if, instead of "(file)", the trigger "**(clip)**" is inserted. This is particularly practical when, while working on a document, text from another document that cannot simply be copied needs to be inserted. It is only necessary to take a screenshot of the relevant passage (Shift-Window-S), and then Freestyle can be instructed with a command such as "Extract for me the text from the image (clip)" to extract the text and deliver it as a response. To make life easier for you, a corresponding command is already preprogrammed as a Word-Helper-Function ("**Clipboard to text**") and can also be used by entering the short command "**insertclip**" or "**iclip**" within Freestyle.
- The trigger "**(rev)**" instructs the add-in to code all **markups in the selected text as such** before the text is passed to the language model. This makes it possible to ask the language model questions about the markups, e.g. "Summarise the changes

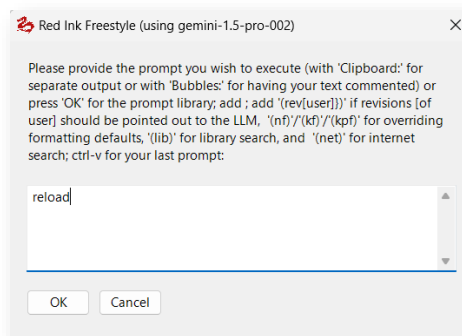


made to the contract in terms of their meaning.". If you only want to have markups from a specific author coded, you can simply enter the name after a colon, exactly as it appears in the text (e.g. "(rev:VISCHER)"). You will have to tell the AI in sufficiently clear terms what to do with the markups (deletions or insertions/additions) because it will only see them, but not by default know what to do with them.

- With "**(noformat)**" or "**(nf)**" in the prompt, the add-in can be instructed not to store any formatting information about the original text or to remember it, which significantly increases the processing speed for longer texts. As explained above, a character limit can be configured, but not for Freestyle in isolation (see "Maximum text..." in para. 25 above). Conversely, "**(keepformat)**" or "**(kf)**" activates the Keep Format function for the current prompt, while "**(keepparaformat)**" or "**(kpf)**" activates the second function described in para. 25 for preserving formatting. So, if you want to keep the format of an existing text (as far as possible) in Freestyle, you must request this with these codes in the prompt; the above configuration settings for preserving formatting deliberately do not apply to Freestyle. If "**(sar)**" (*same as replace*) is used, the special function for preserving dynamic fields, references, and character formatting is activated, as if it were a matter of replacing the text.
- With the addition of "**(all)**", the add-in can be instructed to use the text of the entire document for the query, even if not all of it has been selected.
- With the addition of "**(iterate)**", the add-in can be instructed to process the command not in one piece, but step by step. If this is selected, the add-in asks for the number of paragraphs to be processed per step. So, if the value 3 is selected, for example, the add-in will not execute the desired command on the entire selected text, but first only on the first three paragraphs, then the next three paragraphs, and so on. This does not work with Clipboard and Pane either. Bubbles, however, does work with it. With regard to markups, only the Diff and Regex methods are supported (the user is automatically prompted if they have not been chosen). Iterate is especially useful for large texts, which can be divided into smaller blocks for better output in this way. Predefined commands such as Translate, Shorten or Improve do not offer Iterate, but via Freestyle this can be emulated to a certain extent with the corresponding prompt ("Shorten each paragraph for me as much as possible (iterate)"). The most reliable value is 1.

41 These additions also work when they are specified in the prompts of the prompt library. There, the special output formats (prefixes "Clipboard", "Markup" and "Bubbles") can also be activated via a checkbox.

- 42 If the language model used is multimodal and supports **image generation**, Red Ink automatically saves an image delivered by the LLM in a (sequentially numbered) file on the desktop and inserts a note in the text, which describes where the file has been saved (provided that image encoding and format are supported by Red Ink). This output can also be redirected, e.g., to the pane.
- 43 To repeat the last Freestyle command with the last selected model (but with the current selection), execute **Freestyle (redo)**.
- 44 Insofar as this is configured for a model using the "TokenCount" or "TokenCount\_2" parameter, it is possible to log the tokens used for a freestyle request and have them multiplied by a currency amount. This can be used to pass on corresponding expenses to a client or a cost center. The token costs may be low, but if certain tasks can be completed more quickly through the use of AI, it may make sense in some cases to charge for the value added, to a certain extent based on the principle that the AI's "thinking" effort also has value. This is particularly relevant for Freestyle, e.g., when the AI is asked to create a memorandum or conduct an analysis. The output appears in a text file named "**redink-cost.txt**" on the user's desktop. Further entries are appended each time.
- 45 A number of "command line" commands can also be executed via Freestyle; some of them are for administrators only. They are simply entered instead of the prompt and confirmed with OK:



- "**model**" outputs the primary model currently in use and the model's current timeout value.
- "**terms**" outputs any preconfigured usage restrictions or permissions in the INI file. They are also displayed when you move the mouse over the Red Ink logo in the menu bar.
- "**version**" provides information about the current version of the add-in; this information also appears when you move the mouse over the Red Ink logo in the menu bar. The "version" command also shows the expiry date of the current licence for Red Ink (this information can also be accessed via Settings and then "About Red Ink"). This menu item also provides further information, e.g. about the third-party libraries used.

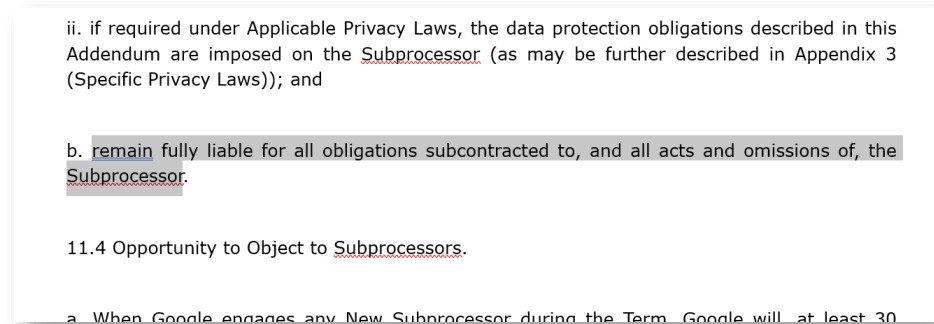
- **"switch"** can temporarily swap the primary and secondary AI models; this is also possible via Settings.
- **"clearlastprompt"** clears the cache of the last executed free-style prompt.
- **"cleanmenu"** removes any existing Red Ink context menus and rebuilds them if they are enabled.
- **"reload"** ensures that the add-in reloads the configuration file (e.g. because something has been manually adjusted in the meantime; this also happens automatically after saving the configuration in Settings).
- **"reset"** resets the local configuration file so that only the minimum required entries are present and the other values are set to the default values.
- **"settings"** calls up the function for manually adjusting the settings.
- **"encode"** can be used to encrypt API keys and private keys so that they do not have to be stored in plain text in the configuration file. To do this, the key in plain text must be marked in Word (see para. 280 et seq. below).
- **"decode"** can decode them if the keyword is known (see para. 280 et seq. below).
- **"inipath"** allows the directory for a central configuration file to be written in the registry (see para. 257 et seq. below).
- **"codebasis"** allows you to write the keyword in the registry if it is not hard-coded in the programme code (see para. 280 et seq. below).
- **"domain"** shows the current domain in which the add-in is running and whether and to which domains it is restricted, if this should be programmed in as a security function (see para. 280 et seq. below).
- **"speech"** starts the Transcriber (see para. 70 et seq. below).
- **"voices2"** opens the window for selecting two voices for using the Google Text-to-Speech function, **"voices"** for selecting one voice; this function is normally not needed, because the selection is opened automatically for both the podcast and audiobook functions; these commands can be helpful if only voices are to be selected. For more see para. 92 et seq. below.
- **"createpodcast"** starts the function for creating podcasts (see para. 92 et seq. below).
- **"read"** starts the function for creating audiobooks, i.e. the selected text is read aloud (see para. 92 et seq. below).

- **"readlocal"** will have the integrated speech-to-text function read the selected text (or abort an ongoing output); no data is sent to the Internet (unlike when using Word's natural language read aloud or Google text-to-speech function), but the voice does not sound natural.
- **"voiceslocal"** allows you to select the voice to be used for "read".
- **"anonymize"** executes the anonymization function (in mode 3 or 4), otherwise optionally used when calling language models, without calling a language model on the selected text. This allows testing the anonymization. This can also be accessed via the "Analyze" menu item.
- **"generateresponsetemplate"** or **"generateresponsekey"** is used for the automated creation of the templates that are needed in the "Special Service" configuration file for processing the JSON strings returned by the respective service (see para. 275 et seq. and Annex 2). To use the function, first copy an exemplary JSON string in the typical response structure of the relevant service into a Word document, and then draft a description of what the template should be able to do or how the output should look based on the respective fields and values. Both are then selected and the command to be executed via Freestyle. All will then be passed to the current LLM for generating the template, along with the necessary information (i.e., the program code that processes the templates is also passed to the LLM).
- **"insertclipboard"**, **"insertclip"**, **"iclip"** or **"clipboard"** executes the command that passes the clipboard contents to the LLM and asks it to extract the text from it (for images with text or audio or video recordings) or to describe the content in text (for images). The function is also available as a Word Helper. The LLM must be configured for this type of function (processing binary objects).
- **"redinktest"** will read the text that is contained in the file "redinktest.txt" on the user's desktop and show it in a Window. It can be used to test the rendering of Markdown-formatted text. The Window uses RTF. If the content is inserted into the document, it will be rendered separately.

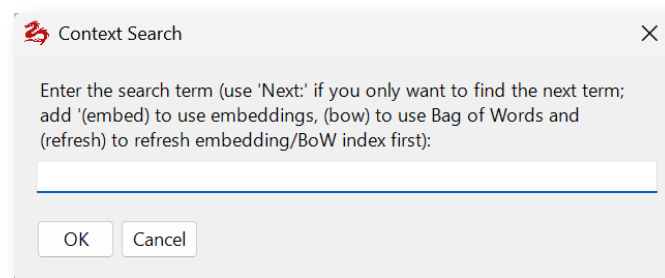
#### **D. Search by topic: Context Search**

46 The Context Search function can be used to ask the add-in to search and highlight for specific topics in the current text. Unlike Word's built-in text search, this function also finds places that do not exactly match the entered search terms but cover the same topic. For example, the Context Search keyword "Liability" in a contract can also find text passages in which the word does not appear in exactly this form, but something that has the same meaning (e.g. the sentence "... remain fully liable for all obligations ..."). The add-in specifies the hit by se-

lecting it: When entering Context Search terms, it is not necessary to enter a complete prompt. It is sufficient to enter the terms in context.



- 47 When entering Context Search terms, it is not necessary to enter a complete prompt. It is sufficient to enter the terms in context. However, it is also possible to narrow the search ("liability but not audit" will not find any audit clauses, although these may be related to liability).



- 48 This context search is performed by the configured primary language model, i.e., for longer texts, this can take some time. Alternatively, Red Ink also offers a vector search and a so-called bag-of-words search:

- With **vector search**, the current document is first "vectorized," i.e., the text is divided into so-called chunks (e.g., groups of two sentences), which are then stored in memory in a multidimensional data space (the "embedding space") at a virtual coordinate corresponding to their meaning. This allows the user to search in their own words for sentences that say the same, but possibly with different words. LLMs use this technique as well. If context search is used for this, it performs this vectorization on the current text once (however, it must be repeated as soon as the text is changed in any way). A prerequisite for using the vector search is also the installation of a suitable model for determining the meanings of sentences (more on this below). If vector search is available, it can be activated with the trigger "**(embed)**" and is performed using this technique. This does not require an LLM.
- The **bag-of-words search** is a simple method in which a text is first divided into small elements (words, tokens), which are recorded in a list. Then, it is counted how often each of these words

occurs in the document, without considering the order or context. This allows for a quick keyword search across the entire text. However, it only finds keywords that actually occur as such, so strictly speaking, it is not a context search. It can be activated with the trigger "**(bow)**".

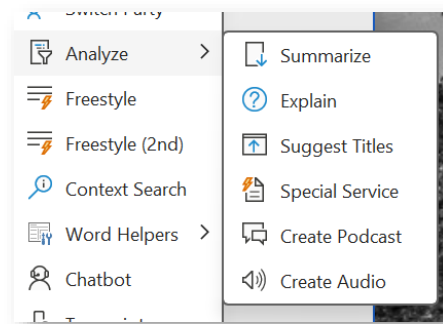
With both search methods, the trigger "**(refresh)**" can be used to prompt the system to regenerate the index (e.g., after changes including moves within the text).

- 49 For efficiency reasons, the entire document is searched at once by default. All hits are marked with a word comment. For vector search and BoW search, the relevance of the hit is also indicated; there, parameters such as how many hits should be displayed and what minimum relevance they should have can also be specified (if none are found, the relevance is expanded). For vector and bag-of-words searches, you can specify how large the chunks should be (e.g., two sentences each, with one sentence of overlap).
- 50 If you only want to display the next hit, prefix the search query with "**Next:**", which, however, only really makes sense with the normal context search using an LLM.
- 51 When the Context Search window is opened, the last search query appears automatically.
- 52 Incidentally the Chatbot Inky (para. 130 et seq. below) can also perform such contextual searches. They are programmed slightly differently and can therefore lead to different results.
- 53 To activate the vector search, a **suitable model** with a tokenizer file must first be installed. The starting point is the path specified in the "LocalModelPath" parameter (e.g., "D:\ModelsInUse"). There, the sub-directory "embed" must be created, and within it, the model must be saved as "model.onnx" (in the official source, the file will typically have a different name but also end with ".onnx") and the tokenizer file as "vocab.txt".
- 54 An open-source model is provided on <https://apps.vischer.com/> as a separate download ("all-MiniLM-L6-v2-onnx"); it can alternatively also be downloaded from HuggingFace (only the two mentioned files are required). This standard model has 384 dimensions (Float32 Array), a maximum sequence length of 256 tokens, input tensors of type Int64, shape [1, seqLen] "input\_ids", "attention\_mask", and "token\_type\_ids", and the output tensor of type Float32 (384), with an ONNX opset of  $\geq 11$ . It uses a Wordpiece tokenizer. It supports various languages, but in our experience, it is only of limited suitability for specialized texts. We will continue to look for and test further models here.

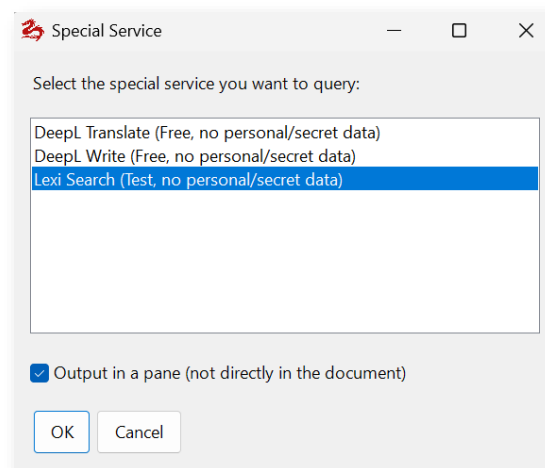
## E. Access to other data sources and services: Special Services

55 Red Ink can also be used in Word to access online services from third parties or within your own company, for example, to retrieve legal information (e.g., from services such Lexi Search), to use specialized AI services (such as the translation service from DeepL), or to retrieve information from an internal knowledge system (e.g., from a server with a vector database containing all internal knowledge documents).

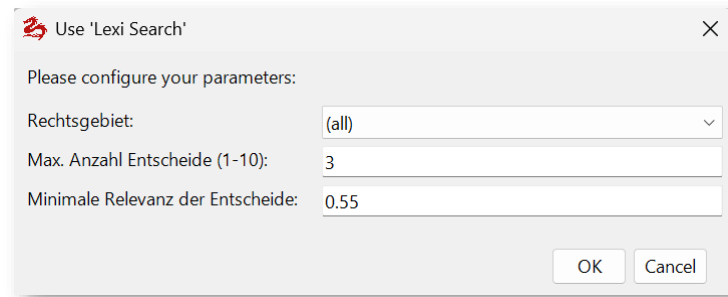
56 Access to this is via the **Analyze** command and there via **Special Service**:



57 It only appears when such "special services" are configured, and they can only be used if a text passage has been selected beforehand to be transmitted to this service. If this is the case, the list of available services is displayed:

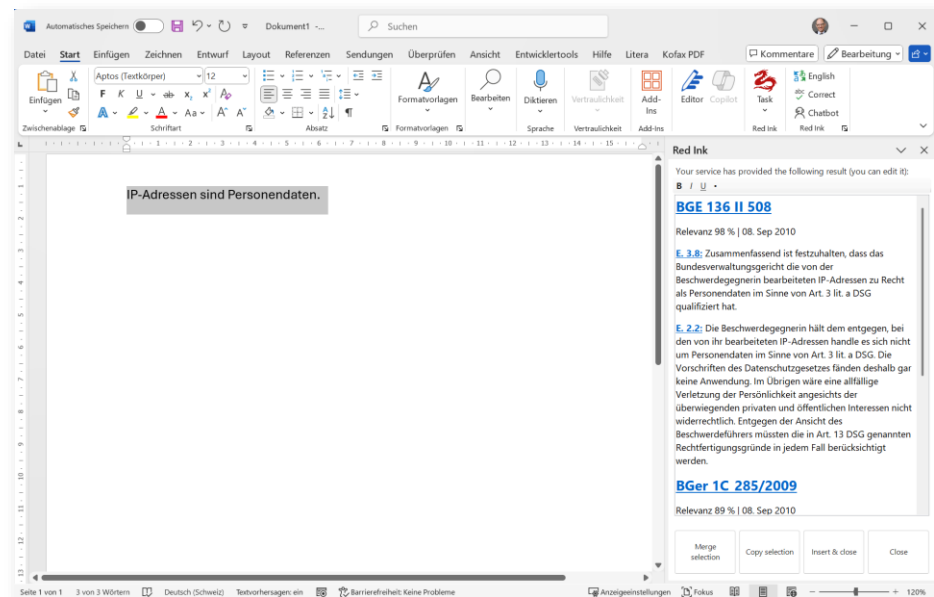


58 The desired service is selected (here, e.g., a service that provides Swiss court decisions that match the statement in the selected text passage [www.lexisearch.ch]), and any parameters for the query can be entered. These are different for each service. Here is the example of Lexi Search:



59 If configured accordingly, a **Query Assistant** can support the query. If selected (it appears as the last item in the query above), the selected text is sent to the primary LLM beforehand with the request to extract the necessary search terms for the query (the prompt used for this can be defined per special service in its configuration). The determined search terms are displayed in a window before use and can still be changed. This is also helpful if you want to ensure that no confidential data is sent to the special service.

60 The query is then executed and the result is either inserted directly into the document or a new pane is opened to the right of the document where the text can be read and edited and further used. The pane can remain open as long as desired. The content, however, is overwritten with the next request.



61 At the bottom of the pane, there are several buttons: **Merge Selection** allows you to merge the selected text with the help of AI into the text selected in the document (e.g., in the above example, the court decision with citation could be inserted). If a text passage is selected both in the document and in the pane, and the button is pressed, a window opens in which the prompt for merging can be entered. A separate prompt can be configured for each service. Furthermore, there is



a button to insert the text selected in the pane into the clipboard and a button with which the original text delivered by the service (i.e., without subsequent edits) can be inserted into the document with the original formatting.

- 62 The **configuration** is done via a separate configuration file, which is structured similarly to the configuration file for models and whose path is stored in the configuration file ("SpecialServicePath"). The specifications are the same as for an LLM and the file is structured the same way as the configuration file for alternative models (see para. 267 et seq.). The only difference is that here, for each service, up to four additional parameters and an individual MergePrompt can be recorded, which are queried as shown above. The configuration entry for Lexi Search looks like this, for example:

```
[Lexi Search (keine Klienten- oder Personendaten)]
APIKey = xxxxxxxx
APIKeyPrefix =
APIKeyEncrypted = False
Model = Lexi Search
Endpoint = https://www.lexisearch.ch/api/v1/search
HeaderA = Authorization
HeaderB = Bearer {apikey}
Response = response
APICall = {"search": {"query": "{promptuser}", "filters": {"decision__law_field": "{parameter1}",
"top_k": {parameter2}, "min_score": {parameter3}}}, "locale": "de"}
Timeout = 200000
Parameter1 = Rechtsgebiet; String; (all); (all), Zivilrecht<civil>, Strafrecht<criminal>,
Öffentliches Recht<public>
Parameter2 = Max. Anzahl Entscheide (1-25); Integer; 5; 1-25
Parameter3 = Minimale Relevanz der Entscheide; Double; 0.55
MergePrompt = Integriere den selektierten Auszug aus einem Bundesgerichtsentscheid als Zitat so
in meinen Text, dass es diesem als Beleg mit Quellenangabe dient, wie dies in einer juristischen
Fachschrift passen würde
```

- 63 Each parameter entry has the same structure and is separated by a semicolon:

- Description of the parameter (displayed to the user);
- Type of the parameter (String, Boolean, Integer, Double);
- Default value;
- Optional: Values from a drop-down menu, separated by commas (with the parameter value to be used in <...>) and in the case of a numerical value, the permitted range (the parameter will then be automatically clamped).

- 64 The recorded value of the parameter is then inserted in the APICall string at the relevant position of the placeholder (e.g. "{parameter2}"). This way, the query command for the service can be controlled individually. The interface's response is then inserted either in the document or in the pane (in the case of formatting in Markdown format, also with corresponding formatting).

- 65 If the parameter value to be used itself contains ";", ",", "<" or ">", then these characters must be "escaped", i.e. they must be provided with two backslashes in this case, e.g. "\\," or "\\<". In this way, more

complex JSON expressions can also be inserted in the angle brackets (here using the example of LexiSearch):

```
timeout = 200000
Parameter1 = Rechtsgebiet; String; Alle; Alle<>, Zivilrecht<civil>, Strafrecht<criminal>, Öffentliches
Recht<public>
Parameter2 = Gerichte; String; Bundesgericht; Bundesgericht<["CH_BGE"\\, "CH_BGer"]>, Kanton
Zürich<["ZH_OG"\\, "ZH_HG"\\, "ZH_KG"]>, Alle <["CH_BGE"\\, "CH_BGer"\\, "ZH_OG"\\, "ZH_HG"\\, "ZH_KG"]>
Parameter3 = Max. Anzahl Entscheide (1-25); Integer; 5; 1-25
```

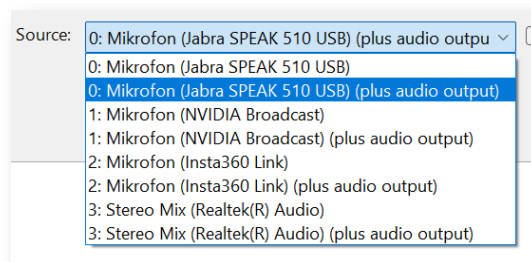
- 66 An empty pointed bracket can also be inserted (as shown in the example); in this case, an empty string will be inserted at the relevant position. This is also the case if the text is "(no selection)", "(keine Auswahl)" or "---". If, on the other hand, the API requires that all options be listed when all options are selected, these must be stored accordingly (see the example above for Parameter2 the parameter value for "Alle").
- 67 The MergePrompt is the prompt displayed during Merge Selection, which can be edited by the user. If it is missing, the default MergePrompt is used, which is either stored in the configuration file "red-ink.ini" ("SP\_MergePrompt") or otherwise predefined in the add-in. In addition to the MergePrompt, the prompt in the parameter "SP\_Add\_MergePrompt" is passed to the LLM on top, with the necessary information about the text to be inserted and the text of the main document.
- 68 The QueryPrompt is the prompt with which the Query Assistant extracts the search terms from the selected text. It is sent to the LLM without any further additions. It should indicate that the source text, from which the search terms are to be determined, will subsequently be passed between the tags <TEXTTOPROCESS> and </TEXTTOPROCESS>. Example: *"MergePrompt = Extract from the TEXTTOPROCESS (provided to you between corresponding tags) precise and language-preserving search terms for the purpose of finding relevant court decisions addressing the same legal topic in a database of court decisions. Provide only the bare-bones search terms, separated by space, and nothing else, no wildcards, no quotes, no boolean operators, no comments, no commas."*
- 69 The "Response" parameter can contain complex evaluations, which can also be used to determine how the result appears in the pane (in the example above, the service itself provides a suitably Markdown-formatted text, but in many other cases this is not the case). See para. 275 et seq.

## F. Transcriptor

- 70 Red Ink has built a transcription feature into Word. It currently supports the open-source solution **Vosk** and its various speech-to-text models for converting speech to text, the OpenAI solution "**Whisper**" and the Cloud-based STT-models from **Google** (V1, via Vertex) (see below). Once the models are configured, the Transcriptor function is available in the tile menu. This can be used, for example, to transcribe

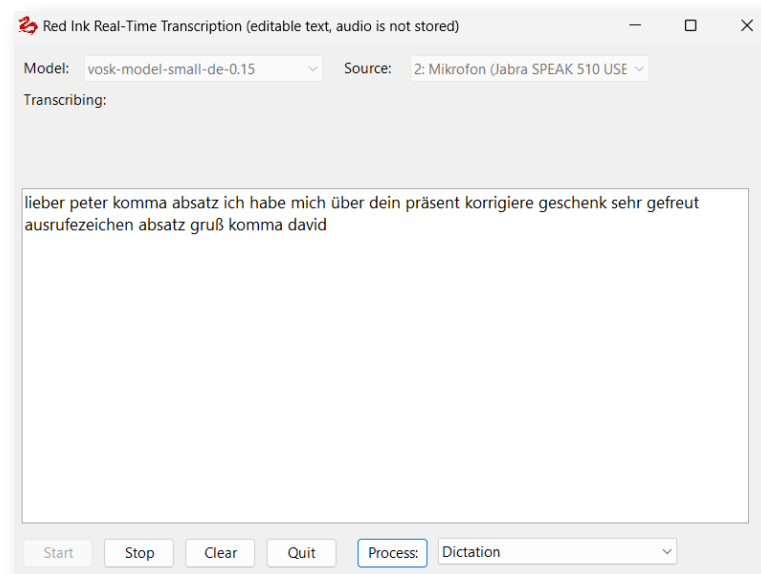
meetings (including online meetings), to dictate texts or to transcribe videos or lectures (e.g. to create a summary).

- 71 The live transcription function was deliberately programmed so that the **audio signal is not stored**, but continuously forwarded to speech recognition. This means that the user has no access to the audio signal via Red Ink and cannot listen to it again (not even via a temporary file – the processing takes place exclusively in the main memory, as is also the case with the video conferencing solution). This is important because, depending on the legal system, it may determine whether the user of the live transcription needs to obtain the consent of the other participants in the conversation. For example, in Switzerland, according to Art. 179ter of the Swiss Criminal Code, participants in a non-public conversation are only prohibited from recording it without the consent of the others.
- 72 When the Transcriptor function is selected, a window opens in which the model (i.e. the language) must first be selected at the top, followed by the **audio source**. The microphone can be selected here, with or without audio output, respectively. If audio output is selected, the Transcriptor will process the signals transmitted from the main output device configured in Windows (i.e. the device used as the normal loudspeaker), in addition to the microphone input; with the button "Dev" (for Device) the audio output device can be changed from the default main audio device to a different one (for example the one used for videoconferences). For transcribing video conferences, either choose the microphone with Audio-Output ("(plus audio output)", see image below) or "Stereomix" or "Stereo-Mix" should be selected, as this combines the microphone's own voice and the sound of the other participants. However, not all computer systems support this so there is no guarantee that the Transcriptor can actually hear all the inputs. This may be because a video conferencing software "books" a microphone exclusively for itself, for example, and the Transcriptor can no longer connect. The ability to exclusively reserve audio devices can also be deactivated in Windows. What can also work in practice is that a video conference is heard over loudspeakers, the speaker uses one microphone (e.g. in the loudspeaker), while another microphone records the whole dialogue and it is transcribed from there.



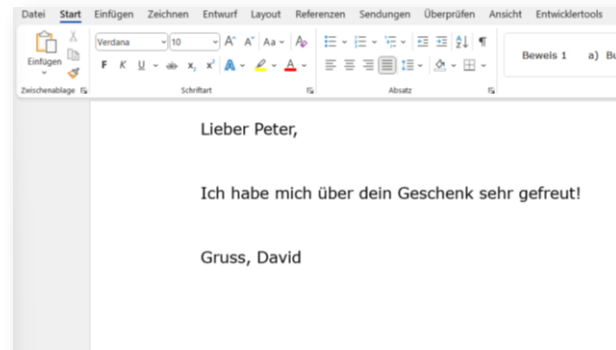
- 73 Once the configuration is complete, transcription can be controlled with "**Start**" and "**Stop**". Once started, the text appears as soon as the

audio signal has been evaluated by the speech recognition. Here, Vosk and Whisper behave differently. Vosk displays how it constructs the sentences as it hears them. This can be seen in the upper area under "Transcribing". Once it has heard enough, the recognized text appears in the window below and can be freely edited there even during transcription. Whisper only delivers its text once it has finished transcribing a section and also takes longer (depending on the model and computing power of the device), i.e., it is necessary to wait longer until something appears. It can therefore happen that the transcription continues for a while after the conversation is already over. In this case, it is simply best to let it continue. Pressing "Stop" prematurely can abort the transcription.

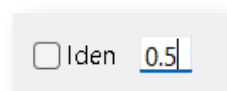


- 74 With Vosk, the model tuned to the **language must be selected**. Because Whisper can work with multilingual models, this is not necessary there. However, before transcription, in this case, you are additionally asked which language will be spoken. If this is not known or Whisper should figure it out itself, enter "auto", otherwise the two-letter ISO code for the language in question, i.e. "en" for English, "de" for German or "fr" for French. If the language is specified, the result of the speech recognition is generally better. Swiss German is not an official language, but it is supported by the larger Whisper models (Medium, Large) under "de". When using Google, the desired language must be selected from a list.
- 75 If you want to **transfer the transcription to the current document in Word**, you can use the "Process" button. In the selection on the right, you can choose how the AI should edit and clean up the text, i.e. whether it should edit it as a dictation, as in the example above, or whether a summary, minutes or a to-do list should be created. The processing options are based on prompts that the user can define in a separate prompt library (which is structured exactly the same as the

normal prompt library). Some examples are provided. If dictation is used, the following text appears in Word (i.e. spoken commands are also taken into account):



- 76 The Process function processes the text selected in the transcription window or, if nothing is selected, the entire text in the window. The Process function can be used during ongoing transcription. If you want to use text without the Process function, you can select it and copy it to the clipboard and then paste it at the destination.
- 77 Incidentally, it is possible to copy transcripts from other programs into the window to have them processed by the AI. If an **existing recording needs to be transcribed**, this is also possible. The "Load" command is intended for this purpose: A window opens into which the file can be dragged. Then the transcription starts immediately with the selected model. Since more time is available here, larger models can also be used that require more computing power. When using Google, you can choose whether the data from the file is sent to the AI bit by bit (in chunks) or continuously like a live conversation.
- 78 Vosk's and Google's speech recognition at least theoretically also supports **the recognition of different speakers** (Whisper does not), also referred to as speaker diarization. It recognises them by differences in their voices, but of course it does not know who is who. If you want to use this function, you have to install an additional model and then activate the identification of speakers using the checkbox "Iden" in the upper right corner:



- 79 When activated for Vosk, the speaker's number appears in front of each text. The value after the checkbox (from 0.5 to 2.5) can be used to indicate how much the speakers differ so that the system can tell them apart. If the value is too low, different speakers are grouped together; if the value is too high, the system might consider the same speaker to be different people:

0.5 - 0.7: Very forgiving – even slight differences in voice are treated as the same speaker (useful for noisy environments);

0.7 - 1.2: Balanced – a good range for speaker differentiation under normal conditions (suitable for meetings);

1.2 - 1.8: Strict – only texts resulting from very similar vocal patterns are grouped as the same speaker (higher values are not recommended).

In our experience, the quality of speaker recognition is not good, especially in video conferences. Often it is more practical to simply switch it off and let a protocol be created without attribution to individuals.

- 80 With Google, speaker recognition also works in principle, but it runs a little differently. In the box next to Iden, the maximum number of speakers that will occur must be specified (e.g., "3" if three speakers are expected). Google's speech recognition works in such a way that Google only outputs the final text only once the speakers pause briefly. In addition, Google adjusts the final text as needed. So users should not make edits within the results window during transcription with Google with speaker recognition. Also, the transcription should only be stopped once the speakers have stopped talking, otherwise the last part of the already recognized text may be lost. However, we haven't had good experiences with Google's speaker recognition so far, and normally use the transcriber without it (the "Process" function can still assign the text to the individual speakers relatively well, if they are named).
- 81 If a Whisper model is selected, **automatic translation to English** is possible instead of speaker identification, i.e., a kind of simultaneous interpreting. Where the switch for the identification of speakers appears for the Vosk models, the switch for Whisper models changes to "Trans" for "Translation". Furthermore, in the input field to the right of it, you can specify for Whisper how sensitive the speech recognition should be in order to distinguish speech from background noise. A medium value is 0.6, and in noisier environments, a higher value is recommended (0.7-1.0).
- 82 Punctuation marks and upper and lower case are not yet supported with all models. This will follow, along with support for other speech-to-text systems (in addition to Vosk), for example to process dialect. However, upper and lower case and punctuation are not a hindrance, as they are compensated for by the process function.
- 83 Red Ink tries to prevent the computer from going to sleep while the transcription is running. This setting is reversed after the transcription is finished.
- 84 **Installation:** If you want to use the Transcripator, you first have to download a suitable Vosk or Whisper model and set the configuration file of Red Ink accordingly (for the configuration file, see para. 246 et

seq. below). Please note that transcription does not work on all devices, which can have various reasons (e.g. lack of power, lack of inputs). It will also work to varying degrees of success.

- 85 **Vosk models** can be obtained free of charge as a ZIP file from the website <https://alphacephei.com/vosk/models>. The contents of the ZIP file (it is a directory with several subdirectories) are placed in a local directory and the path to this local directory is stored in the Red Ink configuration file using the "SpeechModelPath" parameter. This then resembles the following:

```

vosk-model-small-de-0.15
vosk-model-small-en-us-0.15
promptlib-transcript.txt

```

- 86 **Whisper models** are also available free of charge. For the Transcriptor, you need to get the models that have been ported to C++ (a programming language) for improved performance and efficiency, particularly for devices with limited resources. You can get them at <https://huggingface.co/ggerganov/whisper.cpp/tree/main>. The file names start with "ggml" (the "q" refers to a quantized version, which has a reduced size). Just place the files in the "SpeechModelPath" directory:

```

ggml-base-q8_0.bin
ggml-small.bin
ggml-tiny-q8_0.bin

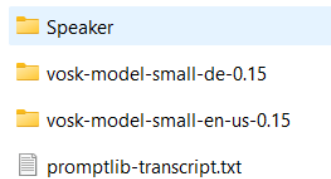
```

In addition to the Whisper models, it is also necessary to store the **runtime libraries** (i.e., some program files) of Whisper.net (source: <https://github.com/sandrohanea/whisper.net>) in the same directory. These are to be placed in a subdirectory called "runtimes" and are included in the installation package (with the subdirectory "runtimes", you can simply drag and drop it from the archive). The installation package can be downloaded at <https://apps.vischer.com>.

- 87 For both Vosk and Whisper, **only the "small" models should be used for live transcription**, since the large models exceed the computing power of a normal workstation computer. If the model is too large, the transcription also takes too long or is too delayed.
- 88 The **prompt library** for processing the transcripts can also be stored in the same directory. It is stored in the configuration file via the parameter "PromptLib\_Transcription". The full path with file name must be specified. Paths can always contain placeholders (see below, para. 256). The prompt library for transcripts uses the same syntax as the normal prompt library (see below, para. 261 et seq.).
- 89 If **speaker identification** is used with Vosk, an additional subdirectory "Speaker" must be created, into which the model (i.e. the directory



with the model) is then copied (e.g. vosk-model-spk-0.4; this is also available on the above-mentioned website):



90 Access to **Google** is already preconfigured, as this speech recognition is cloud-based, i.e., not on the local computer. Red Ink is configured to use Google's data centers in the EU for speech recognition. Google speech recognition is only available if the primary or secondary model for Google's Vertex API and consequently an OAuth authentication is configured (the same access codes are then used). If this is the case, Google automatically appears in the selection of models. When starting the transcription, the respective language must be chosen.

91 In a future version, we will try to offer models from other providers.

## G. Creating Podcasts and Audiobooks

92 The Analyze menu provides access to functions for creating podcasts and audiobooks (i.e., voice recordings of texts). Both require, however, that Red Ink is configured to use **Google's Vertex API** or **OpenAI** (hosted with OpenAI itself, not Azure), as its text-to-speech models are used. It is sufficient, though, if only the secondary Google or OpenAI model is configured. This is necessary because access to the Google and OpenAI text-to-speech models requires the same authentication as the language model.

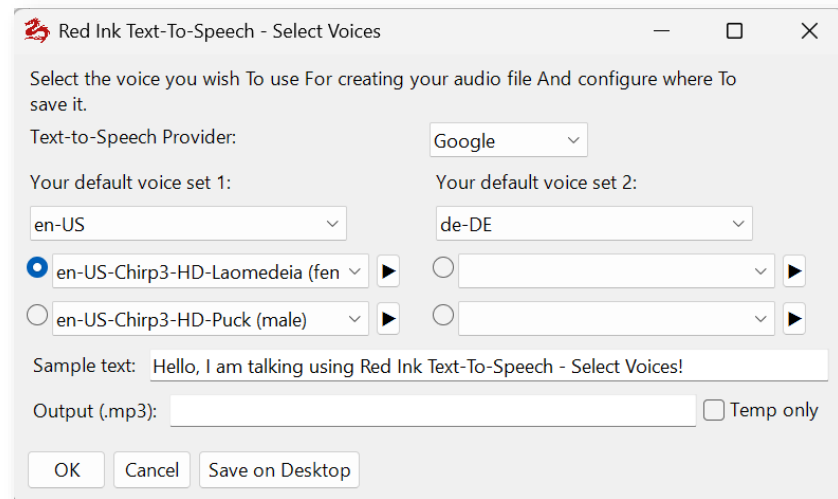
93 The **Create Podcast** function is similar to that of the popular Google service "NotebookLM". The selected text is first converted by the AI into a podcast dialogue between a host and a guest. This dialogue can be edited. Afterwards, this dialogue is converted into an MP3 file with two speakers.

94 The **Create Audio** function reads the selected text paragraph by paragraph and creates an MP3 file based on it. If the selected text contains a host and guest dialogue (the paragraphs must begin with "H:" and "G:" respectively, and the first one must have an "H:"), then the text is read with two speakers, as if it were a podcast (in this way, podcast scripts can be dubbed later). If this is not the case, then you can select whether Red Ink shall switch between two speakers (each time when a new chapter starts, based on what is recognized to be a title) or just one speaker. JSON-encoded text can also be processed, as long as it corresponds to the format required for the Google Text-to-Speech API, such as the multi-speaker encoding (if enabled). JSON texts are only supported up to a maximum length of 5,000 characters (Google's Long-Speech API has not been implemented). Other texts may be longer; to bypass the character limit, Red Ink dubs texts with Create Audio and podcasts paragraph by paragraph. A short pause is made af-



ter titles in normal texts. Lines with pure numbers or special characters are ignored.

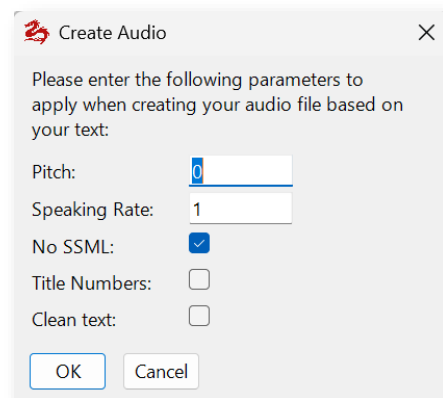
- 95 Alternatively, Create Audio can also be used to **add audio to a Powerpoint file with speaker notes**. To do this, the function is called without having selected any text beforehand. In this case, Red Ink asks whether a text file or a presentation should be converted to audio. If so, a window opens and the desired file can be dragged onto the window (or opened via the dialog box). If it is a text file, Red Ink opens its content in a new document; if it is a Powerpoint file, the speaker notes on each slide are read out and the audio integrated into the presentation in such a way that they can be played back automatically in Powerpoint later. It is then also possible to create a film from it – a video presentation with an artificial speaker (or two speakers, if alternating voices are selected for each slide).
- 96 In all cases, the user can select the desired voices for speech generation. Depending on the selected server location (which is controlled via the "TTSEndpoint" parameter, see 246 et seq. below), more or fewer languages and voices are available (at Google they are retrieved from the server; at OpenAI they are hard-coded and the same for all languages). First, the language must be selected, then the available voices are shown, from which two can be selected:



- 97 That two voices can be selected twice here is pure convenience: The tool remembers all four last selected voices, so that they reappear next time. The user can thus save their "favorite" voices for two languages and does not have to search for them again each time. If a podcast is being voiced, the user must choose whether they want the left or right set of voices; when reading a text aloud, it is possible to choose between all four voices.
- 98 With the button to the right of the voice selection, the respective voice can be heard; the text typed into "**Sample text**" is read aloud, including any SSML coding contained within (so it can be easily tested

whether they are supported by a voice – if nothing is heard, they do not work).

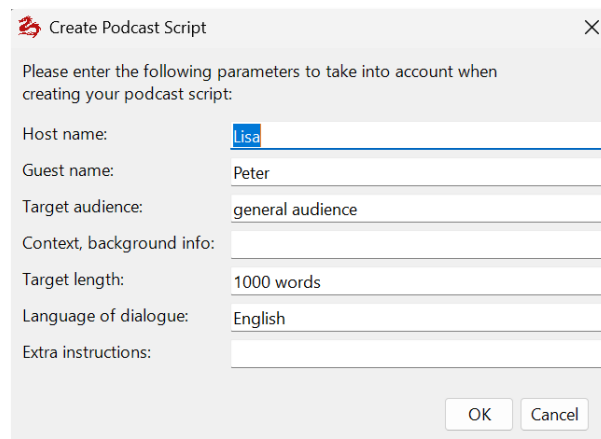
- 99 In the "**Output**" field, you can specify where the MP3 file should be written and what it should be called. The "**Save on Desktop**" button adjusts the path so that the file is saved on the desktop. If "**Temporary**" is selected, the generated MP3 file is played and then immediately deleted (however, temporary files are always created when generating MP3 files, which are then deleted ; the "%TEMP%" directory is used).
- 100 Not all voices sound equally natural; with Google, the largest and best selection exists for English and on the US endpoint. Not all voices support SSML coding (this allows text to be supplemented with commands for emphasis and pronunciation) or changing the pitch and speaking rate; OpenAI does not support this currently. If generating speech does not work, it may be because one of these parameters has been changed or SSML has been used even though the voice does not support it. These parameters are queried before generation:



- 101 The default value for "**Pitch**" is 0; a value of, for example, -0.5 reduces the pitch slightly, which usually makes the voice sound a bit richer. The "**Speaking Rate**" is normally at 1; a value of 1.1 makes the speaker talk a little faster, while 0.9 slows it down a bit. As said, this is only support by Google and only with certain voices. If "**No SSML**" is checked, it ensures that no SSML commands reach the speech generation, thus avoiding errors (the generated podcast dialogues contain SSML coding by default) because not all voices support them.
- 102 The parameter "**Title Numbers**" appears only with Create Audio and ensures that any numbering in titles is read along; in most cases, however, they will be rather disturbing (if necessary, the text can be provided with spoken numbers beforehand, e.g., if chapter numbers are to be read).
- 103 If you want each paragraph to be submitted to the LLM for "cleaning" before it is read, for instance for removing chapter references and other content that is not easily read out, you can select "**Clean text**" and will get the opportunity to define a prompt. The clean text function can

also be used if the text to be read aloud contains overly complicated sentences (which can lead to an error message). To avoid this, the LLM is prompted to make two sentences out of a complicated sentence, for example, without changing the content.

- 104 If a podcast script is generated, various parameters can also be recorded:



- 105 The **host's and guest's names** are required because the two roles will address each other in the generated dialogue. The target audience as well as the context or background of the podcast ("**Context, background info**") will also be provided to the AI for generating the podcast. The "**Target length**" contains information on the desired length, whereby the length of the podcast also depends on how much source material is available; in our experience, certain language models have difficulty adhering to length instructions, if time units are used. It is better to use the number of words. The "**Language**" specifies the output language, and in the "**Extra instructions**" field, further commands can be inserted that are passed to the prompt for generating the dialogue (along with "PreCorrection").
- 106 The generated podcast will be displayed in a separate window and can be edited there. If this is not cancelled with "Cancel", Red Ink will try to add voice to the script. It can also be copied to a normal Word document via the clipboard and voiced later via "Create Audio". Create Audio automatically recognizes podcast scripts by the alternating "H:" (for Host) and "G:" (for Guest)..
- 107 All dubbed MP3 files are played back after dubbing with an internal MP3 player. Cancellation is possible when using Create Audio with the "Cancel" button. The dubbing of a podcast runs (invisibly) in the background, i.e. work can continue; with Create Audio, it is shown which paragraph is currently being dubbed (if an error occurs, the beginning of the paragraph is displayed). Red Ink notifies you as soon as the podcast is ready for playback. If errors occur during dubbing, the corresponding sequences are missing in the final result (or the file cannot be played at all). If, for example, one voice is not compatible with the selected settings, only the other voice will be heard. In such cases, the

dubbing must be repeated with a new selection of voices, paying attention to the standard values for the parameters and not using SSML or having it filtered out.

108 When audio generation (Create Audio and Podcast) is running, Red Ink tries to set the computer so that it can no longer fall into **sleep mode**. This setting is reversed after the end of audio generation.

109 Regarding the **protection of your data**, it should be noted that with Google, the audio content is generated on a different system than the responses of the language model. While the latter may run in your own country, depending on the selected "Endpoint", the speech content may be generated on a foreign server.

## H. Integrated anonymization function

110 Red Ink has a simple, built-in anonymization function that can anonymize all texts sent to a model or special service and re-identify the returned texts. With this function, confidential content can also be used with service providers who are not sufficiently trustworthy or do not provide the necessary contractual assurances. The anonymization runs completely locally and is currently based on predefined search terms where the hits are replaced by placeholders. We are experimenting with anonymization models that also run on a local computer with moderate computing power, but have not yet found a suitable model that can be implemented reasonably and delivers the required quality.

111 Two things must and can be configured:

- For which model or special service the anonymization should take place and how: This can be stored for each model in the model's configuration, and can also be determined by the user with the file "redink-anon.txt" on their desktop.
- Which search terms should be replaced by placeholders and then reinserted. This is determined by the user with the file "redink-anon.txt" on their desktop.

112 Red Ink uses a simple but powerful configuration via the file **redink-anon.txt** on the user's desktop to determine both the "WHEN" and the "HOW" of anonymization. By default, *no* anonymization is active ("none; 0") until the mentioned file is either created and customized by the user or it is provided for the corresponding model or special service in its configuration (parameters "Anon" and "Anon\_2" in "redink.ini").

113 The **structure of the file** divides its content into sections marked with square brackets:

- What follows **[All]** applies to all models and services;
- What follows **[ModelA, ModelB, etc.]** takes precedence for exactly these models. Thus, a section can be defined for one or more models. They are separated by commas. The model name

corresponds to the one from the configuration of the model or special service;

- Lines starting with ";" are ignored. This can be used for comments.

114 Within each section, the anonymization mode and type are first defined on one line:

```
Anon = <mode>; <type>;
```

where:

- <mode> is a parameter such as *none*, *silent*, *ask*, *askshow*, *show*, and
- <type> is a number from 0 to 4 (see below).

The parameters "Anon" and "Anon\_2" in the configuration of the models and special service are, by the way, also determined according to the same structure.

115 The **search terms and patterns** for which placeholders should be inserted are then listed, either as simple terms or search terms with wildcards, or as regex lines. Some examples:

- Regex: \b[A-Z]{2}\d{4}\b
- "Max Mustermann" (exact text)
- Client\* (Wildcard \* → any combination of letters/numbers)
- ACME\*{{Company}} (instead of the placeholder "redacted" the placeholder "Company" is used)

116 The **placeholders** always have the same format:

```
<<prefix>_<GroupID four-digit>_<SubIndex>>
```

will result in "<redacted\_0003\_1>" for the first entity of the third rule. The group IDs are assigned internally and sequentially, sub-indices count newly found, different occurrences of a term for each rule. At the same time, the module defines in a list which original texts correspond to which placeholder.

117 For the user, the placeholders are not necessarily relevant, because the system automatically remembers which term has been replaced by which placeholder and uses this term again at the end. However, the use of user-defined placeholders can help the language model to understand the context even with anonymized texts and thus provide a better answer. If, in the example above, all occurrences of ACME like "ACME" are replaced by a placeholder with the word "Company" instead of "redacted", the model knows that it is a company name and can take this into account, even if it does not know it. Because all occurrences of ACME have the same group ID, this context is also preserved. Where these things are not all required, the definition of custom placeholders can also be omitted.

118 The supported **modes** are:

- **none**: no anonymization (this is the default);
- **silent**: automatic, without asking (the user doesn't notice it and doesn't receive the anonymized text for prior review and adjustment; those who want to know how the anonymization function processes their text can use the Anonymization function in the Analyze menu);
- **ask**: asks via yes/no dialog whether to anonymize and then does so silently (like silent);
- **askshow**: asks whether to anonymize and if so, the anonymized text is displayed before use so that it can be adjusted (i.e., manually anonymized) or canceled. If anonymized manually, the result must also be manually re-identified, i.e., the placeholder has to be replaced manually by the desired value;
- **show**: always anonymizes and displays the anonymized text for review before use.

119 The **types** at a glance are:

- **0**: no anonymization (default value);
- **1**: The user can enter their search terms (e.g. a name), separated by commas, in an input window; the last used search terms are displayed (Regex does not work here);
- **2**: Like 1, but the input field is empty each time;
- **3**: The user is not asked for search terms, but the file `redink-anon.txt` on the user's desktop is accessed directly and an error is displayed if it is missing. This is the only type in which Regex works;
- **4**: The user can enter their search terms as in 1 in an input window; the search terms provided for the model are displayed, which can be adopted (Regex does not work here).

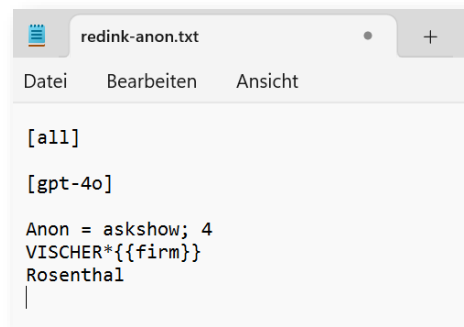
120 The anonymization is automatically called in Word, Excel and Outlook based on the configuration before text is passed to a model or special service, but **only with regard text selected in the user's document**. The prompts to the model, including everything that is sent to the model via the freestyle command (such as document content inserted using "{doc}" or the content of a file with "(file)" or the clipboard "(clip)") is **not anonymized**.

121 Anonymization uses by default the configuration of mode and type stored in the model or special service. However, if the user has the file **redink-anon.txt** on their desktop, the values there override the default configuration. If the local `redink-anon.txt` file contains information both under "[all]" and for a specific model, the information for the specific model takes precedence. This way, each user can quickly and

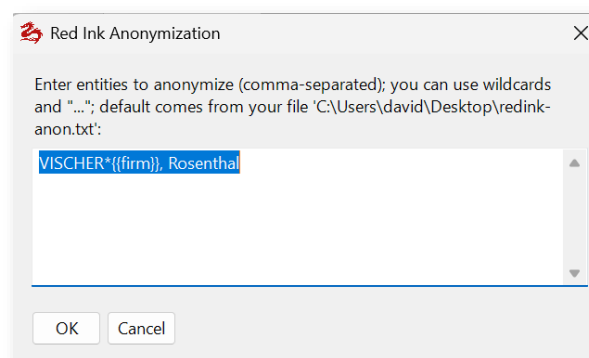
VISCHER

easily make the setting that suits them best; only the file redink-anon.txt needs to be adjusted with a simple editor.

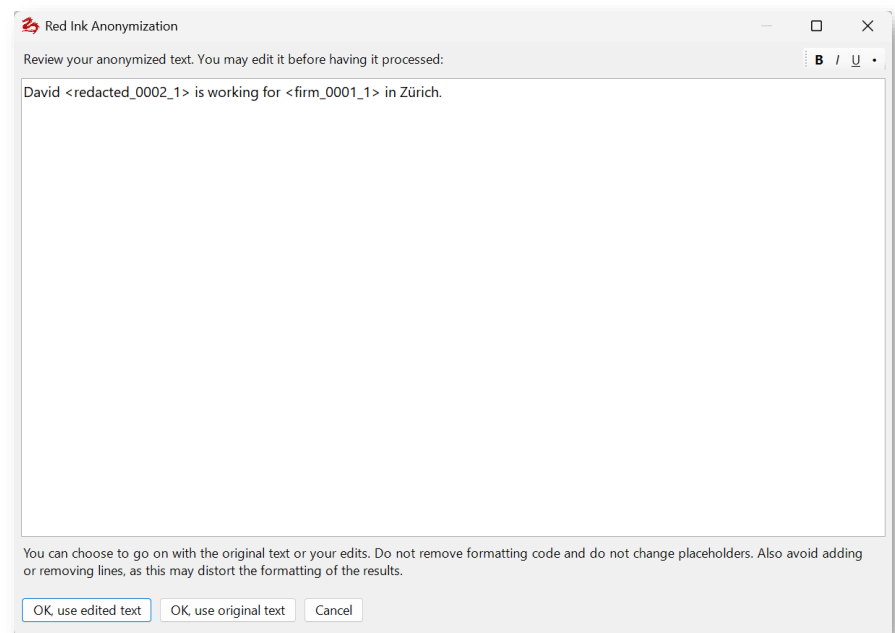
- 122 Here's an example (anonymization is only used when using gpt-4o, the user is asked whether to anonymize, they can enter the search terms based on the content of this file and the result is displayed beforehand; all occurrences of "VISCHER" and the search term "Rosenthal" are anonymized, whereby "firm" is used as a placeholder for "VISCHER", whereas the standard placeholder redacted is used for Rosenthal):



- 123 Once anonymization is complete (the entire text is processed), and if the **show** or **askshow** modes were selected, Red Ink then opens an editing window where the anonymized result can be manually checked and adjusted (placeholders and formatting should not be changed). As soon as the window is closed with the corresponding release button, the add-in continues with the final, anonymized text as usual. If canceled, nothing is passed on and it is aborted or the add-in behaves as if nothing had been returned by the language model.
- 124 If the text "David Rosenthal is working for VISCHER in Zurich" and a redink-anon.txt file with the above content were used in Word, the following prompt window appears after the yes/no question of whether to anonymize:



- 125 The result then looks like this:



126 If the language model or the Special Service has returned a text, Red Ink automatically replaces the placeholders with the previous values. It must be manually checked whether the replaced terms are in the right place, because the language model did not see them in plain text. Here, as mentioned, descriptive placeholders can increase the answer quality.

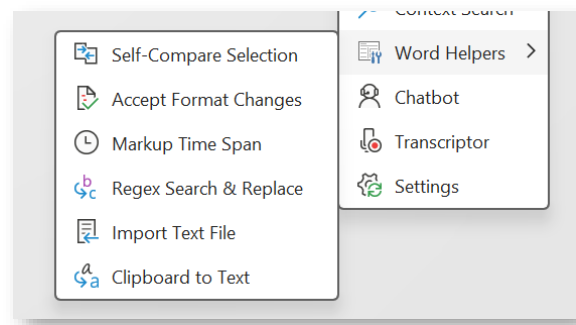
127 The anonymization can also be tested without a language model by using the "**Anonymization**" function in the "Analyze" submenu in the Word add-in. There you can choose between anonymization type 3 and 4. At the end of the text, the table with the assignments of the placeholders is also displayed.

128 Anyone who wants to check for themselves what is actually sent to the language model can run the Red Ink code in their own development environment in debugging mode. If the "Debug" parameter is set to True, the string sent to the endpoint and the string received from it are output.

## I. **Word helpers – practical everyday helpers (almost) without AI**

129 Word helpers are additional functions that (with one exception) actually have nothing to do with AI but are missing from Word and can be very useful in everyday life. That's why we built them in right away. They can be accessed via a dedicated submenu in the menu tile and via the context menu:

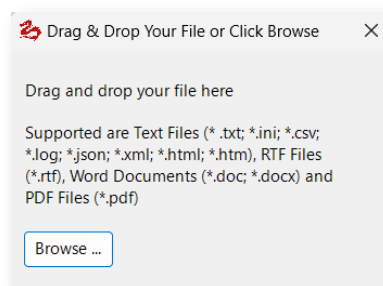




- **Self-Compare Selection:** The first half of the selected text is compared with the second half of the selected text in the same document. If, for example, two or four paragraphs are selected, a markup of the second is compared with the first or the first two paragraphs are compared with the second two paragraphs. In our work, there is often a need to briefly check what has changed between two clauses. Previously, two separate documents had to be created and compared for this. This is no longer necessary with this function. You can use Settings to define which markup technique (Word, Diff or DiffW) is used (see para. 25 above). Since short texts are usually compared here, Diff or DiffW often makes sense in practice, even though the function is less reliable.
- **Accept Format Changes:** In the selected section, all formatting changes are accepted, but only these. They are often disruptive in the display on the screen or in printouts. Although they could be hidden, this requires an additional step each time. So, this way the problem can be eliminated with a single operation.
- **Markup Time Span:** This function can be used to calculate how much time has passed since the first and last markup or comment in the selected area, based on the time entries created by Word creates when it performs markups. When the function is selected, you can specify whether the calculation should only take into account the comments of a particular author (otherwise all comments and markups will be considered). This function can be useful if you want to calculate retrospectively how long a revision took, for example, for entering in a timesheet.
- **Regex Search & Replace:** This function can be used to perform a so-called Regex search. Regex (short for "Regular Expressions") is a powerful text search and manipulation tool that recognises patterns such as phrases, numbers or special formats in text, rather than just finding exact matches as in a normal search. It is not available through

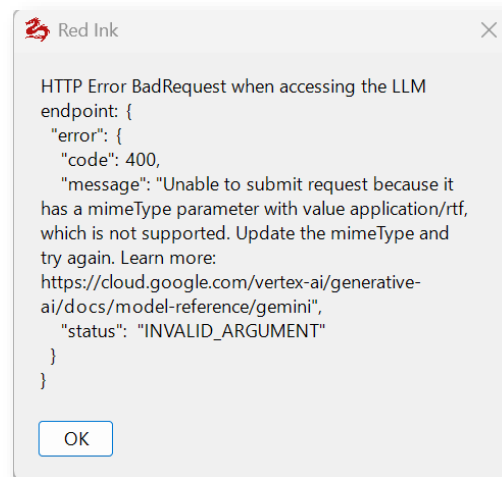
the Word interface. This function allows you to enter one or more Regex search patterns and instruct the add-in to replace hits with one or more texts. The add-in first asks for the search pattern, next the search options and then the replacement texts. If you want to enter multiple search patterns, each one should be entered on a new line (without blank lines), and the appropriate replacement text should also be entered on each corresponding line. The search options apply to all search patterns. If a Regex search pattern is non-compliant, an error message is displayed. If there are no replacement texts, the first match is displayed. The search and replace only takes place in the selected text. If you would like more information about the possible search patterns and options, you can find them on a Microsoft help page, which can be accessed via <http://vischerlnk.com/regexinfo>.

- **Import Text File:** This is the same function that is available in Freestyle and allows you to insert the contents of a text document directly into the current document as text. This is not possible in Word without further ado, especially not with PDF documents (however, the helper only processes text from PDFs that is actually available as such, i.e. that can be searched for; if text recognition is also required because the text is only available as an image, e.g. in the case of scanned PDFs, this must be done beforehand with a separate PDF program or – where the model supports it – with below Word Helper or Freestyle, see below). When the command is selected, the following window opens and the file can be easily imported by dragging it onto the window with the mouse. Alternatively, it can be selected using the button.



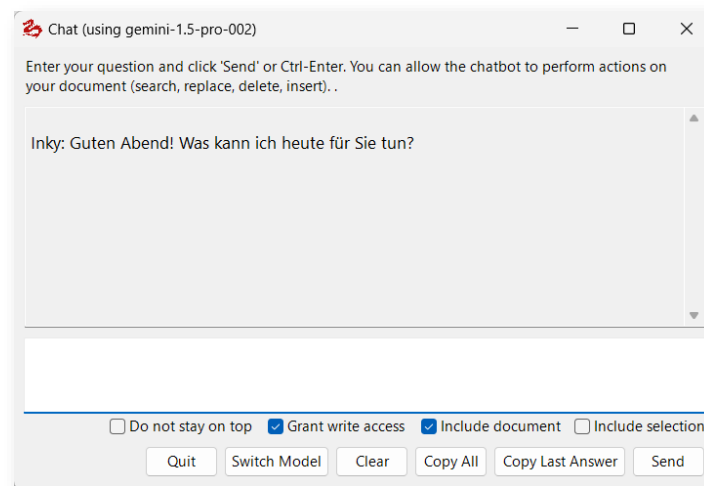
Alternatively to this function, when using language models that can process not only text input but also files, the following "Clipboard to Text" or Freestyle function with the trigger "(file)" or "(clip)" can be used. Then the language model translates the content of the file, which in the case of a PDF, for example, can also work if text recognition is required.

- **Clipboard to Text:** This function, in contrast to the other Word helper functions, is AI-based and only works if the primary model supports the processing of binary objects ("APICall\_Object" parameter). If it is selected, the LLM is asked to convert the contents of the clipboard to text, without anything having to be entered, and regardless of whether the data in the clipboard is an image, a text document, or an audio or video file. This function is very practical, for example, if you want to copy a part from a document that is open next to Word into Word, but this does not work with copy & paste. It is then sufficient to take a screenshot of the area with Shift-Windows-S and select this function – and the text (or image description) is inserted. If the format of the clipboard's contents is not supported, the model returns an error message like this:



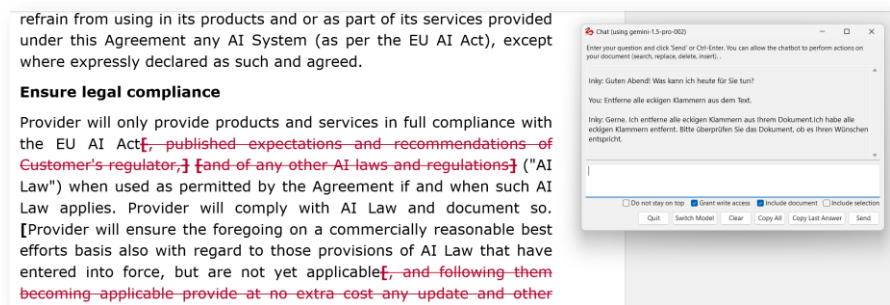
## J. Integrated chatbot "Inky"

- 130 The Red Ink Word add-in offers an integrated AI chatbot. It is accessed and appears in a separate window that can remain open while you continue to work; you can set whether or not it should always remain visible ("Do not stay on top").



- 131 The chatbot "Inky" is on the one hand a normal AI chatbot, which can have a dialogue with the user and also remember the dialogue (within the scope of the configurable limitation), i.e. can include earlier questions and answers. It initially communicates with the user in the language in which Word is configured, but changes if the user speaks in a different language. The chatbot does not yet have its own internet access, but can spontaneously access it when creating texts for short questions, for example, about formulations or to explain terms. The user enters their question in the white input field; it is concluded with Ctrl-Enter or by clicking the "Send" button. The latest answer from the LLM can then be copied to the clipboard and used in the text ("Copy Last Answer"). It is also possible to copy all answers or delete a dialogue if the user wants to change the topic. It is also possible to switch back and forth between the primary and any configured second model (the context is retained). The window can be closed with "Quit", but the previous dialogue will be retained until the next time or until "Clear" is used.
- 132 If the chatbot is to remember something from a text for later, it must be told to do so. It can see the current text (i.e. the active document), but only as long as one of the two checkboxes is activated, i.e. it is not included in the chat history. To enable the chatbot to remember, it will repeat the information when it is asked to remember something. To make it easier to switch between documents, the chatbot also sees the name of the document. But the same applies here: if you want the chatbot to remember something, you have to tell it to do so ("Remember the place where the price adjustment is mentioned and the name of the document"). This can be useful, for example, if two documents have to be compared in parallel (alternatively, the Freestyle function with the addition for importing a second document can be used for this).
- 133 What distinguishes this chatbot from others, however, is that it not only sees the current text (or a selection of it) that the user is editing, but can also change it if this is permitted ("Grant write access") and

the user requests it. Inky can search for and mark text passages, it can delete or replace parts, it can insert text, including before and after certain passages in the text. It is therefore possible, for example, to instruct the chat bot to "mark all instances in the contract where costs are mentioned". It will respond and carry out the command. This is done in Markup mode so that each adjustment can be undone (the execution can be cancelled by pressing "Esc"). For example, the command was given to remove all square brackets from the contract template, which it also executed:



134 The result should, of course, be checked – not only for any content errors but also to see if there are any replacement errors, for example, where the same term has been replaced multiple times. The programming of Inky is independent of the other functions of Red Ink. It is therefore possible that Inky could implement something differently from the Freestyle function, for example.

## K. Further tips for using Red Ink in Word

135 With **longer texts**, using the markup function is often not effective because the AI takes too long to do so and makes too many mistakes or formatting cannot be preserved. Here we have two tips:

- Especially where a larger document is to be selectively commented on or revised, the **"Bubbles:"** function has proven very useful in practice: instead of having a markup made on the text, Red Ink or AI is asked to annotate the text (for example, enter "Bubbles: Go through the text and show me all the places you find linguistically unclear and how I could do better." in Freestyle). Then use **"Apply comment"** in the Improve submenu to have the resulting bubble comments implemented in your text. The advantage of this is that the the tool no longer needs to process the entire text as output (language models usually are only able to generate a fraction of the amount of text as output as they can process input). Another strategy for very large documents can be to work with smaller portions of the text.
- If direct editing of a text is unavoidable, such as with translations, we recommend proceeding step by step. **Only a few paragraphs** are marked at a time, and then the relevant function is selected. To maintain formatting, especially when working with

style sheets, we recommend activating the **Keep paragraph** format option (see para. 24 above). The use of style sheets also has the advantage that documents are formatted better and more consistently. While it is possible to ask Red Ink to also remember the formatting of individual characters (Keep format), this function slows down processing massively because the AI has to be provided a lot of formatting information to retain it, whereas with Keep paragraph format, only paragraph formatting is remembered. Experience has shown that it is easier to proceed paragraph by paragraph and manually adjust individual character and word formatting (e.g., bold). However, if you have to translate a longer document, you should preferably use – if permitted under data protection law – a different tool such as "DeepL", which is specifically designed for this task (it receives the text as a file).

- 136 When using Red Ink, **tables can be created**. This is sometimes advantageous because the evaluation of the results is clearer. In the command, the AI is to be instructed to present the result in the form of a table. This will only be created for text codes with a separator chosen by the AI. However, this text can easily be converted using the Word command "Convert Text to Table..." (simply select the table, open the command and enter the separator chosen by the AI):



- 137 Red Ink itself can handle **tables in text**. If it encounters one in the selected text, it asks whether it should go through the individual text blocks before and after it and the individual cells separately, so that the table is not destroyed. This then generates a corresponding number of queries. This can also be declined. If it is cancelled, nothing happens. If markup is activated, the diff method (not DiffW) is used because it is the most efficient for this purpose.
- 138 If you want **Red Ink to revise tables** (e.g., to have them automatically filled in or adjusted), you should not do this in Word, but rather copy the table into Excel and ask Red Ink there, via the Freestyle func-

tion, to revise the table. This works much better and faster than in Word.

- 139 If you are in Word and want to enter a text in Red Ink that consists of **several parts** that are to be controlled individually (e.g. two text passages in the same document that are to be compared), it has proven useful to mark these with HTML-like "tags" to delimit the content. Example:

`<FIRSTPART>`

*.... Insert the first part of the text.*

`</FIRSTPART>`

`<SECONDPART>`

*.... Insert the second part of the text.*

`</SECONDPART>`

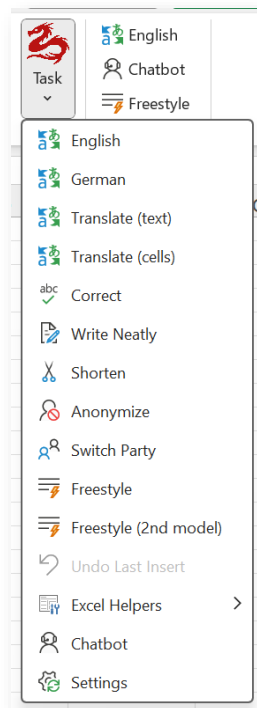
The terms used should be logical or meaningful. The prompt can refer to this and an advanced large language model will be able to distinguish the texts. The add-in also uses this technique internally (by enclosing texts passed to the AI with `<TEXTTOPROCESS>` and `</TEXTTOPROCESS>`).

- 140 The use of **paragraph markers** ("`\n`") in the prompt can also be useful to separate individual elements from each other. This can also help the language model to better understand and implement the prompt. It has also proven useful to clearly **identify individual steps** of an instruction (e.g. "Do (1) an analysis of all changes in the text and (2) create a summary of the three most important changes").

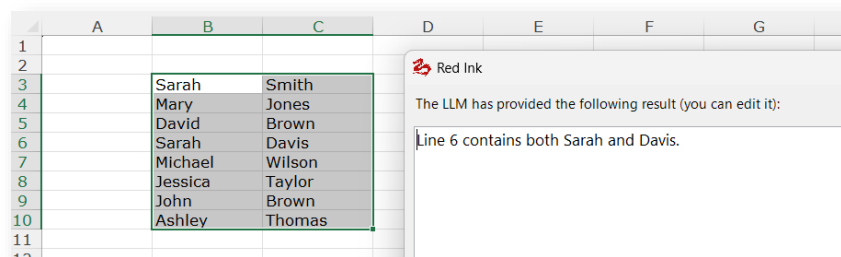
- 141 The add-in for Word also processes various formatting instructions in the **Markdown format** (".md") during output, such as bold, italics, titles or bullets. Various language models support this. In Freestyle, the model can be specifically instructed to use this format for output ("Output in Markdown Format") if necessary, provided it does not do so by itself. However, tables and images are not supported. For tables, proceed as described in para. 135 above.

## **L. Red Ink functions in Excel**

- 142 The add-in works in Excel in the same way as in Word, although the functions are slightly different. There are also predefined functions and a Freestyle command. Access is possible via a context menu (if not disabled) and via the tiles. Shortcuts can also be defined as in Word (see above):



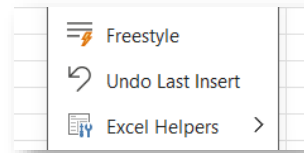
- 143 As we are working with cells here, Red Ink provides for two different methods for executing functions: The add-in can be instructed to proceed **cell by cell**, which is useful for translations or anonymisations, for example. Or it can be instructed to consider the **selected cell range as a whole**, for example if a formula or text content is to be inserted based on the existing content of the worksheet or if a specific piece of information is to be searched for ("Which row contains both Sarah and Davis?", in the following example the question was asked in English):



- 144 All predefined functions (e.g. the translation function) proceed cell by cell because this is the main use case. The Freestyle function (para. 150 below), on the other hand, considers the selected cell range as a whole by default, unless a cell-by-cell approach is requested
- 145 When proceeding cell by cell, it must be ensured that the cells are not **locked**. If all selected cells are locked, an error message is displayed. Otherwise, only the unlocked and non-empty cells are processed.
- 146 Attention: **Excel is not able to undo** the amendment of the cells by the add-in. You may therefore want to create a copy of the content

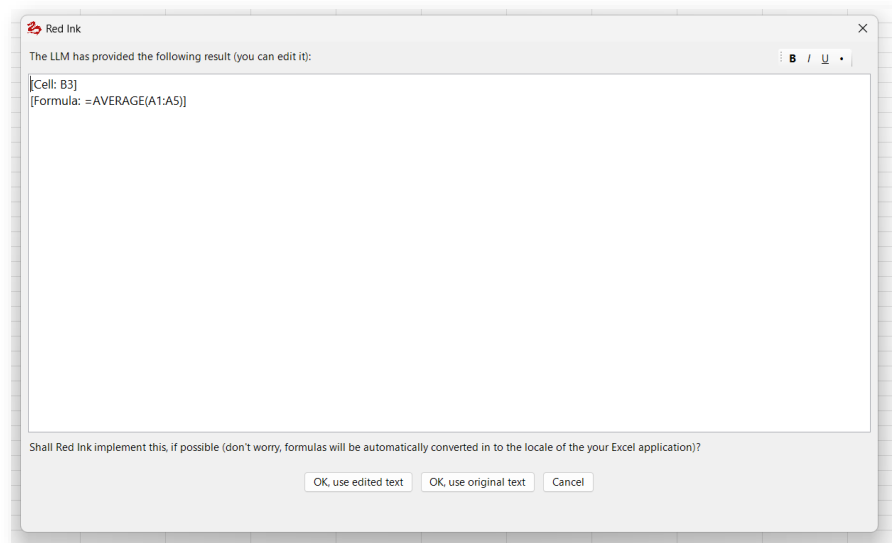


(e.g. on a separate worksheet) before using the AI. Alternatively, you can use the **Undo Last Insert** command that is accessible via the Red Ink main menu (however, be aware that you can't restore what Red Ink inserted after you have used this command should you again wish to have back the content generated by the AI):



- 147 In contrast to Word, the add-in in Excel does not take into account any **formatting** within the text of a cell, i.e. it is lost. However, the formatting of the entire cell is not changed.
- 148 In the menu, a distinction is made between "**(text)**" and "**(cells)**" in the translation functions. The difference concerns the question of whether the function only processes cells that contain text (including numbers) or also those that contain formulae ("(cells)"). If formulae are also included, the add-in will take this into account. The language model is instructed to only translate the texts within formulae, not the words in the formulae, as these would otherwise no longer work. However, the AI can also be asked to adapt formulae using Freestyle. Before this is used productively, this function should be experimented with so that it can be determined what works best and how for the specific case.
- 149 In Excel, the similar "**Write Neatly**" function is available instead of "Improve ". It differs in that it has a slightly different internal prompt, which is designed to convert keywords into complete sentences and, if necessary, to take into account a context (which is queried beforehand, but is optional). The newly formulated sentences replace the previous text in the cell.
- 150 The **Freestyle** function (see para. 33 above) is also available in Excel. However, the prefixes and triggers familiar from Word do not work here because they do not make sense in Excel or are not normally required. Basically, the add-in will transfer the content of the entire selected cell range, including values and formulae, to the AI together with the command. In this way, the AI can be asked, for example, what the range in question means, how it works or whether it calculates correctly, for example. The AI can also be asked to create a formula for a specific task or to fill out a form ("Fill out C5:C11 like D6:D11 has been filled out, but take into account the use case in C3 and the seven questions on the left."). The result is displayed in a window and can be edited. If formula or cell values are returned, the add-in can also attempt to implement them immediately (they are specially coded with square brackets for this purpose). If you do not want certain of the suggested adjustments, you can delete them in the editor

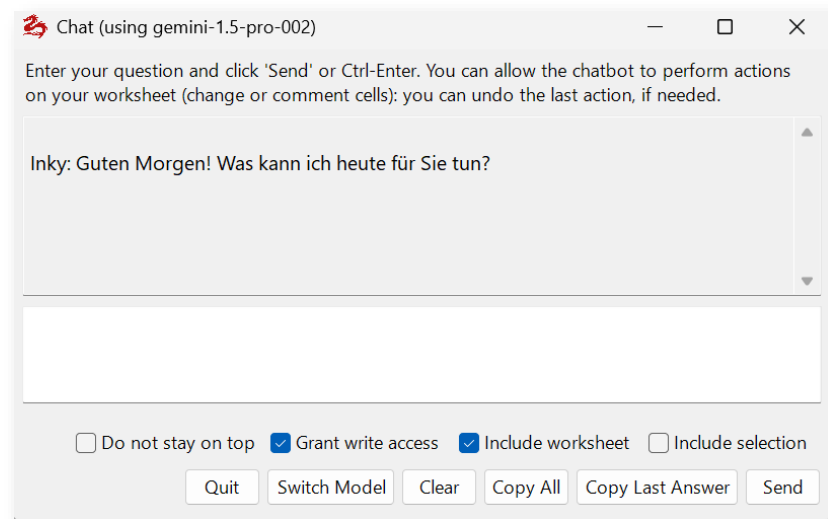
before execution or adjust them manually. The result of the AI is also copied to the clipboard. Adjustments can also be made outside the selected cell range with this command, but the AI only sees the selected cells. If the AI is to include existing content in its considerations, the relevant cells must be selected. If the trigger "**(color)**" is added, existing color information is also communicated to the AI (e.g. "Fill in all cells with a light blue background with X. (color)"). However, it is also possible to use Freestyle without any selection if the contents of the worksheet are irrelevant. The command "Calculate the mean value of A1:A5 in B3." produces the following result even without any selection (if "OK" is then pressed, Red Ink will insert the relevant formula in B3; if necessary, it will be converted to the local language format):



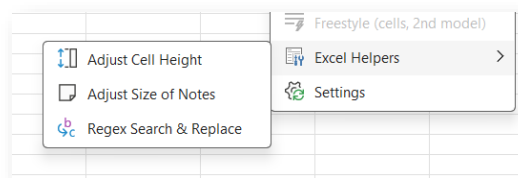
- 151 However, if Freestyle is only to proceed cell by cell, the command must be preceded by the prefix "**CellByCell:**" or "**CBC:**". If only cells with text content or a purely numerical value are to be edited, the prefix "**TextOnly:**" must be used.
- 152 In Excel, the Freestyle feature, similar to Word, can output to a pane. For this, either "**Pane:**" must be prepended, or in the usual output, the "**Transfer to Pane**" button is used to move the content to a pane. In the pane, it is also possible with the respective button to only apply the selected content with the square brackets in Excel.
- 153 As in Word, you can also work with the prefix "**Bubbles:**" in Excel. The output will then be added in the form of comments to the relevant cells. Also as in Word, the content of external documents can also be read into the prompt by using the placeholder "**{doc}**" (e.g. "Fill cells A1:C20 with the content of this text: {doc}"). The triggers "**(file)**" and "**(clip)**" for inserting a file or the clipboard contents are also supported.
- 154 Freestyle normally accesses the current worksheet. If data from other worksheets is also to be included, the trigger "**(addws)**" must be in-

cluded in the prompt (the position does not matter). If it is specified, the user can, after entering the prompt, choose which of the currently open worksheets they want to additionally pass to the AI (it does not matter if the worksheets are in different files). Optionally, all other open worksheets can also be passed to the AI.

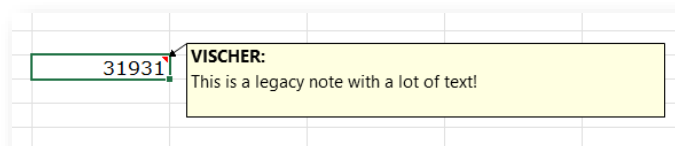
- 155 The **prompt library** and the **Ctrl-P** command are of course also available for Freestyle in Excel to re-insert the last command used.
- 156 A practical tip: If **several cells are to be translated** or otherwise edited, the Freestyle function can alternatively be used, so that all cells to be translated are marked and freestyle is given the command "Translate to French" (as an example). In this case, all cells are transmitted to the language model at once and it delivers the adjustment instructions in one go, which can then be implemented. This works much faster than if Red Ink translates each cell individually. However, this does not work with merged cells. In this case, they are emptied (due to the way Excel handles merged cells); this can at least be prevented by manually deleting the corresponding commands, which are supposed to insert empty content, in the window for editing the commands before they are executed.
- 157 To prevent long waiting times, Red Ink automatically limits the selection of rows and columns to those areas that are actually used. If a cell pass takes too long, it is also possible to abort it with **"Esc"**.
- 158 Furthermore, the add-in has a **chatbot** functionality that works analogously to the chatbot within Word (para. 130 et seq.) and is called accordingly. The chatbot in Excel can also make direct changes to the current worksheet, if permitted (with the corresponding checkbox "Grant write access"); it can fill cells with content and formulas and insert comments. Depending on whether this is selected, the entire current worksheet or the current selection is transmitted with each question and each task to the chatbot. For larger worksheets, it may therefore be advisable to work only with selections of cells, because otherwise the transmission takes a long time (a warning is also issued). When making selections, the AI is automatically informed of the cell and font color information (so that it can be referenced). Unlike the chatbot for Word, the checkboxes "Include..." do not have to be selected here for the chatbot to have write access to the worksheet. However, if it does not have access, it cannot take existing cell contents into account when writing. If you wish to undo a change made by the chatbot, use the "Undo" function in the Red Ink menu. This works only with the last change made by the chatbot, and not with comments. If additional worksheets (besides the current one) should be included, the trigger **"(addws)"** can be added to the current input. Red Ink will then display the worksheets that are also currently open and one (or all) can be selected. Their content will then be sent with this request and will only be available for this one. The trigger must therefore be repeated for subsequent requests if necessary.



159 The add-in also has three Excel helpers:



- **Adjust Cell Height** adjusts the cell height in the selected range to fit the text, even if it spans multiple cells. Excel's Autofit function cannot do this; for merged cells, it always sets the cell height to the height of one line of text, which means that the text is then not fully visible.
- **Adjust Size of Notes** adjusts the size of notes stored with the selected cells. This can be a tedious task to do manually. A minimum size is defined, and the font size is taken into account up to a certain point.

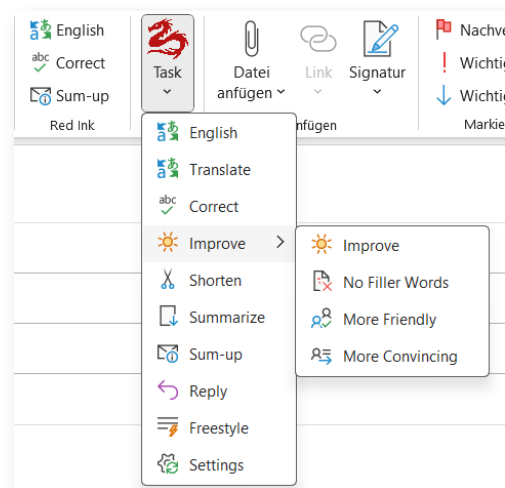


- **Regex Search & Replace:** This function works the same way as for the Word helper (see para. 129), one cell at a time in the selected range.

160 Various configuration values can be set in **Settings**, in a similar way to Word (see para. 26 et seq. above). If they are saved, a local configuration file is created only for Excel. Otherwise, a centrally defined configuration file (if any) or that of the Word add-in is used.

## M. Red Ink functions in Outlook

- 161 The add-in works in Outlook in the same way as in Word, although the functions are slightly different here. The functions can only be accessed via the tiles because Outlook does not support any add-in-specific context menus. Therefore, it is not possible to define any shortcuts the way it is in Word and Excel. The tile of Red Ink is located both in the main menu and in the menu of the window that opens when composing an email (i.e., when drafting a new email, replying to an email, or forwarding an email in a separate window). If an email is edited only in the pane area on the right-hand side of Outlook and a Red Ink command is selected, Red Ink opens the respective email in a separate window and only then executes the command. It should also be noted that only HTML and RTF emails are supported, i.e., emails that can contain formatting (this can be set in the "Format Text" tab). The tile then appears there (it is positioned slightly differently and the Quick Access tile is placed in front of it because Outlook would otherwise collapse it if necessary):



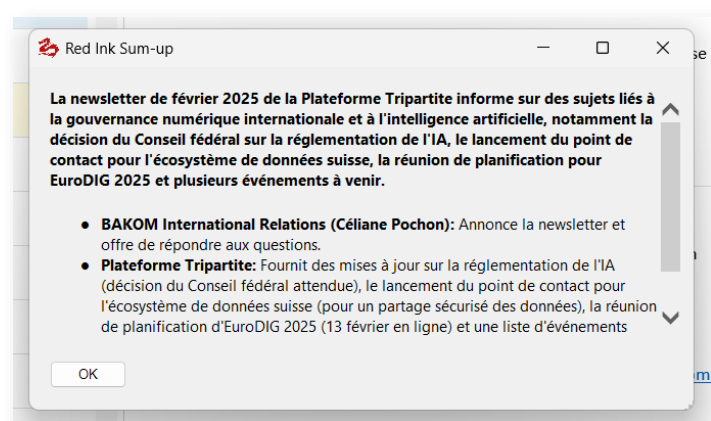
- 162 The Outlook add-in offers a selection of the same functions that are available in Word (see para. 22 et seq. above), i.e. pre-programmed functions and the Freestyle function, which can be used to enter any prompt, with or without selected text (but no helpers).
- 163 A practical tip for using **Translate**: For target languages that distinguish between formal and informal forms of "you" (such as German), the AI is instructed to make the right choice based on the context and the words already used. If a person is addressed by their last name or a greeting formula with a full name is used, it will assume that the formal form is necessary. When translating, the salutation or greeting formula should therefore be selected as well, so that the AI can take this into account. It only sees what is selected. This also works in Word.
- 164 As in Word (under World Helpers), Outlook also offers the **Clipboard to Text** function, provided that the configured primary model supports

this function. It works by asking the model to convert the contents of the clipboard into text (e.g., text in a screenshot, text in a voice message). If an email is currently being composed, the function inserts this text there. If this is not the case, the clipboard is filled with the text extracted or generated by the AI when this function is called, and it can then be inserted into any application.

165 **Reply** is an Outlook-specific function used to prepare replies to e-mails. First, you have to select the parts of the previous e-mail chain to which you want to reply. The top e-mail of the selected area should be the e-mail to which you are replying directly (if nothing is selected, the entire e-mail chain will be used). The add-in will take this sequence into account. Specific instructions and information for formulating the response can be entered in the window that opens. The response is inserted at the top and can be edited as required.

166 The **Sum-up** function summarises the email chain (by contact). The AI is instructed to do so in the language of the mail. You must be replying to the relevant email or forwarding it. Then select the relevant parts (or the entire email chain will be taken into account) and select the "Sum-up" function. The AI will display a summary of the email chain at the top of the email. This can be useful for longer emails to get a quick overview.

167 Unlike all other Red Ink for Outlook features, the functions Sum-up and Translate can also be used with **emails that are not open for composing**. It is sufficient to select an email in the Outlook overview for it to be displayed in the field on the right. If Sum-up is then pressed, Red Ink creates a summary or a translation (as the case may be in the language entered) and displays it in a separate window:



168 If **multiple e-mails** have been selected (with none of the opened), then sum-up will go through all selected e-mails (and try to extract only the latest one of each mail-chain to save time) and provide a short overview of the most important and most urgent of these e-mails. E-mails that have already been answered are disregarded.

- 169 Certain functions (in Freestyle by adding "Markup:" before the prompt) work with a markup function, although Outlook does not actually offer one. In this case, the AI's text is displayed first, followed by "MARKUP:" and a markup that shows the insertions and deletions made by the AI in colour. Because these markups technically cannot be "accepted" within Outlook, the output is for information purposes only. It must be deleted manually by the user as soon as they are satisfied with the text. As in Word (para. 25), Settings can be used to specify which markup method should be used, although in Outlook only "Word", "Diff" and 'DiffW' are available because only these make technical sense here. For "Diff", the same character limit is used as in Word and handled in the same way. However, it is possible to configure independently of Word whether markup should be created automatically for the pre-programmed functions where it makes sense (e.g. for "Correct" and "Shorten"). As in Word, it is also possible to cancel the "Diff" markup output with the "**Esc**" key if it takes too long. In practice, "DiffW" has proven to be ideal for everyday use.
- 170 As in Word (para. 25), the configuration can be used to determine for all functions (except Freestyle, Reply and Sum-up) whether the text generated by the AI should **replace the existing text** or be added to it.
- 171 Further configuration is possible, as in Word (para. 25), to determine whether the add-in should attempt to preserve basic **formatting** (such as font and bullet points) when using the translation, correction, improvement and abbreviation functions, or whether it should work in pure text mode (which means that special formatting will be lost in the output; Red Ink, however, tries to preserve simple formatting such as bold text). The latter takes more time because all formatting must also be transferred to the AI (which occurs in HTML format). To keep the response times within limits, only the most important formatting is retained. When it is reproduced, it is possible that it will not exactly match the previous display. The function for limiting the preservation of formatting to a certain number of characters is also available here and can be configured via Settings or switched off with 0.
- 172 The **Freestyle** function is also available in Outlook, but with a reduced scope compared to Word. However, as in Word, you can work with the prefixes "**Markup:**", "**MarkupWord:**", "**MarkupDiff:**" and '**MarkupDiffW**', but not "MarkupRegex:", because this does not make sense for Outlook (see para. 25 and para. 39). Another option is "**Replace:**", which inserts the language model's answer in place of the selected text (e.g. "Replace: Find a more polite way of saying this sentence").
- 173 The triggers available in Word are not available in Outlook because they are not normally needed in that environment. However, the functions for preserving formatting are not supported in Freestyle for Outlook (unlike the Word version). But Freestyle for Outlook does include

the prompt library, and the last prompt written in the current email window can be inserted using **Ctrl-P**.

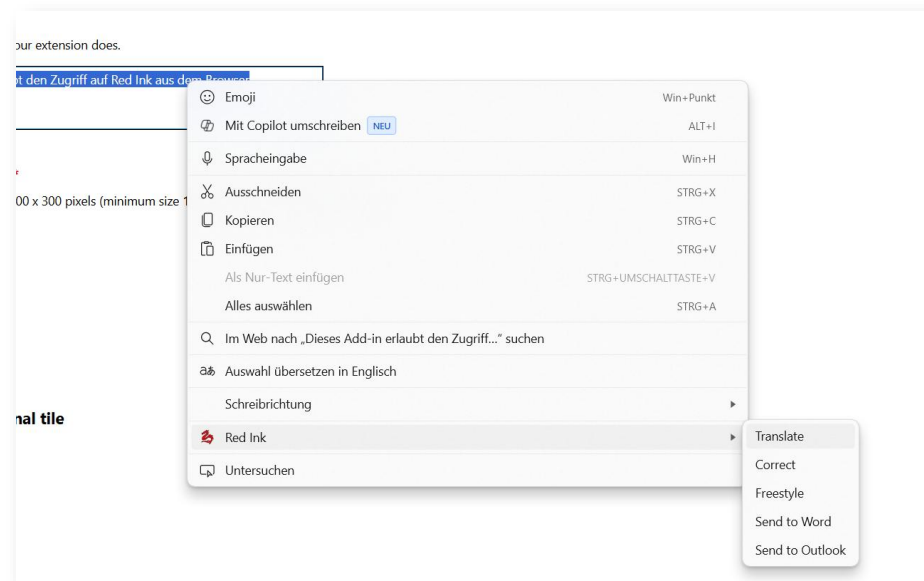
174 There is no dedicated Freestyle command in Outlook for accessing the **secondary** and, as the case may be, **alternative language models**. Instead, this other model can be selected by adding the trigger **"(2nd)"** if it has been configured. The prompt then goes to the secondary language model. You can see which one this is by going to Settings (see para. 26 et seq. above) or by moving the mouse over the Red Ink logo. It is also possible to switch the two models in Settings. If alternative models are configured, the user can select which one to use.

175 If you want to change the settings of Red Ink in Outlook, you can do so temporarily or permanently using the **Settings** function (for the individual values, see para. 26 et seq. above). If the settings are simply changed, they will only be retained for as long as Outlook is not closed; after that, the pre-configured settings are applied again. If you don't want this, you can save the settings in a local copy of the configuration file in Settings. This is done automatically as soon as the configuration is saved in Settings. In this case, the "local" redink.ini file (in the Outlook directory and for Outlook only) is overwritten (but not a possible central file, as it can be provided via the registry, see para. 257 below); the local file is given priority when reading. It is also possible to access the other configuration parameters via an "Expert" window, but we do not recommend this (it is better to make adjustments directly in the configuration file). If no configuration file is found for the add-in for Outlook (or Excel), the add-in will look for one for Word. Therefore, if you want to use a separate configuration for Outlook, it is best to save a separate configuration file for Outlook (in Settings or manually).

## **N. Browser extension**

176 Red Ink also has an extension for Chromium-based browsers (e.g., Edge, Chrome). Once installed, users can select text in their browser (e.g. when working on a website) and have Red Ink perform certain actions on the selected text:





177 The commands are:

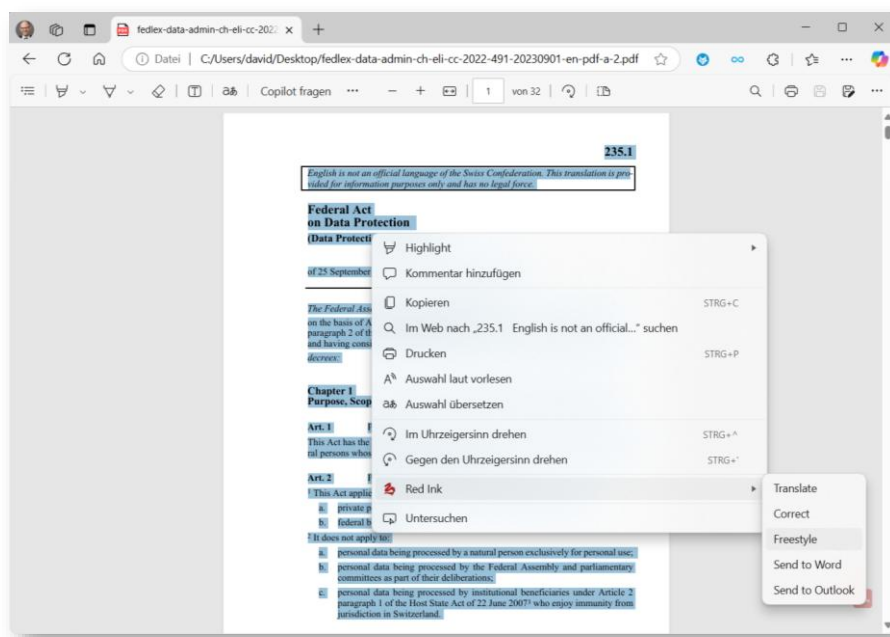
- **Translate:** A window opens and the user enters the desired output language. The translated text is displayed in place of the selected text.
- **Correct:** The text is linguistically corrected and a markup is displayed in a window. If the user does not press "Esc", the selected text is replaced by the corrected text (without markup).
- **Freestyle:** You can enter any prompt (the prompt library is also available). However, formatting commands and functions such as "bubbles" does not work here, of course. By default, the AI's response is displayed in a separate window (and placed on the clipboard), just like when the prefix "Clipboard:" is used in Word. However, if you want Red Ink to send the generated output back to the browser for insertion, you must prefix the prompt with "Insert:". Markups are also possible, similar to Correct. For this, "Markup:" must be prepended; this prefix also includes an "Insert:" command, i.e., the content is sent back to the browser, unless cancelled beforehand. Freestyle is the only command that can be used even if no text is selected in the browser.
- **Send to Word:** The text selected in the browser is inserted at the current position in the document in Word (without copy and paste).
- **Send to Outlook:** The text selected in the browser is inserted at the current position in the document in Outlook, provided that a window for composing an email is open (without copy and paste).

178 The way the browser extension works is that the installed software simply sends the selected text in the background to the Red Ink add-in for Outlook (or to the add-in for Word if you use the "Send to Word" function), and the text is processed there. Outlook must therefore be

"running" with the add-in for the browser extension to work. If it is not, nothing happens. Depending on the situation, however, it may be that the window that opens in Outlook when using, for example, Freestyle or Translate is not noticed at first glance (Red Ink is programmed to push itself into the foreground, but this may not always work).

179 After processing, the response from Red Ink is sent back to the browser extension, which then replaces the selected text (if you don't want this to happen, you have to press "Esc" first, which opens a dialogue within Red Ink). However, depending on how the page displayed in the browser is programmed, the returned text may be inserted in the wrong place (e.g. above the input field). Certain pages block the display of the Red Ink context menu (e.g. when working in Google Docs). Note: The browser is not blocked while Red Ink is working. If you continue working in the browser in the meantime, it may happen that the text returned by Red Ink is inserted in the wrong place.

180 In everyday use, the browser extension has proven itself, above all in the **analysis of website content**, but also **of PDF documents**. Here, the relevant web page is simply to be selected partially or completely (e.g. using Ctrl-A) and then the Freestyle command is selected. When the Freestyle window pops up, the question can be entered. If a PDF document is to be analyzed, it should be opened in the browser instead of the PDF reader (provided that no PDF plugin is used in the browser that interferes). There, too, the entire text can then be selected and queried using Freestyle (e.g. "Where is the question of consent dealt with?"):

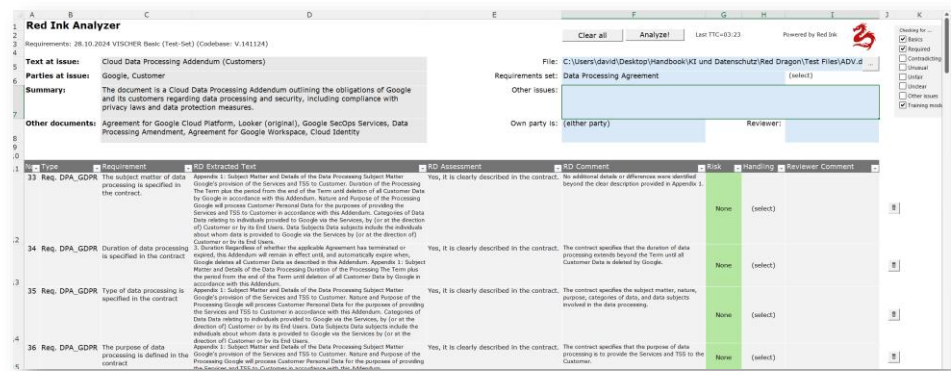


## O. Using Red Ink in other programs

181 The Excel add-in can, with its helper program, also be used by other applications to access language models. This makes it very easy to

build solutions in Excel that use a language model, because the user programming them no longer has to worry about the interface. All this is done by Red Ink, which runs in the background whenever Excel is used. Furthermore, the add-in with the helper also offers the option of reading the contents of PDF files (like the corresponding Word helper). This can be useful because Excel itself does not offer this option in VBA.

- 182 One example of such an application is the **Red Ink Analyzer**, a tool that can be used to analyse legal texts (e.g. contracts or data protection declarations) for predefined requirements or other problems, and to evaluate documents systematically using AI (e.g. to create summaries of evidence in a case, to extract specific information from a series of documents or to have documents from an internal investigation checked for certain suspicious content). We offer this tool for commercial use for a moderate licence fee.



- 183 To access the LLM interface from an Excel application, the add-in for Excel and the helper must be installed and loaded (for the helper, see para. 228 below).

- 184 The **LLM interface** can be used by other VBA modules in Excel or directly from an Excel cell. The command has the following syntax:

Answer = LLM(SysPrompt, UserPrompt, Model, Temperature, Timeout, SecondAPI, Hidesplash)

- 185 The parameters are as follows:

Key	Type	Description
Answer	String	The output of the language model, with escape characters already removed.
SysPrompt	String	The system prompt, where the prompt is cleaned for JSON before use.
UserPrompt	String	The user prompt, where the prompt is cleaned for JSON before use.
Model	String, optional	Model name (if the information is required for the API or supported by it); if the default value is to be used, then ' ' should be entered.
Temperature	String, optional	Temperature (if the information is re-

		quired for the API or supported by it); if the default value is to be used, then '' should be entered.
Timeout	Long, optional	Timeout in milliseconds; if the default value is to be used, then enter 0.
SecondAPI	Boolean, optional	True if the optionally configured secondary language model is to be used for the query, otherwise False; default value is False.
Hidesplash	Boolean, optional	True if the splash window with the Red Ink logo should not be displayed during a query (appears with larger queries); default value is False.

186 In Excel, the above function can be called up as follows from another module, for example:

```
result = Application.Run('redink_helper.xlam!LLM', SysPrompt,
UserPrompt)
```

Or:

```
result = Application.Run('redink_helper.xlam!LLM', SysPrompt,
UserPrompt, "", "", 0, False, False)
```

187 If you want to test the LLM interface of the helper, you can run the procedure "TestLLM". After a short time, a window should appear with an answer from the respective language module.

188 The function for reading the **text from files** has the following syntax:

```
Answer = GetFileTextContent(Filename, ErrorInAnswer)
```

189 The parameters are as follows:

Key	Type	Description
Answer	String	The text content of the file (or the error code, starting with "Error").
Filename	String	The full file name with path, with environment variables automatically replaced.
ErrorInAnswer	Boolean, optional	True if the function should return the text "Error" plus a description of the error in the event of an error, otherwise it returns an empty string.

190 In a module, the function is called up in the same way as in the above example for the LLM function.

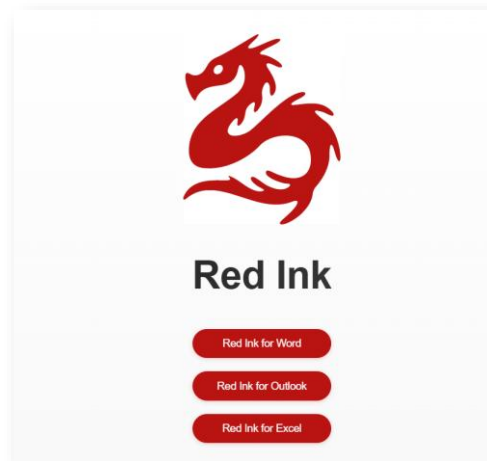
191 Finally, the **Adjust Cell Height** function is also available, which automatically adjusts the height of the currently selected cells to the text content, even if they are connected (Excel cannot do this itself). It is called up using the following procedure and in a module analogous to the example above for the LLM function:

```
AdjustHeight()
```

### III. INSTALLATION

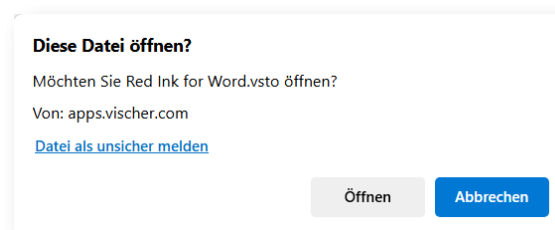
#### A. For those who can't wait: The one-click installation

192 The quickest and easiest way to install Red Ink is via the website <https://apps.vischer.com/>, where the respective add-in can be installed with a single click on the corresponding button:



193 Caution: This type of one-click installation currently **only works with the Edge browser**, but not with Chrome, Firefox or any other browser.

194 The browser will typically warn you because it doesn't recognize this program. You must click "Open" in this example to run the installation:



195 The installation requires no further input. Once it is finished, Word, Excel or Outlook must be restarted and Red Ink is available. At the first start, the minimum information is configured, i.e. the secret access code (usually the so-called API key) to the language model of choice must be entered. Those who have a subscription to "ChatGPT" can get it from OpenAI. It has to be inserted in the form. This is usually sufficient and you can work with Red Ink. Red Ink can be further configured via the Settings function. More on this is described in the following paragraph of section in detail, as is the installation of the optional helper files. Updates can be installed by clicking the buttons again but will also be proposed automatically after a certain time. Those who want to install Red Ink in the company may need appropriate authorization, because the security filters may block an installation.

## B. The installation in more detail

196 There are three ways to install Red Ink:

a) **Method 1:** Installation directly via an installer on the Internet

This is done as described above by using our deployment server <https://apps.vischer.com> and is the easiest installation method. It will not install the two helper files, should they be desired; this is possible later on, though. The disadvantage of this method is that it may be blocked by security settings, especially in organisations. This method is therefore more suitable for **private individuals** and **small businesses**.

b) **Method 2:** Download of the installation package and complete installation carried out locally

The files and two helper files (if desired) are first downloaded and installed from a local source or copied to specific directories (which is also possible later). This method is more suitable for **larger organisations** that provide Red Ink to their employees but do not want to give them access to the files as in method 1.

c) **Method 3:** Installation via image or software distribution

This is intended for organisations that want to control Red Ink even more closely or distribute it in a pre-installed form. We will not go into this in more detail here. However, Red Ink also has functions so that it can be used in **organisations with numerous users** (see, among others, para. 246 et seq. below).

197 In particular, access to the installation files from both method 1 and method 2 may be **blocked for security reasons**. In such cases, only the administrator of the IT environment can help and release the files (e.g. based on their digital signature from us, VISCHER AG)

198 It is also possible that **Windows Defender** (or another antivirus program) may erroneously identify individual installation files (in particular the Excel helper) as a threat (specifically, the 2020 Trojan "O97M/Sadoca.C!ml") and move them to quarantine. The same applies to the helper files. Depending on the Windows security settings, such false alarms can be overridden (and the files can be put on a "white list"). Sometimes it also helps to restart the computer and wait a little. We have observed such false alarms mainly when Red Ink is installed multiple times in succession. We have reported the problem to Microsoft.

199 When Red Ink is first installed, a **minimal configuration** is carried out. Otherwise, the default values are used. It is also possible to configure Red Ink in a much more differentiated way by manually editing the configuration file. Red Ink can also be automatically distributed in organisations and has functions so that it can be used for several users at the same time.

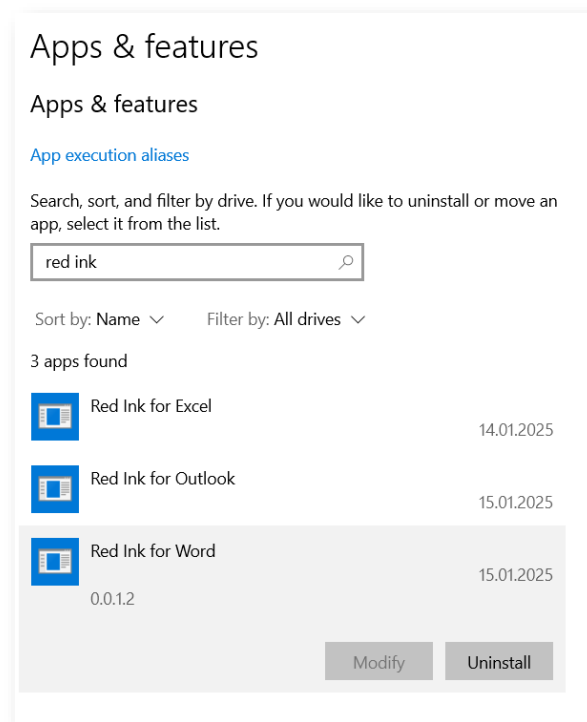
200 For installing the files of the **Transcriptor** (speech models and, as the case may be, program libraries) see para. 83 et seq. above.

201 Red Ink also has an **update function**. It depends on whether method 1 or method 2 has been used:

- For method 1, the update is done by clicking the installation links or by using the built-in update function (in the Settings menu). The add-ins are also configured to check for updates and install it every three days (this can be changed via the configuration file). Updates are performed by the add-ins either by internally calling <https://apps.vischer.com> (the user will simply have to confirm the update, but does not shown more) or by calling the Microsoft Edge browser and the VSTO installer stored on <https://apps.vischer.com/> (i.e., the same happens as clicking the red buttons on the same page). The execution of the installer must be confirmed manually in the browser.
- For method 2, the update is done by downloading a new version of the installation package and copying it over the existing installer directories ("word", "excel" and "outlook"). After that, the installation can be repeated as for the initial installation, or the built-in update function (in the Settings menu) can be used, which basically does the same thing. The add-ins are also configured to check for updates every three days (this can be changed in the configuration file). However, updates from the add-ins directly require, firstly, that the update path is stored in the configuration file ("UpdatePath = "). Secondly, this update path must be the same as the one from which the add-ins were originally installed. Otherwise, the user must first uninstall their previous add-in before they can install the new version.

The updates are possible without terminating Word, Excel and Outlook, but will be effective only after a restart.

202 Currently, the add-in cannot check whether an update was successful; it can only initiate the process. Success can be seen from the installer, which pops up a window, or it can be checked manually. The Windows menu "Add or Remove Programs" shows each add-in, along with a version number of the program files (here: 0.0.1.2). After an update, this number must be higher:



- 203 If in doubt, uninstall and reinstall the add-in. The configuration is retained.
- 204 Red Ink is **uninstalled** via the Windows function described above (press the "Uninstall" button in the screenshot above). The helper files can either be deleted via the settings function or manually and are thus uninstalled.
- 205 We recommend that everyone also signs up for the Red Ink **mailing list** at <https://vischer.com/redink> to stay up to date on our tools.

### C. Preparation: API access

- 206 To use Red Ink, you need access to a suitable language model from a new generation (such as "gpt-4o" from OpenAI or Microsoft or "Gemini 1.5 Pro" from Google). Normal access to "ChatGPT" or "Copilot" is not enough. What is needed is a so-called API access. API stands for "Application Programming Interface" and in this case refers to an interface that is accessible via the internet or a local network and to which Red Ink (or other software) can send requests for the language model. The technical term "endpoint" is also sometimes used.
- 207 Getting API access to a language model is not difficult. Many organisations already have one in operation for other applications (e.g. via the "Azure OpenAI Services" if they use Microsoft's online services) and can also use it for Red Ink. If you don't have such an API access, you can subscribe to one for a very small fee, for example from OpenAI (<https://openai.com/api/>) or Google (<https://ai.google.dev/gemini-api/docs/api-key>). If you have a "ChatGPT" account, you can do this via that account. In contrast to services such as "ChatGPT" or "Copilot



for M365", the use of a language model via API is usually paid for based on usage, although experience shows that the costs for many users will be lower than with a subscription to one of the chat services. If you want to know more, ask a good AI chatbot for help.

208 Here is our personal assessment of the most popular providers:

<b>OpenAI</b>	+ Very easy to subscribe and set up
Model family: GPT	- Not for professional and official secrets
	- LLM server not in Switzerland
<hr/>	
<b>Microsoft (Azure)</b>	+ Many companies already have it
Model family: GPT	+ LLM server also in Switzerland
	+ Supplemental agreement for professional and official secrets available
	- Subscribing and setting up is somewhat more complex
	- Switching off abuse monitoring only possible for large customers (relevant for professional and official secrets)
	- Important model GPT-4o not in Switzerland
<hr/>	
<b>Google (Vertex)</b>	+ Supplemental agreement for professional and official secrets available
Model family: Gemini	+ LLM server also in Switzerland
	+ Switching off abuse monitoring also possible for medium-sized and special customers (relevant for professional and official secrets)
	- Subscribing and setting up is somewhat more complex
	- Models are less popular

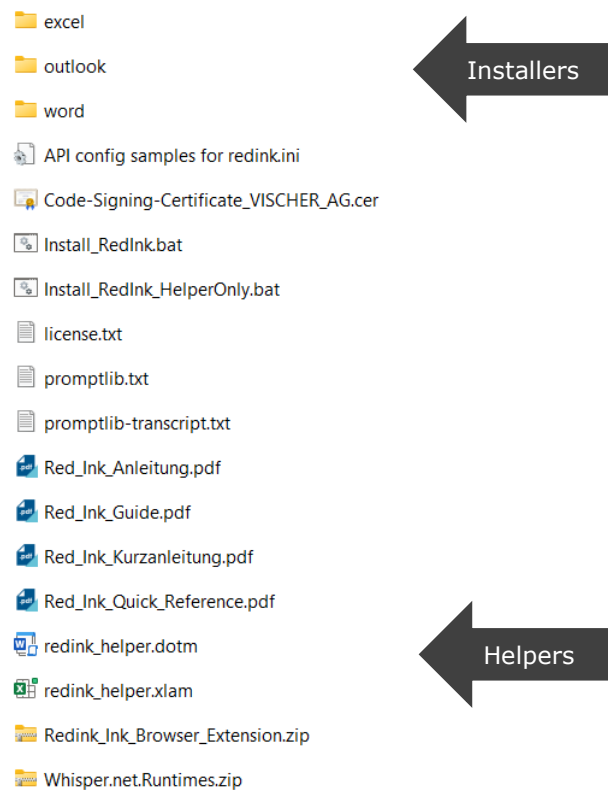
209 Anyone who wants to use Red Ink to process personal data or even data subject to professional or official confidentiality should obtain API access that meets the relevant requirements. For example, we as a law firm use access via Vertex API from Google with a special contract. We are happy to advise on (Swiss) legal issues in this regard.

210 If you have API access, you can have an appropriate "API key" generated, i.e. a secret character sequence that serves as an access key and must be stored in Red Ink so that the tool's queries are answered by the API. Google Vertex and certain other providers require a special authentication procedure, where further information is needed. The API key or the additional information should be readily accessible for the first use of Red Ink.

#### **D. Step 1: Download the installer/installations package**

211 For **method 1**, the steps are described above (para. 192 et seq.).

- 212 For **method 2**, the latest version of the installation package "red-ink.zip" can be downloaded from the same address <https://apps.vischer.com>. It is labelled as a "Local Installation Package". It is a ZIP file. Its contents must be unzipped into a temporary directory (e.g. on the desktop or in a newly created directory for "Red Ink") (but see para. 216 below). It contains the following files, among others:



- 213 The three directories contain the installers for the three add-ins. These are the minimum requirements. The rest is optional.
- 214 The installation package also usually includes:
- This guide (in German and English)
  - A quick reference (in German and English)
  - The VISCHER AG digital code signing certificate
  - A sample prompt library
  - A sample prompt library for the Transcriptor
  - The optional "Install\_RedInk.bat" and "Install\_RedInk\_HelperOnly.bat" batch files to simplify installation
  - A file with the licence details (also with details of the licences of the program libraries used)
  - The two optional Red Ink helpers for Word and Excel

- A file with sample entries for the API configuration for OpenAI, Azure and Google Vertex
- A file with sample entries for configuring various "Special Services"
- The licence terms

215 The **batch file** (".bat") can be used to perform the following steps 2 and 4 with a single click. It copies the two helper files as specified in step 4 and then runs the three installers as specified in step 2. If you want, you can click on these and, if everything works, you will have completed steps 2 and 4. You will first be asked whether you want to continue. If an error occurs, it will be displayed (e.g. if Word is still running while the helper is trying to copy, it may block). If necessary, Windows will ask whether the batch files are trusted. This is to be confirmed accordingly.

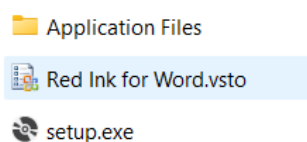
216 To enable updates from future versions of the installation package, the installation must always be carried out from the same file path. It is therefore worthwhile saving the files from the installation package to a fixed directory created for Red Ink on the local computer or on the organisation network, from where the installation is then carried out.

217 **Note for users of method 1:** You can also download the installation package to access the additional files or you can download some of the files (e.g. the prompt library sample) directly from <https://apps.vischer.com>.

#### E. Step 2: Run the installer

218 In **method 1**, the installer is already executed in step 1 with the confirmation that you trust the file. There is nothing more to be done here.

219 In **method 2**, there is a file with the extension ".vsto" in each of the directories "word", "excel" and "outlook". This is the add-in. It will look something like this:



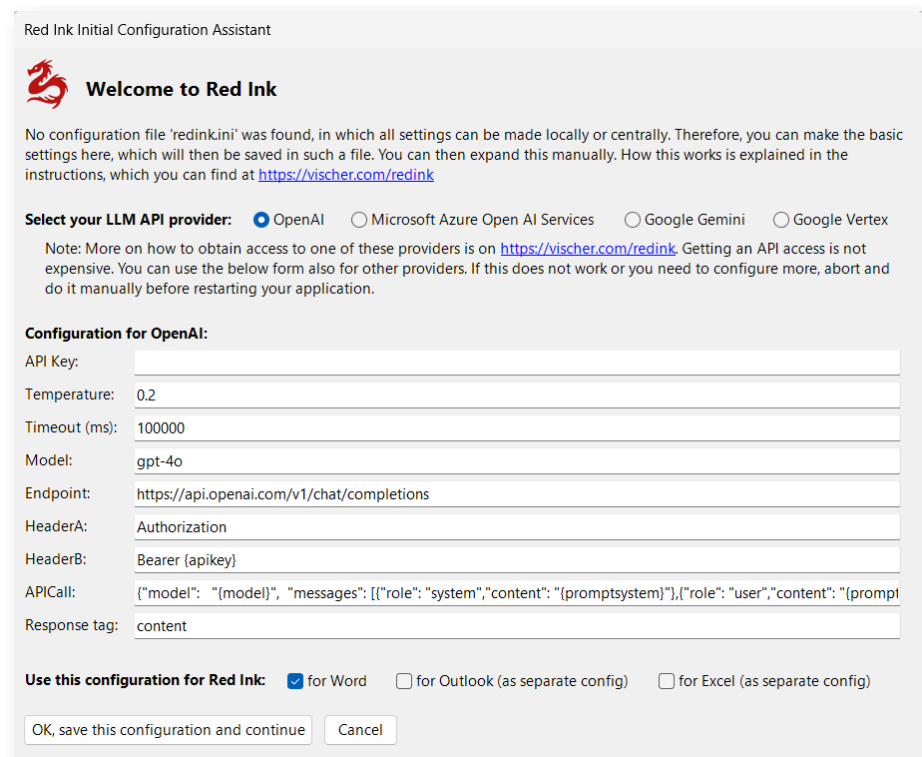
220 To install, the ".vsto" file should be executed ("setup.exe" also works, but is more likely to be blocked by security settings). The installation should be complete after a few seconds. Nothing needs to be entered.

221 The following applies to **both methods**: If an earlier version of the add-in is already installed on your computer, the installation will not work if it has been installed from a different source or in a different way. In this case, you must first uninstall the earlier add-in. This is also relatively easy to do in Windows using "Add or Remove Programs" (see para. 202 above).


222 The configuration file is not overwritten during installation and is not removed during uninstallation.

### F. Step 3: Initial configuration using the wizard

223 When Red Ink is started for the first time and there is no configuration file yet, the following window appears, which can be used to create a minimal configuration (it may look slightly different in the current version; in Word it appears only when a document is opened or created for the first time):



Red Ink Initial Configuration Assistant

 **Welcome to Red Ink**

No configuration file 'redinkini' was found, in which all settings can be made locally or centrally. Therefore, you can make the basic settings here, which will then be saved in such a file. You can then expand this manually. How this works is explained in the instructions, which you can find at <https://vischer.com/redink>

**Select your LLM API provider:** ☒ OpenAI ☐ Microsoft Azure Open AI Services ☐ Google Gemini ☐ Google Vertex

Note: More on how to obtain access to one of these providers is on <https://vischer.com/redink>. Getting an API access is not expensive. You can use the below form also for other providers. If this does not work or you need to configure more, abort and do it manually before restarting your application.

**Configuration for OpenAI:**

API Key:

Temperature:

Timeout (ms):

Model:

Endpoint:

HeaderA:

HeaderB:

APICall:

Response tag:

**Use this configuration for Red Ink:** ☒ for Word ☐ for Outlook (as separate config) ☐ for Excel (as separate config)

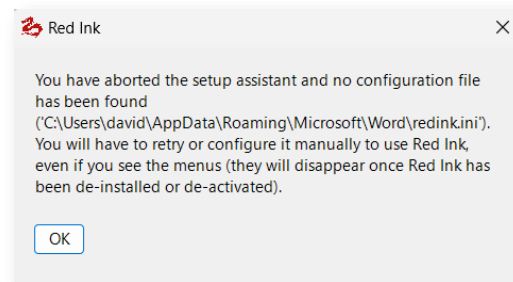
224 First, select one of the providers. The most common providers have been pre-programmed with their typical information, as far as this is generally valid. However, we cannot constantly update this information, which means that it may not match the latest specifications. If you get stuck here, you can also ask a good AI chatbot for help. The use of API access is not intended for end users and may therefore be somewhat complicated. If your provider is not listed, the fields can also be filled out with information from a different provider (selection via the radio buttons is only relevant for the preselection of the values, it is not used further). The content of the parameters corresponds to the parameters of the configuration file (these are described in detail in para. 246 et seq. below).

225 For example, if you have an API key from OpenAI, you just need to enter it in the relevant field and you're ready to go.

226 The information must be saved before it can be used. This involves writing a local configuration file for Red Ink. You can choose whether it

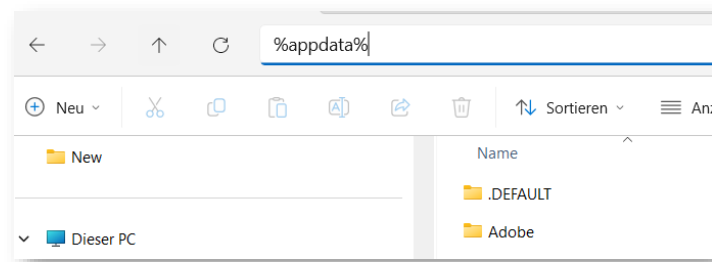
should be written only for the current add-in or also for the add-ins of the other two Office products.

- 227 If this is not done or is cancelled, it is not a problem. The Office application can still be used, but the wizard will appear every time until the configuration is set, the configuration file is stored in the relevant directory, or until the add-in is uninstalled (using "Remove Programs..." in Windows) or disabled (within the Office product). After cancellation, the following error message also appears:



#### G. Step 4: Install helper (optional, can be done later)

- 228 The helpers are not necessary, but anyone who wants to use the context menu in Word and Excel for Red Ink or the keyboard shortcuts, or who wants to use certain functions from Excel for their own Excel applications (see para. [181](#) above), needs them. These are digitally signed programme files for Word and Excel that contain macros in Visual Basic for Applications (VBA).
- 229 There are **two methods** to install the helpers. The easiest method is the installation via the settings function in Word and Excel. There, a button "**Install Helper**" appears. If it is pressed, the latest helper version is downloaded and copied to the respective directory. The disadvantage of this method is that it can be blocked by security functions. If this is the case, the helper must be installed manually. This is described in the following.
- 230 For manual installation, the helpers are also included in the installation package. To install them, simply copy them into the appropriate Office directory:
- 231 The Word helper "redink\_helper.dotm" is entered here:  
C:\Users\vorname.name\AppData\Roaming\Microsoft\Word\Startup
- 232 The Excel- helper "redink\_helper.xlam" is entered here:  
C:\Users\vorname.name\AppData\Roaming\Microsoft\Excel\XLSTART
- 233 Replace the red text "vorname.name" with your own username. The easiest way to display the directory is to enter the characters "%appdata%" in Windows Explorer and confirm with the Enter key:



A directory will automatically open giving access to the data of all applications. Select "Microsoft", then Word or Excel and the above directories "Startup" or "XLSTART". It may be that one or both of these directories are missing; in this case, simply create them (these are the directories in which VBA files are stored so that they are executed when Word or Excel is started).

234 **If you don't want to do this manually**, you can run the batch file "Install\_RedInk\_HelperOnly.bat" from the installation package. It will copy both files to the correct location.

235 The respective files are automatically loaded and activated when Word and Excel are started. It can be uninstalled by deleting it. However, it is possible that the local security settings may block the execution of the files, even though the data has been digitally signed by VISCHER. In most organisations, this will be the case. If this happens, the security settings must be supplemented with a corresponding individual approval (simplest for program files that are digitally signed by us). If necessary, the user must also indicate in Word and Excel that the content must be "activated".

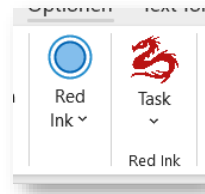
236 The installation of the helpers can also be done later. However, the context menu will not be visible until then. The add-ins each check whether the corresponding helper program is running and then activate it automatically if the context menu is not globally disabled via the configuration file.

## H. Step 5: Using add-ins

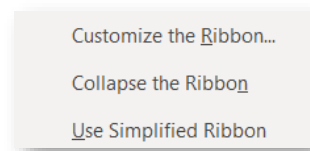
237 The add-ins are now ready to use. In Word and Excel, tiles should be visible immediately when a document or worksheet is open. In Outlook, tiles only appear when a window for composing an email (new email, reply, forward) is open.

238 The **Annex** contains some suggestions on how you can get to know Red Ink better.

239 If there are too many tiles on the ribbon (measured by the width of the window), the Office applications try to condense them. It looks like this:



- 240 This cannot be prevented by the programme. However, each user can remove tiles that are not needed by customising the ribbon and thus create space so that it does not condense (access "Customise Ribbon..." by right-clicking on the ribbon); the position of the tiles can also be defined there (the further to the left they are placed, the less likely it is that they will be minimised):



- 241 In Outlook, the add-in deliberately places its tiles far to the left. If Red Ink is used in a organisation, however, it may be that the position is reset the next time the application is started because the global settings require this.

#### **I. Step 6: Making further adjustments as necessary**

- 242 Red Ink can be configured and customised in a variety of ways. This goes so far that even the internally used prompts can be changed. Some of these settings are available through the Settings function; however, only settings that "normal" users typically need are available there. The other settings can be accessed via the "Expert Config" or even more simply via the "redink.ini" configuration file. This is a simple text file in which all configurations are stored in plain text. Everything else is described in detail in the next chapter, including the location of the configuration file.

#### **J. Installation of the browser extension**

- 243 The browser extension is supported by Chromium-based browsers such as Edge or Chrome. It can be installed either via the Edge and Chrome add-on store or manually. To install them automatically, open Edge or Chrome and go to the respective add-on store:

Edge: <https://microsoftedge.microsoft.com/addons/detail/red-ink-browser-extension/dflpmohocianaolidmcmphfcdpcognni>

Chrome: <https://chromewebstore.google.com/detail/red-ink-browser-extension/doagmfoemngdlbghobfkbobehbodgdoa>

or go to <https://apps.vischer.com>:



- 244 For the manual installation, the files are included in the installation package:

- manifest.json
- Readme.txt
- redink\_background.js
- redink\_icon16.png
- redink\_icon64.png
- redink\_icon128.png

- 245 The files must be copied to a permanent directory (except Readme.txt) and can then be read in the browser, for example in Edge on the internal page "edge://extensions" (or "chrome://extensions" for Chrome), where the developer mode must be activated for this. The extension communicates with Red Ink via a local http connection, using ports 12333 and 12334 from IP address 127.0.0.1. However, in some organisations this communication will be blocked for security reasons.

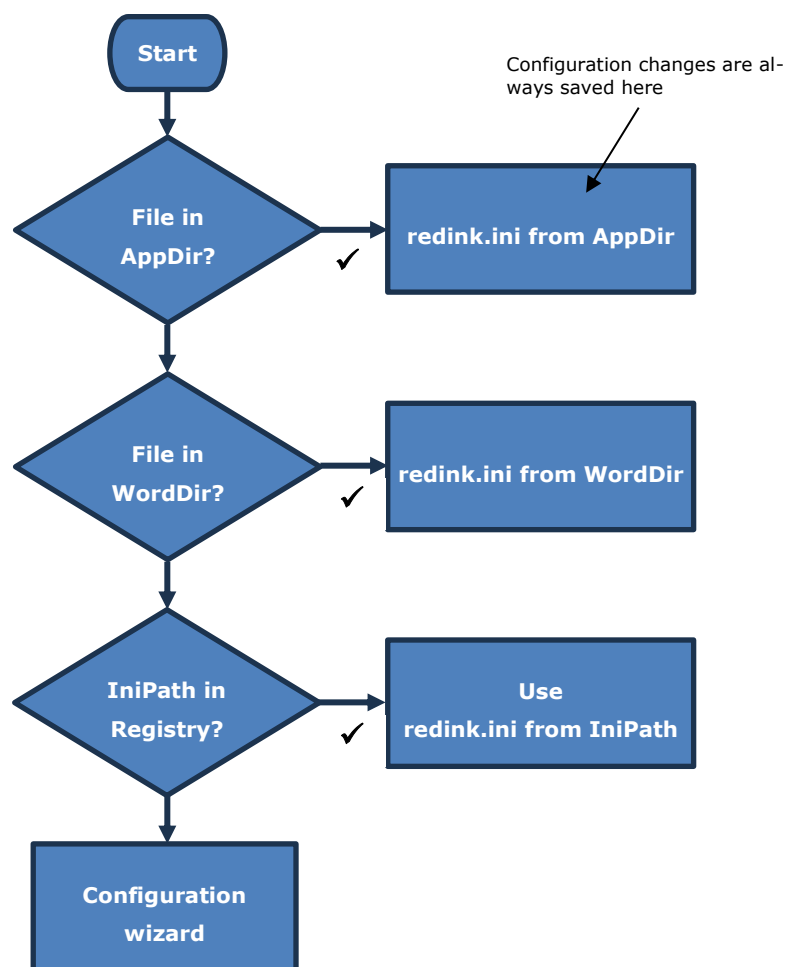
#### IV. CONFIGURATION (FOR ADVANCED USERS)

##### A. Configuration file "redink.ini"

- 246 Red Ink is configured using the parameter file "redink.ini". This is a text file that can be edited with any text editor, in which the parameters for operating Red Ink can be manually adjusted. If it is missing, the respective add-in starts a wizard that is used to enter the minimum parameters needed for use and otherwise applies the default parameters (see para. 223 et seq.). A "redink.ini" file is then automatically generated. For a more specific configuration or for use in an organisation, it is recommended that the configuration file be prepared manually and copied into the relevant directory or made available centrally in a directory that is communicated to all workstations via the registry before the add-in is started. This is also how Red Ink can be distributed in a network.



- 247 Red Ink is flexible about the location of the configuration file. The following concept applies:



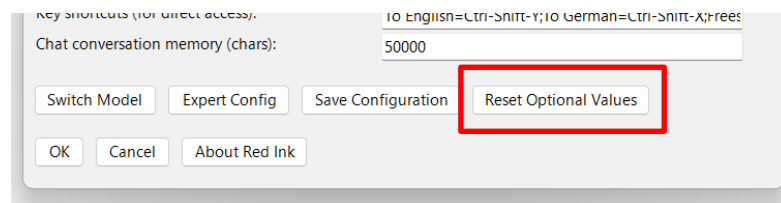
The directories are located in the following places:

Directory	Place
AppDir from Word	%AppData%\Microsoft\Word\
AppDir from Excel	%AppData%\Microsoft\Excel\
AppDir from Outlook	%AppData%\Microsoft\Outlook\
WordDir	AppDir from Word
IniPath	The path stored in the registry under the key "IniPath" (see para 257)

- 248 The priority of the configuration file in the "local" AppDir directory is hard-coded in the add-ins, but can easily be changed in the source code using a constant (requires the applications to be recompiled). In this case, changes are still written to the local AppDir, but are ignored.
- 249 The "redink.ini" file is imported every time it is used. The parameters used are the same for all add-ins, but certain parameters are only used in one or two add-ins. However, they do not affect the other add-ins. By storing the configuration file in separate locations, it is possible to

configure the three add-ins differently. You can also share a common local configuration file (in WordDir). However, as soon as a configuration is saved in Outlook or Excel (which any user can do), a separate local configuration file is stored and used.

- 250 If the user adjusts their own configuration settings via Settings and saves them, these changes are saved in a local copy of the configuration file, which is imported according to the priority set out in the above scheme. If the user wishes to discard these changes and return to the central configuration file, they can do so by clicking a button at the bottom of the Settings dialogue. If no central configuration file is available via the registry, the user can instead reset the optional settings they have configured to the default values (the API access data is not affected):



- 251 The parameters of the configuration file "redink.ini" are as follows (the key must be written exactly as shown, it is case sensitive):

Key		Example value	Meaning
APIKey	=	sk-proj-XXXX	The secret API key for accessing the LLM API, either in plain text or encrypted; if the OAuth 2.0 procedure is used, the private key of the service account used for Red Ink is stored here, (in plain text or encrypted), namely the naked Base64 part (i.e. without the pre- and suffix texts of the PEM format such as "-----BEGIN PRIVATE KEY-----\n"), whereby paragraph marks ("\\n") are ignored or filtered out before use
APIKeyEncrypted	=	Yes	"Yes" if the API key (or the private key for OAuth 2.0) is encrypted, otherwise "No"; for encryption, see section C
APIKeyPrefix	=	sk-proj-	The prefix that occurs with every API key of the respective provider (if one is used; this is for security reasons, because the prefix is not encoded during encryption). When using OAuth 2.0, the prefix is ignored
Endpoint	=	https://api.openai.com/v1/chat/completions	The URL to which the API query is sent using the POST method; if https is used, the transmission is encrypted; if the URL is prefixed with "GET:", the query is made using the GET method; placeholders can also be used in

			the endpoint parameter, and it is possible to configure a double query, i.e., a POST command is executed first and in a second pass (with results of the first command) a second GET query; the two endpoints are then separated by the character " "; this is especially relevant for Special Services; details on this configuration and the use of placeholders are described in para. 275 et seq.
HeaderA	=	Authorization	The first part of the http header; the API key can be used as a placeholder (as shown in the next example); it's the field name of the header (without ":"); it can be empty
HeaderB	=	Bearer {apikey}	The second part of the http header; the API key can be used as a placeholder (as indicated); it's the value of the field listed in HeaderA; it can be empty, if Header A is empty, too
Response	=	content	The JSON segment that marks the LLM response in the API response; if needed (esp. when using "Special Service"), templates for more complex analyses can be defined here; details are described in para. 275 et seq.; if an access requires two endpoint calls, two response templates, separated by " ", can be entered here; this must then be coordinated with the "APICall" and "Endpoint" parameters; if "Response = JSON" is specified, the model's response is output as such (for debugging purposes); regardless of the "Response" parameter, Red Ink extracts binary image objects (and saves them) and source citations (i.e., hyperlinks, which are then appended) from the language model's response
APICall	=	{'model': '{model}', 'messages': [{ 'role': 'system', 'content': '{promptsystem}' }, { 'role': 'user', 'content': '{promptuser}' } ], 'temperature': {temperature}}	The syntax in which the LLM API transmits the query, using curly brackets to indicate the corresponding placeholders for the system prompt, user prompt, model and temperature; if no distinction is made between system and user prompts, both should be specified one after the other; besides {promptuser} and {promptsystem}, the placeholder {userinstruction} can also be used, which contains the prompt that the user enters in Freestyle (which, depending on the model, can be practical because it can be placed in the user prompt in

			<p>this way, even though it is normally in the system prompt); The placeholder {objectcall} is finally used to record the call sequence for transmitting file content (see "APICall_Object"); if required (especially when using "Special Service"), two API call strings can be configured here (separated by " ", which results in a POST command being executed first, then a GET command; details are described in para. 275 et seq.; the parameters "Endpoint" and "Response" must then also be extended)</p>
APICall_Object	=	<pre>, {"inlineData": {"mimeType": "{mime- type}","data": "{encodeddata}"}}}</pre> <p>or</p> <pre>multi- part:model:{model };prompt:{prompts ystem} {promptus- er};filefield:image[]</pre>	<p>The sequence, which is integrated at the correct location in the API call (where "APICall" contains the {objectdata} placeholder) and serves to transmit the content of a file to the model; the placeholder {mimetype} is automatically replaced by the add-in with the MIME type (e.g., "image/jpeg"), the placeholder {encodeddata} is automatically populated with the Base64 code of the file; as an alternative to JSON encoding, multipart encoding is also supported; in this case APICall is not needed (it can contain any value), and APICall_Object must begin with "multipart:", followed by the fields separated by ; and after a colon their placeholders (if ";" is needed, ";" is the escape character); only if this parameter is defined is the trigger "(file)" available in Freestyle</p>
Timeout	=	200000	The timeout for a request to the API in milliseconds
Temperature	=	0.2	The temperature, as far as it can be indicated (inserted above)
Model	=	gpt-4o	The model name (inserted above, if necessary also in the endpoint if there is a placeholder there)
MaxOutputToken	=	8192	This is where you can set the maximum number of tokens an output from the language model can have. The add-in will warn the user in Word and Outlook if a text is to be edited that probably exceeds this number of tokens. This may result in the output being truncated by the language model, which in turn may mean that it is not complete; most models can generate much less output than they can take in input; when set to 0, no checking is performed

VISCHER

Anon	=	askshow; 4	This parameter configures the integrated anonymization function for the model in question; the mode (none, silent, ask, askshow, show) and the type (0-4) of anonymization must be specified (see details in para. 110 et seq.); if the parameter is not set, no anonymization takes place; the user can override this parameter with a local file named redink-anon.txt on their desktop
TokenCount	=	candidatesTokenCount; thoughtsTokenCount; 5; CHF	Information to configure the Freestyle function, which logs the tokens used in each case and multiplies them by a currency amount; first, the JSON segments containing the desired token values (one or more) are to be listed, then optionally the multiplier, and then the specification of the currency or another designation, which then appears in the "redink-cost.txt" file on the desktop
OAuth2	=	Yes	'Yes' if the OAuth 2.0 procedure is to be used for authentication instead of a normal API key; in this case, the "APIKey" contains the private key of the service account used for Red Ink, which is used to request the necessary access token; the other keys must then subsequently be configured
OAuth2ClientMail	=	red-ink@earnest-racecars-212313-x4.iam.gserviceaccount.com	The 'client email' parameter, which contains the email address of the service account used for Red Ink and which must be sent to the authentication server using the OAuth 2.0 process to request an access token
OAuth2Scopes	=	https://www.googleapis.com/auth/cloud-platform	The "scopes" parameter that must be sent to the authentication server after the OAuth 2.0 procedure to request an access token; it specifies the resources for which the access token is requested.
OAuth2Endpoint	=	https://oauth2.googleapis.com/token	The address of the OAuth 2.0 server that performs the authentication
OAuth2ATExpiry	=	3600	The lifetime of an access token in seconds (a new access token is requested shortly before the token expires); if not configured, the value 3600 is assumed
DoubleS	=	Yes	If the "sharp S" is to be replaced by a double S by default; the default is "No"
Clean	=	No	If set, invisible spaces and double spaces in the LLM's output

VISCHER

			are automatically removed; such codes can be used as watermarks – this command then removes them
PreCorrection	=	As an additional instruction but only if you generate text in German language, replace any appearance of 'personenbezogene Daten' with 'Personendaten'	An optional command that is given with each API query (in English), e.g. to make standard corrections; this command is given with the first LLM query.
PostCorrection	=	All references in the text provided to you for processing that refer to Vischer have to be in all-caps, like VISCHER.	An optional command that is automatically executed on the response after each API query; if this command is entered, there will be twice as many queries, which will result in higher costs and take more time; therefore, "PreCorrection" is preferable ; here the complete (system) prompt is to be entered, while the text to be processed is passed to the AI as a (user) prompt enclosed in the two tags "<TEXTTOPROCESS>" and "</TEXTTOPROCESS>".
APIDebug	=	False	If set to True, the full response of the LLM (or Special Service) will be written into a ASCII file on the user's desktop (it will always be overwritten); default is False
Language1	=	English	This is the first standard language in which the translation function of Red Ink can be accessed directly. The text should be in English; if nothing is specified, this is "English"
Language2	=	German	This is the first standard language in which the translation function of Red Ink can be accessed directly. The text should be in English; if nothing is specified, this is "German"
ShortcutsWordExcel	=	To English=Ctrl-Shift-E;To German=Ctrl-Shift-D;Freestyle=Ctrl-Shift-P;Self-Compare Selection=Ctrl-Alt-C	This allows you to define keyboard shortcuts for the individual menus in Word and Excel so that the functions can be accessed without right-clicking (which is faster). The text of the menu must be entered exactly as it appears (in Word's Freestyle, by specifying the model where it appears in the context menu ), then the key combination; supported are Ctrl, Alt, Shift, all letters, numbers, F-keys and various navigation and other keys (but not from the number pad); if you want AltGR, you should write Ctrl-Alt; the combination is displayed in the menu as a tooltip. Note: Red Ink cannot overwrite certain default assignments, i.e.

VISCHER

			they simply do not work; furthermore, the keyboard shortcuts only work if the helper is running for Word or Excel (see para. 228 et seq. above)
UpdateCheckInterval	=	-1	The number of days after the last update that the add-ins should wait to check for new updates or attempt to do so; a 'silent' check only works if an installation has been made directly via the Internet (so-called ClickOnce installation = method 1, see para.212); if an update path is configured for local updates, reminders for updates will also be displayed, but the user will be asked each time whether an update should be attempted (the online update function is overridden if such a path exists); default is 7; if the value is set to 0, there is never a prompt; if it is set to -1, there is a prompt at every start (recommended only for ClickOnce installations); if there is no update path, there is no prompt for local updates.
UpdatePath	=	X:\Updates\RedInk\	you want to enable local updates directly from the add-ins, you have to specify the path to the installation files of Red Ink here, as they can be obtained from the installation package (para. 212); the subfolders "word", "excel" and "outlook" are expected in this directory; the add-ins must have been installed from the same path, otherwise the update will not work.
AlternateModelPath	=	X:\Configuration\all models.ini	Here you can optionally configure a configuration file with the details for further language models, which can be selected in Word when calling "Freestyle (2nd)" as an alternative to the secondary language model; this parameter must contain the full path including file name of the configuration file of these alternative language models; its format and content are described in para. 267 et seq.
SpecialServicePath	=	X:\Configuration\specialservices.ini	Here, services can be configured that can be accessed in Word via the menu item Analyze; these can be legal information systems, but also specialized AI models or internal vector databases; the functionality and configuration are described in para. 55 et seq.)
SpeechModelPath	=	X:\Speech\	The path where the Vosk and Whisper speech-to-text models are stored, in case the Tran-

VISCHER

			scriptor is to be used; the Vosk models are available at <a href="https://alphacephei.com/vosk/">https://alphacephei.com/vosk/</a> models (as a ZIP file) and for Whisper at <a href="https://huggingface.co/ggerganov/whisper.cpp/tree/main">https://huggingface.co/ggerganov/whisper.cpp/tree/main</a> ; Red Ink expects this directory to contain a subdirectory for each model with the name starting with "vosk-model" or "ggml"; the value is optional; the prompt library for transcripts can also be stored in this directory
LocalModelPath	=	X:\Models\	The path where models used by Red Ink are stored locally (or on a network drive); this can be, for example, the embedding model for "Context Search" or an anonymization model; the models are each stored in a subdirectory specified by Red Ink, in "Context Search" it is called, for example, "embed" (see there)
TTSEndpoint	=	<a href="https://texttospeech.googleapis.com/v1/">https://texttospeech.googleapis.com/v1/</a> ; <a href="https://api.openai.com/v1/audio/speech">https://api.openai.com/v1/audio/speech</a>	The URL of the Text-to-Speech-API of Google and OpenAI, if the function Create Podcast or Create Audio is to be used; using it requires for Google or OpenAI that the primary or secondary API is configured for the Vertex API of Google or OpenAI; if both Google and OpenAI are configured, the URLs are to be separated by ";"
ContextMenu	=	Yes	Whether the context menu should be displayed in Word and Excel when helpers are available; default is 'Yes'
UsageRestrictions	=	You may use Red Ink for all kinds of data, including professional secrecy data.	This is any text that appears in the add-in when the user moves the mouse pointer over the Red Ink logo; it can be used to alert the user to usage restrictions.
SecondAPI	=	Yes	"Yes" if a second model is to be configured and made available via Freestyle; this is optional; the same or a different API can be configured as for the main model; if no second model is needed, then set to "No"
APIKey_2	=	sk-proj-XXXX	As above, for the second model
APIKeyEncrypted_2	=	Yes	As above, for the second model
APIKeyPrefix_2	=	sk-proj-	As above, for the second model
Endpoint_2	=	<a href="https://api.openai.com/v1/chat/completions">https://api.openai.com/v1/chat/completions</a>	As above, for the second model
HeaderA_2	=	Authorization	As above, for the second model
HeaderB_2	=	Bearer {apikey}	As above, for the second model
Response_2	=	content	As above, for the second model



VISCHER

APICall_2	=	{'model': '{model}', 'messages': [{ 'role': 'user', 'content': '{promptsystem}{promptuser}' }], 'temperature': 1.0}	As above, for the second model
APICall_Object	=	, {"inlineData": {"mimeType": "mime-type"}, "data": "{encodeddata}"}	As above, for the second model
Timeout_2	=	200000	As above, for the second model
Temperature_2	=	1.0	As above, for the second model
Model_2	=	o1-preview	As above, for the second model
MaxOutputToken_2	=	8192	As above, for the second model
Anon_2	=	none; 0	As above, for the second model
TokenCount_2	=	candidatesTokenCount; thoughtsTokenCount; 5; CHF	As above, for the second model
OAuth2_2	=	No	As above, for the second model
OAuth2ClientMail_2	=	red-ink@earnest-racecars-212313-x4.iam.gserviceaccount.com	As above, for the second model
OAuth2Scopes_2	=	https://www.googleapis.com/auth/cloud-platform	As above, for the second model
OAuth2Endpoint_2	=	https://oauth2.googleapis.com/token	As above, for the second model
OAuth2ATExpiry_2	=	50	As above, for the second model
MarkupMethodHelper	=	2	Specifies which method should be used to create a markup when using the Word Helper Self-Compare Selection: 1 = Word's Compare function; this inevitably causes windows to open temporarily, and other add-ins (such as iManage) interfere with the process because they are not programmed to comply 2 = simple diff algorithm; it is less problematic and faster than Word for short texts, but not as reliable 3 = the same diff algorithm, but the result is displayed in a window, which is the fastest method The default is 3 For further details, see para.25 above
MarkupMethodWord	=	1	Specifies which method should be used to create a markup when using the various pre-programmed functions of Word (if one is created): 1 = Word's Compare function; this inevitably causes windows to open temporarily, and other

VISCHER

			<p>add-ins (such as iManage) interfere with the process because they are not programmed to comply (= default)</p> <p>2 = simple diff algorithm; it is less problematic and faster than Word for short texts, but not as reliable</p> <p>3 = the same diff algorithm, but the result is displayed in a window, which is the fastest method</p> <p>4 = an LLM- and regex-based method, which, depending on the LLM, works as a diff for larger texts and, above all, leaves formatting intact; however, it is less reliable in terms of content</p> <p>The default is 3</p> <p>For further details, see para. 25 above</p>
MarkupMethodOutlook	=	2	<p>Specifies which method should be used to create a markup when using the various pre-programmed Outlook functions (if any is created):</p> <p>1 = Word's Compare function; because Outlook does not recognise revision marks, the markup is displayed in colour.</p> <p>2 = simple diff algorithm; it is less problematic and faster than Word for short texts, but not as reliable</p> <p>3 = the same diff algorithm, but the result is displayed in a window, which is the fastest method</p> <p>The default is 3</p> <p>For further details, see para.25 above</p>
MarkupDiffCap	=	20000	<p>Specifies the maximum text length for which the diff markup method should be used (if DiffW is not used) and for the markdown conversion of selected text, because it is not suitable for very long texts; however, it is possible to decide to use it anyway in individual cases (otherwise DiffW); default is 20'000; for further details, see para. 25 above</p>
MarkupRegexCap	=	30000	<p>Specifies the maximum text length for which the Regex markup method should be used, because it can take a long time and is less reliable for longer texts; however, it is possible to decide to use it anyway in individual cases (otherwise Word Compare); default is 30'000; for further details see para.25 above</p>
MarkdownConvert	=	Yes	<p>Indicates whether, for texts in Word, various formatting such as bold, italic, or underlined</p>

VISCHER

			should be converted to Mark-down beforehand so that they can be converted back into formatting afterward, since most LLMs support this and Red Ink supports Markdown when inserting text; this is enabled by default, but limited to the number of characters as set by MarkupDiffCap
KeepFormat1	=	No	Specifies whether the translation function in Outlook and Word should try to preserve the basic formatting (characters, lists) or whether it should work in plain text (in which case formatting may be lost). If the formatting is retained, but processing takes significantly longer because the formatting has to be coded into the text that is processed by the AI (and it takes longer for the AI to do this); the default is "No"; for further details, see para. 25 above
KeepFormat2	=	Yes	Specifies whether the correction, improvement and abbreviation functions in Outlook and Word should try to preserve the basic formatting (characters, bullets) or whether they should work in plain text (in which case formatting may be lost ); if the formatting is retained, the processing takes significantly longer because the formatting has to be coded into the text that is processed by the AI (and it takes longer for the AI to do this); the default is "No"; for further details, see para.25 above
KeepParaFormatInLine	=	Yes	Specifies whether Word should try to encode the formatting of the paragraphs of the original text into the text of the paragraphs when editing selected text, so that the add-in can try to restore at least the paragraph formatting when the AI is run ; this is less far-reaching than KeepFormat, and also takes less time and uses less data; even without this setting, the add-in will try to remember the paragraph formatting where it fits; default is "No"; for further details, see para. 25 above
KeepFormatCap	=	5000	The KeepFormat and Keep-ParaFormatInLine functions can be automatically switched off in this way for longer texts, so that not too much time is spent on them which could result in Red Ink being stuck; specified is the number of characters above which the automatic shutdown

VISCHER

			occurs (a value of 0 means no check, a value of 1 always dispenses with saving the format in the text itself); default is 5000; for further details, see para. 25 above
ReplaceText1	=	Yes	Specifies whether the translation function in Outlook and Word should replace the text to be translated with the translation or insert the translation afterwards; default is "Yes"; for further details see para. 25 above
ReplaceText2	=	No	Specifies for some of the other pre-programmed functions (such as corrections, abbreviations) in Outlook and Word, whether the selected text should be replaced by the AI's response or whether the response should be inserted afterwards; default is "No"; for further details, see para. 25 above
DoMarkupWord	=	Yes	Specifies whether Word should automatically display a markup of the changed text when certain predefined functions (such as corrections, abbreviations) are used (which may take some time); default is "Yes"; for further details, see para 25 above
DoMarkupOutlook	=	Yes	Specifies whether in Outlook, for certain predefined functions (such as corrections, abbreviations), a markup of the modified text should also be displayed automatically (which may take some time); default is "Yes"; for further details, see para. 25 above
PromptLib	=	%APPDATA%\Microsoft\Word\promptlib.txt	If a value is specified, the add-in will load the stored prompts from the specified file when the Freestyle function is used and offer them to the user for selection if they do not specify a prompt (see below for the file format); the file path with the name of the file must be specified (it must be in txt format); for further information, see para. 261 below
PromptLib_Transcript	=	X:\Speech\promptlib_transcript.txt	The path to the prompt library for the process function in the Transcripator; the text file follows the same syntax as the "normal" prompt library on the preceding line; for the Transcripator, see para. 70 et seq. above
ISearch	=	Yes	"Yes", if an internet query should also be possible in the Freestyle function; the default is "Yes"

VISCHER

ISearch_URL	=	<a href="https://duckduckgo.com/html/?q=">https://duckduckgo.com/html/?q=</a>	The URL that should be used for the internet search; default is DuckDuckGo
ISearch_ResponseMask1	=	<a href="https://duckduckgo.com/html/?uddg=">duckduckgo.com/html/?uddg=</a>	This value specifies which characters are found to the left of the URL in the HTML code of the search engine results page (the value is used to identify the results); the default is as shown on the left
ISearch_ResponseMask2	=	&	This value specifies which characters are found to the right of the URL in the HTML code of the search engine results page (the value is used to identify the results); the default is as shown on the left
ISearch_Name	=	DuckDuckGo	The name of the search engine (partially displayed to the user); the default is as shown on the left
ISearch_Tries	=	10	How many search results the AI should look at (top to bottom) when using the internet function until it has the amount of content defined below; default is 10, maximum is 30
ISearch_MaxDepth	=	2	How deeply the add-in should dive into a website that is visited as a result, because with many of the more complex websites the information is only found on sub-pages; default is 2, maximum is 10
ISearch_Timeout	=	3	How long the add-in should devote to a search hit (in seconds); the default is 3, the maximum is 60
ISearch_Results	=	2	After how many successful search hits should the search stop? The default is 4, the maximum is 15
ISearch_Approve	=	No	Before sending anything to the search engine, the add-in can ask whether it is allowed to do so (and display the search query); this can help to maintain confidentiality; the default is "No"
ISearch_SearchTerm_SP	=	You are a ...	Prompt used to determine the appropriate internet search terms; placeholders can be used (see also below): {OtherPrompt} = instruction {CurrentDate} = actual date
ISearch_Apply_SP	=	You are a ...	Prompt, which is used to implement the Freestyle command with the internet search results; placeholders can be used (see also below): {OtherPrompt} = instruction {SearchResult} = search results

ISearch_Apply_SP_Markup	=	You are a ...	Prompt, which is used to implement the Freestyle command with the internet search results when a text has been selected; placeholders can be used (see also below): {OtherPrompt} = instruction {SearchResult} = search results
Lib	=	Yes	"Yes" if the library search should be activated; default is "No"
Lib_File	=	C:\users\username\Documents\library.txt	The filename with full path where the library file is located (in txt, doc, docx, or rtf format)
Lib_Timeout	=	30000	LLM timeout where this is used for the library function (in milliseconds); default is 60000
Lib_Find_SP	=	You are a ...	Prompt, which is used to extract the relevant information from the library for the library function; the following placeholder can be used (see also below): {OtherPrompt} = instruction {LibraryText} = library contents
Lib_Apply_SP	=	You are a ...	Prompt to apply the found content to the user's request; the following placeholder can be used (see also below): {OtherPrompt} = instruction
Lib_Apply_SP_Markup	=	You are a ...	Prompt to apply the found contents to the user's request for an existing text as markup; the following placeholder can be used (see also below): {OtherPrompt} = instruction
LicensedTill	=	30.9.2025	You can enter the end date of your licence for Red Ink here; after that, Red Ink will no longer work with this setting (with a reminder appearing from seven days before); if nothing is specified, the maximum lifespan of the application is as stored in the code ; the date is displayed in the "About Red Ink" window, which can be accessed from the Settings window; the tool does not check whether the date is plausible – the parameter is thus intended as a reminder
SP_Translate	=	You are a ...	Prompt for translations; the following placeholder can be used (see also below): {TranslateLanguage} = language
SP_Correct	=	You are a ...	Prompt für corrections (see also below)
SP_Improve	=	You are a ...	Prompt for linguistic improvements (see also below)
SP_NoFillers	=	You are a ...	Prompt for removing filler words and redundancies (see also below)

SP_Convincing	=	You are a ...	Prompt for more convincing wording (see also below)
SP_Friendly	=	You are a ...	Prompt for more friendly wording (see also below)
SP_Shorten	=	You are a ...	Prompt for abbreviations; the following placeholder can be used (see also below): {ShortenLength} = output length in words
SP_Summarize	=	You are a ...	Prompt for summaries; the following placeholder can be used (see also below): {SummaryLength} = output length in words
SP_Explain	=	You are a ...	Prompt for analyzing a text (see also below)
SP_SuggestTitles	=	You are a ...	Prompt for suggesting titles (see also below)
SP_SwitchParty	=	You are a ...	Prompt for the Switch Party function; the following placeholders can be used (see also below): {OldParty} = previous Partei {NewParty} = new Partei
SP_Anonymize	=	You are a ...	Prompt for anonymisation (see also below)
SP_WriteNeatly	=	You are a ...	Prompt for completions in Excel; the following placeholder can be used (see also below): {Context} = captured context
SP_FreestyleText	=	You are a ...	Prompt for the Freestyle function if text is selected (without the additions, see below); the following placeholder can be used (see also below): {OtherPrompt} = Instruction
SP_FreestyleNoText	=	You are a ...	Prompt for the Freestyle function if no text is selected (without the additions, see below); the following placeholder can be used (see also below): {OtherPrompt} = Instruction
SP_ContextSearch	=	You are a ...	Prompt for the context search function; the following placeholder can be used (see also below): {SearchContext} = context
SP_ContextSearchMulti	=	You are a ...	Prompt for the context search function for when all hits are to be found in one part of a text; the following placeholder can be used (see also below): {SearchContext} = context
SP_RangeOfCells	=	You are a ...	Prompt for the range function in Excel; the following placeholder can be used (see also below): {OtherPrompt} = instruction
SP_MailReply	=	You are a ...	Prompt for composing an e-mail reply in Outlook; the following placeholder can be used (see also below): {OtherPrompt} = notes

SP_MailSumup	=	You are a ...	Prompt for summarising an email chain (see also below)
SP_MailSumup2	=	You are a ...	Prompt for summarising multiple selected emails (see also below)
SP_Add_KeepFormulasIntact	=	Beware, the text contains ...	Additional prompt to instruct the language model to retain formulas contained in Excel cells
SP_Add-KeepHTMLIntact	=	When completing your task, leave any HTML tags ...	Additional prompt to instruct the language model not to remove HTML codes (they are needed to store formatting)
SP_Add_KeepInlineIntact	=	Do not remove any coding with ...	Additional prompt for the language model to instruct it not to remove other formatting codes that are encoded in the text (they are needed to store formatting)
SP_Add_Bubbles	=	Provide your response to ...	Additional prompt for the language model for Freestyle in Word to format the output so that it can be inserted into comments on the text; the add-in then goes through these individually and adds the comments
SP_Add_Slides	=	You shall provide your output ...	Additional prompt that instructs the language model for Freestyle in Word to format the output in the form of additional pages for a PowerPoint file; it provides the information for creating the slide pages in the form of a JSON string, which Red Ink can then implement
SP_MarkupRegex	=	You are an expert text comparison ...	Additional prompt used to instruct the language model for Freestyle in Word to compare two texts and to display the differences as a Regex search pattern so that the "Regex" markup method can be implemented in Word's Freestyle
SP_Revisions	=	When you are asked to ...	Additional prompt that explains to the language model for Freestyle in Word how it can recognise markups in a text if the relevant function has been activated ("(rev)").
SP_ChatWord	=	You are a helpful ...	Basic prompt for the chatbot within Word
SP_Add_ChatWord_Commands	=	You help the user ...	Additional prompt to explain to the chatbot which commands to use and how to use them to format text
SP_ChatExcel	=	You are a helpful ...	Basic prompt for the chatbot within Excel
SP_Add_ChatExcel_Commands	=	You help the user ...	Additional prompt to explain to the Excel chatbot which commands to use and how to use them to work with worksheets
SP_Podcast	=	You are a profes-	Prompt for the creation of pod-



		sional podcaster ...	cast scripts based on a text; the following placeholders can be used (see also below): {Language} = Language {TargetAudience} = Target audience {DialogueContext} = Context {Duration} = Duration {ExtraInstructions} = Additional instructions/prompts
ChatCap	=	50,000	How many characters from the previous dialogue the chatbot is given with each new question (in addition to the question, the base and additional prompts and the current text); default value is 50,000
SP_MergePrompt	=	You will be provided a text to insert into another text ...	The standard prompt used in Word for the LLM-based merging of text selected in a pane and text selected in the current Word document; it is displayed to the user in an input window and can be modified by them; it is sufficient to reference the text to be inserted and the other text; the additional prompt SP_Add_MergePrompt then provides the model with the more specific instructions
SP_MergePrompt2	=	You will be provided a text to insert into another text ...	Same, but for the "Apply comment" function
Add Merge Prompt	=	The text to insert or merge ...	The prompt addition, which is appended to the prompt for merging texts so that everything works. It tells the model that the text in the Red Ink pane is sent between the tags "<INSERT>" and "</INSERT>", and the text in the document between the tags "<TEXTTOPROCESS>" and "</TEXTTOPROCESS>", so that the user doesn't have to specify this
SP_InsertClipboard	=	You will receive a binary object ...	The prompt used in the Word helper "Clipboard to Text" to convert the clipboard contents to text

252 For the **switch values**, "Yes" and "True" are equivalent to one another, as are "No" and "False".

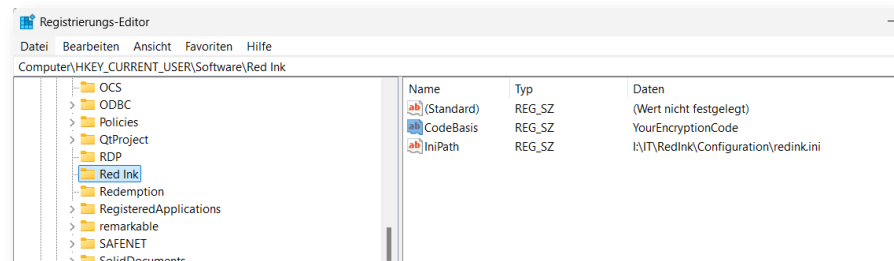
253 The **prompts** that Red Ink uses for its basic functions do not normally need to be configured. However, they can be changed via the parameters if they do not fit the model used or if the model does not follow them sufficiently. The initial access to this is via the "Expert Config" function in Settings, where they can be copied out manually (for various reasons, they are not written to the configuration file when the configuration is saved, provided they correspond to the current standard of the respective version). The prompts can be seen to be structured as system prompts, and the selected text is passed as a user

prompt (if the model differentiates at all). The user prompt normally contains the text between the tags "<TEXTTOPROCESS>" and "/<TEXTTOPROCESS>". For certain other functions, other tags are used, as can be seen from the standard system prompts. These tags should also be used for changes to ensure that the prompts work optimally. Most of the system prompts also contain the standard placeholder "{INI\_PreCorrection}"; the value of the corresponding parameter, which can also be stored in the configuration file, is entered in place of this. The user can also define it via Settings.

- 254 The **key** for encrypting the API key or private key and the IniPath for the central configuration file are not stored in the configuration file (see paras. 280 et seq. and paras. 257 et seq. below). Some local parameters, such as the **last Freestyle prompt** or, in Word, the latest parameters for the chatbot, are also not stored in the configuration file (but in the local installation).
- 255 In the configuration file, lines starting with a **semicolon** are ignored. If mandatory values (keys) are missing, the respective add-in prompts the user to enter them. If the file is missing, the setup wizard is started (see para. 223 above).
- 256 File paths can use the following **placeholders** (these can be used to make the above variables for files more flexible):

%APPSTARTUPPATH%	The directory from which the respective application was started
%APPDATA%	The Applications data folder of the respective user, e.g. C:\Users\Username\AppData\Roaming; this placeholder is best suited for Red Ink because it can be used to build the path to the local copy of the prompt library; in the roaming directory, there is a directory "Microsoft" and under that, the directories for Word, Excel, and Outlook, where a local copy of "redink.ini" may also be found.
%USERPROFILE%	The profile directory of the respective user, e.g. C:\Users\Username
%WINDIR%	The Windows directory, normally C:\Windows
%TEMP%	The directory for temporary data
%HOMEPATH%	The directory of the user profile, but without the drive specification
%DESKTOP%	The Desktop folder of the user

- 257 As mentioned above, it is possible to define a **central path for the configuration file** "redink.ini" for all add-ins, so that the configuration file is loaded from a central location (and can thus be managed uniformly for all users). To do this, an entry must be made in the Windows registry on all devices in the organisation, as follows, under the "IniPath" key of the "Red Ink" entry:



258 The registry path is already stored as a constant in the code of the add-ins (i.e. in their shared library), as is whether the entry in the registry should take precedence over an existing "redink.ini" file in the AppDir directory for the configuration file. As shown (see para. 247 above), the latter case is programmed by default, i.e. if a configuration file for each add-in is present in the (local) default directory, this is used. Otherwise, the one in WordDir is used and then the one in the path recorded in the registry (typically the central copy of "redink.ini"):

```
Public Const RegPath_Base As String = "HKEY_CURRENT_USER\Software\" & AN3 & "\"
Public Const RegPath_CodeBasis As String = "CodeBasis"
Public Const RegPath_IniPath As String = "IniPath"
Public Const RegPath_IniPrio As Boolean = False ' True if the registry path shall have priority
```

259 An organisation that wants to **distribute Red Ink** should therefore only let users install the three add-ins (and the helpers if necessary) and make the corresponding entry in the user's registry in a central directory where a copy of "redink.ini" can be found for everyone. If a user wants a different configuration, they can do so by saving the configuration in Settings, thus decoupling themselves from the central configuration file (however, they can also "abandon" their local configuration at any time via Settings and return to the central configuration).

260 Within the add-in for Word, the **registry entry** can be written locally by writing down the path in a Word document (without "redink.ini"), selecting it and calling up Freestyle. Instead of the prompt, enter "ini-path" as the command line instruction (see para. 42 above). This will write the selected value in the appropriate location in the registry, provided this is allowed.

## B. Prompt library

261 The Word add-in supports the use of a prompt library if the corresponding parameter ("PromptLib") points to a valid path with a corresponding text file.

262 The text file must be saved as a plain text file (not a Word file; save it as a.txt file in Word) and must have the following content:

- Each line must start with the title or a short description of the prompt (e.g. "contract comparison"). This is followed by the character "|" (AltGr-7) and then the prompt (it may contain the separator);

- A new line must be used for each prompt;
- Blank lines and lines beginning with ";" are ignored.

263 The prompt library can be stored on your own computer, in the same directory as the configuration file or in a shared directory on the network. If the add-in does not find the file, an error is reported when the function is used (not when the add-in is started).

264 If changes are made to the file (which is also possible within the respective add-in using the "Edit" button in the prompt library selection), these will be taken into account the next time the function is accessed because it is reloaded each time. This means you can experiment with the prompts and adjust them straight away if they are not yet correct.

265 The prompts should be formulated in such a way that they can also be used in Freestyle, i.e. the prefixes and triggers can also be used. The installation package contains a number of sample prompts in a prompt library file for illustrative purposes – use them at your own risk. If you have a good prompt, please let us know so that we can add it to the collection if appropriate. We have already expanded the sample prompt library with a few interesting prompts that were submitted. The latest version can be obtained from <https://apps.vischer.com>.

266 The prompt library used by the Transcriptor works with the same file format.

### **C. Other alternative language models**

267 In Word, when calling "Freestyle (2nd)", it is possible to choose between further, alternative language models in addition to the possibly configured secondary language model (para. 38). For this purpose, a configuration file must be defined in which the various alternative language models are stored with their configuration, and the path for this file must have been specified with the parameter "AlternateModelPath =" in "redink.ini".

268 Format and content of the configuration file for the alternative language models correspond to that of "redink.ini" (see para. 246). For each model, the necessary information must be stored there in the same way as for the primary model (e.g., with the parameter "APIKey = ...." or "Endpoint = ..."). The only difference is that each model configuration must be introduced by a line containing the description of the model in square brackets:

```
[Perplexity Sonar Pro: Will also search the Internet (3.3 Min. Timeout, USA)]

APIKey = pplx-xxxxxx
APIKeyPrefix = pplx-
APIKeyEncrypted = True
Model = sonar-pro
Endpoint = https://api.perplexity.ai/chat/completions
HeaderA = Authorization
HeaderB = Bearer {apikey}
Response = content
APICall = {"model": "{model}", "messages": [{"role": "system", "content": "Follow the user's inst
they are drafted like a system prompt."}, {"role": "user", "content": "{promptsystem}
{promptuser}"}], "temperature": {temperature}, "top_p": 0.9, "search_domain_filter": null, "return
"return_related_questions": false, "top_k": 0, "stream": false, "presence_penalty": 0, "frequen
Timeout = 200000
Temperature = 0.2

[Perplexity Sonar Reasoning Pro: Will reason and search the Internet (6.6 Min. Timeout, USA)]

APIKey = pplx-xxxxxx
APIKeyPrefix = pplx-
APIKeyEncrypted = True
Model = sonar-reasoning-pro
Endpoint = https://api.perplexity.ai/chat/completions
```

- 269 This description will then be shown to the user when they have to choose their model when calling "Freestyle (2nd)" (if a path to the configuration file has been set in "redink.ini").
- 270 If the user selects such an alternative model, it will be reset back to the originally defined secondary model after the execution of Freestyle. This reset can be deselected with a checkbox in the model selection. It then remains the secondary model until Red Ink is restarted or "redink.ini" is reloaded. If the configuration file is saved, the previous secondary model will be overwritten.
- 271 Caution: When configuring reasoning models, make sure that a sufficient timeout is configured and that Red Ink does not display the "thinking process". The user must therefore be patient.

#### **D. OAuth2.0 (e.g. Google Vertex API)**

- 272 If you want to access an endpoint of a language model, you usually have to identify yourself to it using an API key (e.g. with OpenAI and Azure OpenAI Services or with the free Google AI offerings). For endpoints like Google's Vertex AI API, on the other hand, the OAuth 2.0 procedure is used, which offers more security but also makes it more complicated. Red Ink also supports it.
- 273 For this method, a service account must first be created on the server, i.e. an account specifically for Red Ink. To do this, you need to access the administrator console. This type of account also allows access for users without their own user account. A public and private key must then be generated for this account. It should be exportable in the form of a JSON file because you need the private key to store it in the Red Ink configuration file. This JSON file, for example, looks like this:

```
{
  "type": "service_account",
  "project_id": "earnest-5742c7c0-1000-4000-8000-8000",
  "private_key_id": "240e3efc5742c7c0-1000-4000-8000-8000",
  "private_key": "-----BEGIN PRIVATE KEY-----\nMIIEvQIBADANBgkqhkiG9w0BAQEFAASCBKcwggSjAgEAAoGBA\n-----\n",
  "client_email": "red-dragon-test1@earnest-5742c7c0-1000-4000-8000-8000.gserviceaccount.com",
  "client_id": "240e3efc5742c7c0-1000-4000-8000-8000",
  "auth_uri": "https://accounts.google.com/o/oauth2/auth",
  "token_uri": "https://oauth2.googleapis.com/token",
  "auth_provider_x509_cert_url": "https://www.googleapis.com/oauth2/v1/certs",
  "client_x509_cert_url": "https://www.googleapis.com/robot/v1/metadata/x509/red-dragon-test1@earnest-5742c7c0-1000-4000-8000-8000.gserviceaccount.com",
  "universe_domain": "googleapis.com"
}
```

274 From this JSON file, you take the value "private\_key" (without the prefix and suffix, just the naked key; it doesn't matter if the key itself contains line breaks in the format "\n"; they are removed to increase the security of the encryption), "client\_email" and "auth\_uri" as well as the appropriate scopes value (in the case of Vertex, this is "https://www.googleapis.com/auth/cloud-platform"). These are stored in the configuration file, where the private key should be stored as an encrypted string (see below for an example of how to encrypt it using Red Ink in Word). With this information, Red Ink can obtain an access token, which can then be used as a kind of API key until it expires (usually after 3600 seconds, which can also be configured). After that, Red Ink automatically obtains a new access token. The private key must be kept secret. If it is compromised, however, it can be reset via the console of the endpoint provider. Red Ink offers a certain degree of protection with the option of lightweight encryption (see the next section).

## E. Configuration of Advanced API Calls

275 For a normal LLM, it is sufficient to define a simple API call text ("API-  
Call") and specify in which field of the returned JSON string the LLM's  
response will be contained ("Response").

276 For the Special Service features, however, more complex programming is usually required, since Red Ink sometimes has to obtain the responses in two steps, and sometimes has to read them from the returned JSON strings in several steps. Here is an example of a more complex programming (using LexiFind as an example):

```

APIKey =
APIKeyPrefix =
APIKeyEncrypted = False
Model = LexFind
Endpoint = https://www.lexfind.ch/api/fe/de/fulltext-search|https://www.lexfind.ch/api/fe/de/fulltext-search/{id}?session_id={session_id}&page_no=1
&results_per_page=25
HeaderA =
HeaderB =
Response = id;session_id|**Lexfind.ch ({results[*].number_of_results|/} Treffer)**]
(https://www.lexfind.ch/fe/de/search/{id}/{session_id}/de) (hier max. 25):\n\n{% for
texts_of_law_with_matches %}[**{systematic_number} - {matches[0].title}**]
(https://www.lexfind.ch/fe/de/{dta_urls[0].url})\n\n{matches[0].keywords}\nStatus:
{matches[0].info_badge|removed=Entfernt;abrogated=Ausser Kraft;current=Aktuell;not_current=Nicht
Aktuell} - {matches[0].version_active_since} - {dta_urls[0].language} - [{entity.name}]
({dta_urls[0].original_url})\n\nFundstelle: {htmlnocr:matches[0].snippet}\n\n{% endfor %}
APICall =
{"search_text":"{promptuser}","active_only":true,"search_in_systematic_number":{parameter3},"search_in_title":{parameter3},"search_in_keywords":{parameter4},"search_in_content":{parameter2},"entity_filter":{parameter1},"systematic_filter":[],"category_filter":[],"use_global_systematics":true,"direct_search":false}
Timeout = 200000
Parameter1 = Gemeinwesen; String; Bund (CH); Bund (CH)<[27]>, Bund und Kantone (alle)<[]>, Aargau (AG)<[1]>, Appenzell Ausserrhoden (AR)<[3]>, Appenzell Innerrhoden (AI)<[2]>, Basel-Landschaft (BL)<[5]>, Basel-Stadt (BS)<[6]>, Bern (BE)<[4]>, Freiburg (FR)<[7]>, Genf (GE)<[8]>, Glarus (GL)<[9]>, Graubünden (GR)<[10]>, Intlex (Intlex)<[28]>, Jura (JU)<[11]>, Luzern (LU)<[12]>, Neuenburg (NE)<[13]>, Nidwalden (NW)<[14]>, Obwalden (OW)<[15]>, Schaffhausen (SH)<[17]>, Schwyz (SZ)<[19]>, Solothurn (SO)<[18]>, St. Gallen (SG)<[16]>, Tessin (TI)<[21]>, Thurgau (TG)<[20]>, Uri (UR)<[22]>, Waadt (VD)<[23]>, Wallis (VS)<[24]>, Zug (ZG)<[25]>, Zürich (ZH)<[26]>
Parameter2 = Suche im Erlassesstext; Boolean; True
Parameter3 = Suche im Titel/SR; Boolean; True
Parameter4 = Suche in Stichworten; Boolean; True

```

## 277 Red Ink supports the following functions:

- In the "Endpoint" parameter, two endpoints can be defined, separated by "|". The first endpoint is normally addressed with a POST command. If a second endpoint is defined, it is addressed with a GET command, whereby results from the first command are inserted in place of the placeholder. In the example above, these are the values {id} and {session-id}, which are extracted directly from the JSON response of the POST command. These two values are stored in the left part of the "Response" parameter, i.e., the part to the left of "|" (it is used to evaluate the responses delivered by the endpoint). The response of the second endpoint (i.e., the GET request) is evaluated by the right part (more on this in a moment).
- In the "Endpoint" parameter, the first endpoint can also be executed as a GET command instead of POST. For this, the prefix "GET:" must be placed before the URL.
- In both cases, various values can be inserted into the "Endpoint" and "APICall" parameters using placeholders, which are then inserted before the call (in "Endpoint", spaces are replaced with "+"):
  - {model} = Model name
  - {apikey} = API key
  - {ownsessionid} = a unique ID that Red Ink generates on its own
  - {promptsystem} = the command prompt sent to the LLM



- {promptuser} = usually the text that the user has selected (without the "<TEXTTOPROCESS>" tags)
- {userinstruction} = die barebones instruction that has been provided by the user in the Freestyle prompt box
- {temperature} = Temperature (only for "APICall")
- In the case of a Special Service: {parameter1}, {parameter2}, {parameter3}, {parameter4}
- The "Response" parameter can be set in three ways:
  - A single label such as "text" or "response": In this case, the first corresponding value is extracted from the JSON;
  - The label "JSON": In this case, the JSON string is output as received; this is used for debugging or programming more complex templates (more on this in a moment);
  - A complex template, as in the example above: In this case, Red Ink will process the template and construct a corresponding text string. Since Red Ink can output or display Markdown-formatted strings in the pane, the template can be used to incorporate corresponding formatting. Essentially, a template thus consists of text elements with formatting, placeholders for values extracted from the JSON response, and a loop function to extract multiple JSON elements. Further details are included below.

If two endpoints are defined in "Endpoint", two elements must also be defined in "Response", separated by "|".

278 Detailed instructions on how to program "Response" templates can be found in Annex 2. If this is too complicated, Red Ink offers automatic programming. For this purpose, an example output of the endpoint must be specified in Word (use the "Response" value "JSON" for this) and below it, describe what the output of the template should look like. Both are selected and Freestyle is chosen. Enter "generateresponsetemplate" there. Red Ink will give the currently configured model the necessary instructions to program the template. It will appear in Word shortly after.

279 The installation package contains several configuration examples for various special service offers, including the templates.

## **F. Security features**

280 The source code of Red Ink is open access. The configuration file is also open access. Nevertheless, the API key (or the refresh token and the ClientSecret when using OAuth 2.0), which grants access to the LLM API, can be stored in encrypted form. It is also possible to configure Red Ink so that the respective copy only runs in a specific network (which may be necessary in certain cases).



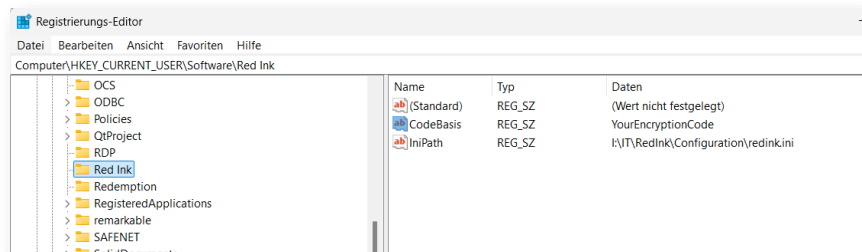
- 281 The **encryption of** the API key, refresh token and ClientSecret ("XOR") is not technically strong. However, it should be sufficient to ward off non-specialised attacks. In addition, these attacks typically assume that parts of the encrypted plain text are known, which is not the case here. For this reason, the prefix that is sometimes used for API keys is masked (for the private key in the OAuth 2.0 procedure, the prefix and suffix are omitted for this reason).
- 282 The key used is a text that can be freely designed (it is not set by default, but in the examples below it has the value "SecretValue"). However, it must be stored in a way that protects it from unauthorised access but is still accessible to the software. This is a challenge for open source software if a more elaborate authentication procedure is not to be used. We have therefore developed the following two methods for storing the key for encryption and decryption:

- **Method 1 (hardcoded):** The encryption key is stored in the code of the add-in itself, defined at the beginning as the "private" constant "Int\_CodeBasis" ("private" so that it cannot be read by another module). In the source code of the add-in library (SharedLibrary), this would look like this, where "YourEncryptionCode" would stand for your own key:

```
' Amend the following two values to hard code the encryption key and permitted domains  
Private Const Int_CodeBasis As String = "YourEncryptionCode"  
Public Const alloweddomains As String = "int.company.com, int2.company.com"
```

This allows an organisation to choose its own key, specify it in the VB.net code, recompile the code and install the finished executable files where the code can no longer be viewed. The key therefore remains hidden and is only available to the add-in itself. The disadvantage of this method is that, although the code can be compiled easily using development environments such as the free "Visual Studio 2022" with free libraries, it requires a certain amount of expertise. Furthermore, the procedure of adapting the code and installing password protection has to be repeated for each new version. However, for an appropriate fee, we can provide versions of the add-ins that have been provided with their own key.

- **Method 2 (registry):** The key for the encryption is not stored in the code itself, but in the Windows registry. This is where most programs store their configuration. A separate entry is created for Red Ink and the code is stored in plain text.



If the value *Int\_CodeBasis* is not set in the code (as provided for in method 1), the add-in will always try to retrieve the key from the registry at the location specified above whenever encryption is active. The path is also set as a constant in the code and could be changed there if necessary. The advantage of this method is that the code does not have to be changed. The disadvantage is that access to the registry must be blocked for those users who should not see the key (e.g. block access to "regedit"). However, this is done in many organisations anyway. However, it should be noted that any application can read the value. For example, if an organisation allows its employees to write macros in VBA themselves, they can create a small script that reads the value. To write the value to the registry, the RegEdit command can be used, but it is also possible to make the entry via the Word add-in. To do this, the desired key (e.g. "YourEncryptionCode") should be typed into a Word document without quotation marks and selected. Next, the Freestyle function should be chosen and, enter "codebasis" instead of the prompt (see para. 42 above). The add-in will write the value to the registry (overwriting an existing value if technically possible) and then the add-in must be restarted to load the new key.

- 283 Method 2 is used by default, although no value is stored in the registry the first time the application is installed. The installer will not set this value either. It must therefore be done manually.
- 284 To encrypt your API key and any private key associated with it, you will need to install Red Ink for Word and then copy and paste the API key etc. into a blank document, select it and then click on the Freestyle function. The command "encode" is entered there (see para. 42 above). Next, the secret code must be entered. Red Ink will write the encrypted API key on a new line in the same document. Attention: Automatic hyphenation must be switched off (the add-in tries to do this). If this is not done, the encrypted text will be copied with additional hyphens and cannot be decrypted correctly. Encrypted text can be decrypted using the "decode" command. It is important that any API prefix in the configuration file has been configured correctly (this only applies to the API key, not the private key). If the API key or the private key is encrypted, it can be entered in the configuration file and the value "APIKeyEncrypted" can be set to "Yes" (in the OAuth 2.0 procedure, this configuration applies to the private key). The same applies

to the second model. Before the private key is used, any line breaks or "\n" are automatically removed from the decrypted text. It therefore does not matter if they happen to be present.

- 285 The second security feature, **run-only-at-home**, is only necessary if method 1 is used above. This security function ensures that the relevant copy of the add-ins only runs in the programmed domains. This counteracts the use of the add-in with the API key etc. in other organisations or on private computers and thus the misuse of your own API key etc. Your own domain is also entered as a "AllowedDomains" constant at the beginning of the code module of the add-ins' shared library (in the example here, the domains "int.company.com" and "int2.company.com"):

```
' Amend the following two values to hard code the encryption key and permitted domains
Private Const Int_CodeBasis As String = "YourEncryptionCode"
Public Const alloweddomains As String = "int.company.com, int2.company.com"
```

- 286 To find out which domain is your own and in which domains Red Ink may be running, you can use the command line command "domain" in the Freestyle function of the Word add-in (see para. 42 above). It will display both.

## V. FAQs

### 1. How is Red Ink different from other AI tools?

Everyone has to judge that for themselves. From our point of view, Red Ink makes it is much easier to control what happens to your own data than with certain other services and products. We have integrated many functionalities directly into the program, and the costs of using AI are much lower than with some AI services, where each subscription per employee and month costs CHF/EUR 15-30. Red Ink usually costs a fraction of that. Finally, for us, Red Ink is the basis for further AI applications that use Red Ink as an interface to the major language models – for example, for other tools that we use to analyse legal texts. We do still use other AI tools in the firm for some tasks.

### 2. How well does Red Ink really work?

This depends to a large extent on the language model used. Red Ink is basically just an interface for accessing a language model as easily and diversely as possible. How well a text is corrected or summarised, whether Red Ink follows the instructions or provides the desired answers, depends on the language model. To ensure that Red Ink can be used effectively, an advanced language model with a good ability to follow instructions should be used.

3. **Sometimes I have to wait a long time for answers – why?**

This is mainly due to the language model; in our experience, there are big differences in speed between models. The more text to be processed, the more time the language model needs, of course. If, in addition, the functions for "remembering" the formatting are used, this will further slow performance because the formatting is stored in the text (to a greater or lesser extent, depending on the method used, see para. 25 above) with the techniques used, which can significantly increase the amount of text to be processed in some cases. The reverse coding and formatting also take a certain amount of time because Word (and Outlook) unfortunately do not support these things very well and therefore we had to find our own solutions for Red Ink. Sometimes, however, waiting times are simply due to the fact that the computer of the respective language model is overloaded or is otherwise stuck for some reason.

4. **Sometimes all the formatting in my texts gets corrupted – what can I do about it?**

In Red Ink, we have implemented various methods to prevent or mitigate this. These can be activated, but they cannot fully overcome the challenge because today's language models do support the processing of formatted texts only to a very limited extent. The methods we use work best with shorter or medium-length texts because they require a lot of processing power, which can slow down the process and overwhelm the language models (for the methods, see in particular para. 24 above). For longer texts, workarounds must therefore be used. One strategy is to work with shorter blocks of text (see, for example, the "(iterate)" trigger within Freestyle), another is to use the commenting instead of the revision function ("Bubbles:") – and then to implement the comments manually. Those who experiment with the tool will quickly find out what works best for their own needs.

5. **I have selected "Keep character formatting" and yet boldface etc. is sometimes lost. Why?**

It is lost because Red Ink in this case not only tries to preserve simple character formatting such as bold, but also the paragraph formatting. This can, depending on how they are defined, override character formatting again. In a first step, Red Ink bolds a word as expected, but this boldface setting is then cancelled when Red Ink assigns the previous formatting to the paragraph in a second step, provided such paragraph formatting contains a corresponding instruction regarding the character formatting. This sequence unfortunately cannot be easily reversed. At best, repeated "Ctrl-Z" can undo the assignment of the paragraph formatting. It should also be noted that the Keep character formatting function is only executed up to the defined number of characters in order not to cause excessively long waiting times.

6. **How can I undo a Red Ink replacement or insertion?**

In Word and Outlook, this is possible using the classic Undo function. However, this may have to be performed multiple times, as Red Ink replaces the content in several steps and also uses several steps for adjusting the formatting (press Ctrl-Z several times!). Excel, on the other hand, does not support the use of the Undo function for adjustments to cells made by Red Ink. Therefore, Red Ink has a dedicated function for this in the Red Ink menu (Undo Last Insert). However, once executed, the AI-generated content cannot be retrieved. This would therefore have to be saved separately before the Undo Last Insert function is used.

7. **Does Red Ink also support source citations?**

Yes, if the AI's response provides citations and these are marked as "citations" in JSON format (the format used internally for responses) (possibly also as a "url" object), then Red Ink will append these citations at the end of the text. This can be used, for example, with Perplexity's models, which provide internet sources.

8. **We occasionally get the error 429. What does that mean?**

This error occurs when too much text per second is sent to the server. It then blocks. This error can occur if multiple people use Red Ink through the same account at the same time or if one person makes multiple queries in quick succession. Red Ink is programmed to wait initially when this error message appears and then to try again up to two more times. The error message only appears if that doesn't help either. You will then have to wait and retry later or increase the waiting time between two requests.

9. **Suddenly, Red Ink's AI functions no longer work. What could be the reason?**

This can have various causes, of course, but if it's not the AI server (temporary unavailability), then one possibility is that Red Ink has lost its configuration parameters from its memory due to a previous error. Although the add-in is programmed to automatically reload these in certain cases, this is not always possible. The easiest way is to close the current Office program completely (i.e., with Word, for example, all open documents) and restart it. This will reinitialize Red Ink. If this does not help either, you should first check whether the latest version is installed (use the Edge browser to go to <https://apps.vischer.com/> and press the red buttons). If this also does not help, then uninstall Red Ink and reinstall it.

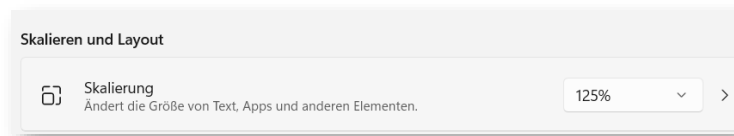
10. **Why does Red Ink sometimes not follow the length specifications for abridgements and summaries?**

Language models sometimes struggle to grasp the length of a text, especially when counted by characters. The reason is that language models themselves do not work with characters but with so-called to-

kens, which are more like syllables or words. We therefore specify abbreviations in words, which language models can handle better. But unfortunately, this is not very reliable.

11. **Why can't I see all of the writing in the dialogue boxes of Red Ink or why is it displayed strangely?**

This can happen if a high value is selected in the display settings for enlarging the font (e.g. 175% or 200%). Also, special resolution settings can have an impact. In this case, Windows overrides the font size that Red Ink selects and you may not be able to see everything as usual. Let us know if this happens, and we'll try to find a solution.

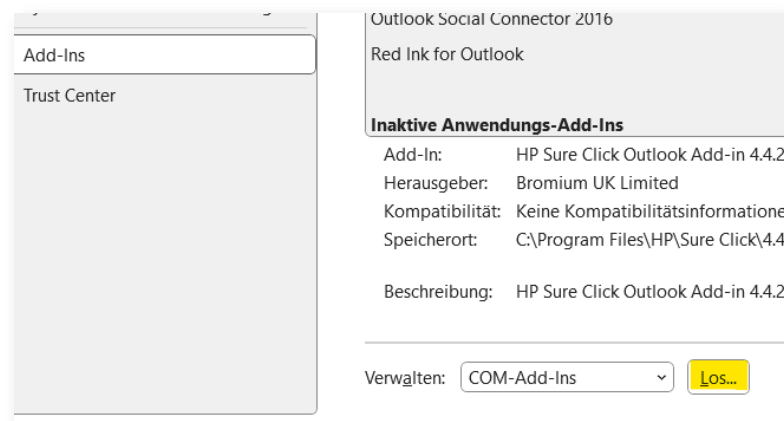


12. **Why when I am using markups in word do new windows suddenly open as if by magic?**

This happens when the Word-internal comparison function is used to create markups. It works by temporarily creating and opening three documents (hiding them would cause an error). That's why they can be seen briefly. However, this is generally not a problem as long as the process is not interrupted. We have also noticed that certain other add-ins can do this. For example, the document management system "iManage" inserts itself into the process without being asked and asks whether the temporary files should be saved, even though Word is instructed by Red Ink not to ask. In our view, this is a bug in iManage, but so far we have not been able to work around it and iManage has not corrected it. If this bothers you, you can try the other markup method(s).

13. **I have installed Red Ink for Outlook, and it appears briefly during loading, but it does not appear in the program.**

This is usually because Outlook automatically disables add-ins that repeatedly take more than 0.5 to 1 second to load. We try to prevent this with a trick, but unfortunately it does not always work depending on the system. Instead, this measure must be switched off by adjusting the Outlook configuration ("Always enable"). Red Ink can also be reactivated afterwards. To do this, select "Options" (File menu), then the submenu "Add-ins", and at the very bottom the option "Manage" or "Go..." for "COM Add-ins". There you can reactivate Red Ink. Outlook does not need to be restarted for this:



#### 14. **Does Red Ink slow down Word, Excel or Outlook?**

No, in our experience not during normal operation. However, we have noticed in the course of our own development work that Word in particular can become slower over time because the standard format template "normal.dotm" grows over time, presumably because it absorbs some kind of data residue and does not dispose of it properly. In our case, the file is between 100-200 KB, but for some people it has grown to 2-3 MB. This slows down Word enormously. The problem is solved by replacing the "normal.dotm" file with a fresh, clean version (it is usually in the "%AppData%\Microsoft\Templates" directory). Of course, it takes a moment for the add-in and the configuration files to load when starting up (which is what happens in Word and Excel when opening a text). It takes another moment to create the context menu, especially in Word; if you find this takes too long, you can switch it off via a configuration setting.

#### 15. **Where is the Red Ink tile in the Outlook Add-in?**

Even with the add-in installed, it primarily appears when a window for composing an email is opened, i.e. to reply to an email, forward an email and, of course, write a new email. In addition, it must be an HTML or RTF email, i.e. an email that can basically also contain formatting. If you want to edit or comment on a "text only" email using Red Ink, you must first switch to HTML mode, save the email and reopen it. However, in most cases today, people work with HTML emails. We have had a second tile for Red Ink for some time now, which can also be used when browsing emails, but only for the "Sumup" function (when multiple emails are selected). If an email is being edited in the pane and then the Red Ink tile is clicked with a command, the email is transferred to a separate window.

#### 16. **Will the "new" Outlook also be supported?**

No. This is because Microsoft provides much less extensive commands for programming add-ins for the new Outlook. So many of the things that Red Ink does (especially with the texts) will not work in the new Outlook. However, according to previous announcements, the current

version of Outlook will be available until 2029. We ourselves do not plan to migrate to the new Outlook; if we did, various other add-ins would probably no longer work either.

17. **Why is the author stored in Word suddenly "Red Ink"?**

This is due to a malfunction. If an error occurs during the execution of markups by Red Ink (e.g. in the chatbot) so that the function is not terminated correctly, it can occasionally happen that the reset of the author information in Word does not work. In this case, it must be done manually. In newer versions of the software, the author "Red Ink" is no longer used.

18. **In the Transcriptor, I get an error message when using Whisper that certain "native libraries" are missing. What do I need to do?**

These are additional program components that must be copied to the same directory as the Whisper language models (i.e., into a subdirectory "runtimes" there). These program components are included in the installation package. For more information, see para. 86 above.

19. **No transcript is displayed in the Transcriptor, although the Transcriptor starts without an error message.**

This can have various reasons. The most common cause is that the selected audio source did not provide any audio data that could be transcribed. If the correct source is selected and it is configured correctly (e.g., volume), this can be due to the fact, among other things, that the source is exclusively "booked" by another application, such as a video conferencing client. In this case, another microphone must be selected, which, for example, records the room sound if conferencing with a loudspeaker. The "Stereo Mix" or "Stereomix" sound source, which basically combines all sounds processed by the computer, has often proven to be practical. But even this is not reliable. In newer versions of the add-in, for each selectable microphone version, we have also included the option "(plus audio output)"; this instructs the transcriptor to mix the currently selected default audio source with the microphone signal (this can be set via the "Dev" button). Nevertheless, there are situations where no reasonable audio source is available (which is why, for example, in a video conference only one's own voice is transcribed, but not the audio of the others). Another reason is that the waiting time is too short or a model that is too large has been selected. The larger the model, the more time elapses until the transcription appears—and sometimes the model will simply be too large. Typically, transcription on a normal PC only works with small or very small models. However, these have the disadvantage that they are prone to errors or are less good. We recommend using cloud-based models for live transcription.



20. **The podcast I generated is missing a voice or no voice is recorded at all.**

This is probably because the podcast script contains SSML commands that the selected voice(s) do not support. The same can happen with adjustments to the default values for Pitch (0) and Speaking Rate (1). In this case, use only the default values and check the box "No SSML".

21. **The tile with the logo has a delayed reaction – why is that?**

This is because the original, white Windows color scheme is not used (but e.g. "Grey"). Red Ink works, but slower. We recommend using the normal color scheme.

22. **The add-in for Outlook says that only HTML and RTF emails are supported – what should I do?**

In the "Format Text" tab of Outlook, a plain text email can be converted into an HTML email.

23. **Why does only a blue circle appear in the menu beside the tile with the logo instead of the direct access buttons?**

If there are too many tiles in the menu ribbon, the Office applications try to condense the tiles by displaying the circle instead. If you click on it with the mouse, the tile is fully displayed. This compression cannot be prevented by the programming. What helps is to remove unnecessary tiles and move the red ink menus to the left (by customising the ribbon).

24. **Why is Red Ink producing results in the wrong language?**

This is because the language model used does not correctly execute the instructions it is given. Certain language models need more instructions here than others or cannot process multiple instructions. It may help to use the "precorrection" parameter to give the language model additional instructions (e.g. "Your output must always be in German"). In the case of Sum-up within Outlook, we have further observed that the speech recognition is misled by Outlook's default entries (such as "Sender" or "Subject", which always appear in the installation language). Also, with the chatbot, the AI does not always follow the instruction to always respond in the language of the last user command.

25. **I have installed the browser extension but the system is not responding when I select a command.**

This can have several causes. First of all, Outlook with the Red Ink add-in must be loaded and active, as the browser extension depends on it (and on the add-in for Word if the "Send to Word" function is to be used). So, Outlook should be started first. In practice, it can also happen that the window for selecting the language, for example, opens but is hidden by other windows or ignored by the user. Finally, it is also possible that security settings prevent communication between the

browser and Red Ink, or the browser extension is disabled (the ports 12333 and 12334 are used by 127.0.0.1 using the http protocol).

26. **Can Red Ink record what users do with it?**

Red Ink itself does not have this function, i.e. it does not currently have a logging function or audit trail. It is conceivable that such a function could be integrated into the tool, especially for use in organisations, but this would require a central server in the network to which the inputs and outputs can be sent. We can implement such a function if it is desired. However, it should be noted that if more than just the commands are to be logged, a lot of text is generated.

27. **Which prompts does Red Ink use and can I change them?**

Yes. All the prompts used by Red Ink can be read and changed in the configuration file. To do this, open Settings in Red Ink and select "Expert Config". The prompts can then be copied. We have described which prompts are used for what in the advanced configuration (see para. 246 et seq. above). We reserve the right to revise and improve the prompts ourselves from time to time, which is why they are not written to the configuration file, provided they meet the standard. However, reformulated prompts are read from the configuration file. Therefore, anyone who has stored their own prompts there will not benefit from our improvements in these cases until their prompts have been deleted from the configuration file.

28. **Can the installation of the helpers for Word and Excel be automated?**

Yes, this is easily done. The two files "redink\_helper.dotm" and "redink\_helper.xlam" just have to be copied from the installation package into the relevant directories of Word and Excel (see para. 228 et seq. above). Everything else happens automatically when Word and Excel are started. This copying process can be outsourced to a batch file, for example. An example is included in the installation package. Alternatively, the "Install Helper" button in Word or Excel can be used; the files are then downloaded directly from our server and copied to the correct directory, if the security settings allow this.

29. **During the installation of the helper for Excel, Windows tells me that the file contains malware. Is that true?**

This is a false alarm from Windows Defender, which unfortunately can happen again and again and is annoying because Windows then deletes the file. We have reported the problem, but so far it has not been fixed. Sometimes it helps to wait; the false alarm does not always occur, even on the same computer. The problem can be solved locally by adding the file in question to the "white list" or by instructing Windows to allow this alleged "threat". We assume that the false alarm is caused by the fact that our Excel helper contains functions for transmitting data, since it must be able to communicate with the AI.

30. **Why does the installation of Red Ink not work on my office computer?**

This is most likely due to the security settings that many organisations use to prevent their employees from installing software because it may contain malicious code or is otherwise undesirable. In this case, contact your IT department or IT service provider. If they trust us as a source, they can authorise the installation based on our digital certificates or our deployment server, <https://apps.vischer.com>. In addition to the actual add-ins, it is important to ensure that the two helper programmes for Excel and Word can also be installed and run if possible. They contain macros and VBA code for Excel and Word, which is also blocked in some organisations. The same procedure should be followed here. However, the helpers are not essential for the basic functions of Red Ink.

31. **Is Red Ink a secure program?**

Everyone has to judge this for themselves. The source code of Red Ink is openly accessible to anyone who wants to check it, so it is possible to see exactly what the software does and which third-party libraries are used (all of which are also open source). If you want, you can compile and use Red Ink yourself based on the checked source code using the files published on Github. However, this requires programming knowledge. Those who trust us, on the other hand, may simply want to rely on the fact that the files we deliver are digitally signed by us (VISCHER AG).

32. **Does VISCHER collect data about the use of Red Ink?**

No, VISCHER does not collect data on the use of Red Ink. Even the built-in function for licence control consists of a date that the user can configure themselves. The built-in function gives VISCHER *no* feedback. The tool has no internal switch that we can use to deactivate it remotely (but the software has a switch that allows an organisation to code its own copy of the tools so that the tool only runs in its own domain, see para. 285 above). However, Red Ink has a function that regularly automatically checks our server for updates and installs them if possible (whereby no registration is necessary, i.e. we do not identify the persons). Otherwise, Red Ink only accesses the installed AI endpoints and – if activated – the configured search engine. Finally, the add-ins for Word and Outlook have the option of receiving and displaying data via a built-in http listener. We use this as described above for the browser extension.

33. **I don't trust the AI providers in the cloud – can I use Red Ink only locally?**

Yes, you can do that without any problem, provided that a suitable language model is operated on your own system or in your own network. Such language models are available free of charge, but experi-

ence shows that they require a powerful server. The tool can be configured on this endpoint.

34. **I received an error message when using Red Ink that I don't understand – what should I do?**

In addition to the usual error messages that can be traced back to an operating error, the software may also experience internal errors. Please note the situation that caused the error and let us know together with the error message. This helps us to improve the tool. With around 10,000 lines of code, it has now become a more complex programme.

35. **Who developed Red Ink?**

Red Ink was designed, developed and written by David Rosenthal in his spare time. Of course, AI was used, as were various libraries, e.g. for encryption functions, analysing HTML code, reading PDF documents or the diff method. In the original version of the add-ins, however, these were developed in VBA, where no libraries could be used. In this first generation of Red Ink (which is still available), everything had to be programmed, except for the crypto functions, for which an external program was used. When we ported the add-ins to VB.net, we were able to use libraries and no longer needed the external program, but we basically had to rewrite the add-ins.

36. **Why was Red Ink developed by VISCHER, a law firm, of all places?**

We initially did this just for ourselves. There were three reasons for this: firstly, we couldn't find any programmes that could do what Red Ink can do. Secondly, we wanted a solution that we could configure with API access so that we could also use it with documents subject to professional confidentiality. Thirdly, we wanted to save costs because a provider of an AI translation solution has massively increased its prices for 2025. Thanks to Red Ink, costs will now be much lower because the tool is much cheaper to use. Meanwhile, positive feedback and a constant stream of ideas have driven us to make the tool even better and more powerful – and to offer it publicly. Since David Rosenthal worked as a software developer in his younger years and since AI can now do a lot of the supporting work in programming, he was able to implement the project in a few months alongside his normal work.

37. **Why is the tool called 'Red Ink'?**

Internally, the tool was called "Red Dragon" after the nickname that a former employee of David Rosenthal had given him (with affection) because he always returned her draft responses to clients with massive (red) markups. Because this is also one of the tool's functions, we gave the tool this name internally. Of course, it can now do a lot more than just improve texts. However, after we registered a word mark for "RED DRAGON", Microsoft got in touch after a while and prohibited its use on

the grounds that the term was too close to the trademark "DRAGON" of their speech recognition product. We did not believe that there was a risk of confusion, but we did not want to invest our resources in a legal battle. Instead, we wanted to create a better product (also as a co-pilot of Microsoft) and changed the name to "Red Ink" before launching, – the colour of error corrections. The logo, which was also registered as a trademark, shows a mythical creature – a cross between a seahorse, a dragon and a sea monster. We call it "Inky", and you can communicate with it via the chatbot.

38. **Who may use Red Ink and under which conditions?**

This can be seen from the licence conditions we have published (see <https://vischer.com/redink> and <https://apps.vischer.com>). During the public test phase, everyone will be able to use it free of charge. After that, it is planned that personal use of Red Ink Generation 2 will be free of charge. If organisations make it available to their employees, we will require licences for a moderate annual licence fee, depending on the constellation (in self-regulation). We will publish details at <https://vischer.com/redink>. However, in our experience, even when the costs for an API access and possible licence fees are taken into account, the use of Red Ink is still significantly cheaper in many cases than equipping all users with subscriptions from certain commercial providers.



39. **Will Red Ink be developed further?**

Yes, anyone who has additional good ideas about how Red Ink can be enhanced should contact me ([david.rosenthal@vischer.com](mailto:david.rosenthal@vischer.com)). The same applies if you discover any bugs, even though I can't provide any support. I myself keep coming up with new ideas for features that would make our work easier and allow us to make even better use of the possibilities offered by large language models.

40. **Where can Red Ink be obtained?**

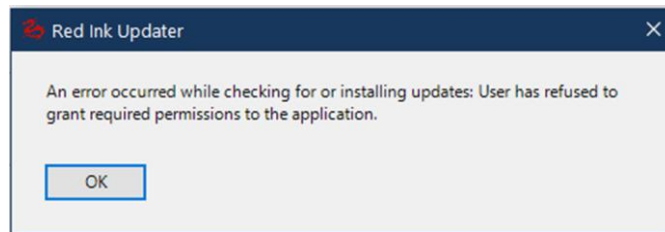
All information can be found at <https://vischer.com/redink> and the files for installation on the deployment server at <https://apps.vischer.com>. The browser extension is distributed through the Microsoft Edge and Google Chrome Add-on-Stores.

41. **Are there updates for Red Ink?**

We offer updates from time to time, but don't guarantee when and how frequently. Red Ink regularly checks for updates itself and installs them if possible. Furthermore, uninstallation and reinstallation is possible at any time. Since the configuration is stored separately, nothing is lost in the process.

42. **I receive an error message from the Red Ink Updater when I start my programs.**

If this error message appears, it is because the installation was faulty or access to the internet is blocked. This error message can also be triggered by Word itself if Red Ink has been installed by another user or as an administrator. In these cases, it helps to uninstall and reinstall Red Ink.



In newer versions of Red Ink, this message is shown once, and the update is only attempted again after a new cycle so that the message does not appear too often.

43. **How can I uninstall Red Ink?**

This is done using the relevant Windows function for uninstalling software. Red Ink appears there separately for each Office application and can be uninstalled within a few seconds. The helpers can simply be deleted. The uninstallation does not remove the configuration files and helpers. These have to be deleted manually if necessary (the locations are described above in para. 228 et seq.). The helpers can also be deleted using the "Remove Helper" function in Settings of Word and Excel (deletion may, however, be blocked by Word or Excel).

**VI. RELEASE NOTES**

287 The following table documents the various works and adjustments to Generation 2 of Red Ink (only Beta Test version or later):

Datum, Build	Release Notes
10.2.2025 W=0.1.1.0 O=0.1.1.0 E=0.1.1.0	Release of Beta Test version
14.2.2025 W=0.1.1.1 O=0.1.1.1 E=0.1.1.1	Correction of the Setup Wizard header information for Google Gemini API; header information is now optional; active deletion of the Interop COM objects in Outlook; more stable securing of the http listener in Outlook; Undo function in Excel; Deselect after insertions in Word; Adjustment of the text in the podcast parameter form
19.2.2025 W=0.1.1.2 O=0.1.1.2 E=0.1.1.2	Support Perplexity citations (when using Sonar etc. as a model) and additional placeholder in API call parameters; fixed bug that caused Word to remain in Extended Selection Mode; fixed bug that prevented Word from always loading the config file (also ensures better loading of the config file in Outlook); fixed bug in Excel menu display; fixed bug in settings display; ad-

VISCHER

	justment to the httpListener in Word (i.e., alignment with Outlook)
22.2.2025 W=0.1.1.3 O=0.1.1.3 E=0.1.1.3	Replacement of the search & replace mechanism in the regex replace functionality, incl. small adjustments also to the chatbot (so that deleted text is no longer replaced, as Word normally does); error messages from bubbles are now inserted into a final comment if the user does not cancel them; Switch Parties and Anonymize now suggest the Regex Replace method; the Regex Cap has been increased to 30k; minor bugfixes
23.2.2025 E=0.1.1.4	Bug fix in the reading Word documents within Excel (API)
25.2.2025 W=0.1.1.4 O=0.1.1.4	Improvements to MarkupRegex; minor bug fixes; improvement of the switch-party prompt; support for more formats of source citations in the LLM's JSON responses
26.2.2025 W=0.1.1.5	Bug fixed with Bubbles; improved Chatbot command implementation
3.3.2025 W=0.1.1.6 E=0.1.1.5	Font size in the chatbot can be adjusted with the mouse; unnecessary blank lines in Excel cell functions removed
10.3.2025 W=0.1.1.7	Improvement of Context Search to find individual instances, including bug fixes.
2.4.2025 W=0.1.1.8 E=0.1.1.6	Markups are no longer displayed as "Red Ink"; improvements to "Bubbles"; if the endpoint is exhausted (429 error) Red Ink will retry up to three times with increasing wait times (5, 10, 30 Seconds).
5.4.2025 W=0.1.1.9 E=0.1.1.7 O=0.1.1.5	Minor corrections and improvements in the chatbot; adjustment of the bubble prompt for better performance; larger prompt box; switching of the updater to direct http call (as the previous function did not work reliably).
7.4.2025 W=0.1.1.10 E=0.1.1.8 O=0.1.1.6	More models to choose from in "Freestyle (2nd)"; introduction of the "Pure:" prefix in Word; "Ctrl-P" to insert the last prompt instead of using the clipboard (in all three add-ins).
7.4.2025 W=0.1.1.11 E=0.1.1.9 O=0.1.1.7	Bug fix concerning the system component (System.ValueTuple).
7.4.2025 W=0.1.1.12	Support for image generation (multimodal Gemini models).
8.4.2025 W=0.1.1.13 E=0.1.1.10	Bug fix regarding Compare Selection Halves; author changed in Word Compare; additional models and "Pure:" also in Excel.
14.4.2025 W=0.1.1.14	Added alternate voices and cleaning of text for "Create Audio".
15.4.2025 W=01.1.15 E=0.1.1.12 O=0.1.1.8	Added "sum-up" for multiple mails. Updated default interval for looking for updates to seven days; fixed bug when automatically inserting content into Excel if such content contained square brackets.
21.4.2025	Excel chatbot; Excel bubbles; Transcription with Google



VISCHER

W=01.1.16 E=0.1.1.14 O=0.1.1.9	Speech-to-Text (V1); direct pasting with the "Clipboard" command; Markdown support in Word expanded; smaller bug fixes in Word Chatbot, in the Transcripator, and in other areas.
22.4.2025 W=01.1.17 E=0.1.1.16	Increased speed of cell data collection in Excel (chatbot, Freestyle); Google Transcription time-limit circumvented; smaller fixes; fontscaling support for various windows (not all) in all add-ins
23.4.2025 W=0.1.1.18 E=0.1.1.17 O=0.1.1.10	Google Transcription time-out problem fixed; bug in cell data collection fixed; turned off Cuda support due to incompatibility
24.4.2025 W=0.1.1.19 E=0.1.1.18	Cosmetic adjustment transcripator; Fix in TTS form; Fix for inserting formulas in Excel
27.4.2025 W=0.1.1.20 E=0.1.1.19 O=0.1.1.11	Reading images, etc. via the trigger "(file)" (Word); cosmetic adjustments (forms)
28.4.2025 W=0.1.1.21 E=0.1.1.21	Transcripator can now also capture signals from the standard output device; fix for TTS dialog; fix for capturing cell contents with more than 255 characters of text
2.5.2025 E=0.1.1.22	Excel Freestyle now also supports including external documents through "{doc}"
3.5.2025 E=0.1.1.23	Excel Freestyle (and the Chatbot, when using selections) now also recognizes drop down options and color codes
4.5.2025 E=0.1.1.24	Bug fix when reading cell content (Excel)
16.5.2025 W=0.1.1.26 E=0.1.1.26 O=0.1.1.13	Interface for external services e.g. for legal information such as lexisearch.ch or DeepL (Word), output of AI results in a pane (Word), possibility to filter invisible characters from AI results ("Clean" parameter), easy selection of the entire document within Freestyle ("all", Word); improvement of the conversion and display of Markdown-formatted texts; improved display of the window with the "Clipboard" command; bug fixes; extension of the Prompt Library
19.5.2025 W=0.1.1.27	When transferring data from browser, the source is mentioned (Word); minor bug fix
20.5.2025 W=0.1.1.28	Fix of new problem inserting text in Word; fix for initializing Red Ink when protected documents are opened
25.5.2025 W=0.1.1.29	Added support for more Special Services, including support for more complex JSON templates; improved and fixed issues pasting Markdown text into Word; hyperlinks from Perplexity etc. are now clickable
27.5.2025 30.5.2025 (github) W=0.1.1.31	Support for complex Special Services (including various templates), support for output in panes in Excel and improved in Word, support for clipboard and file objects in Freestyle also in Excel and partially in Outlook, improvements for podcast and audiobook functions, new transcript prompts, better support for mixing the audio output source in the transcripator, bubbles



VISCHER

	more robust (including ignoring multiple spaces)
2.6.2025 W=0.1.1.32 E=0.1.1.27 O=0.1.1.14	Even more functionality in Special Services (with various further integrations); Outputs in the window can be transferred to the panes (where the "Merge Selection" function or, new in Excel, the selective application of cell commands is also available); local, transparent anonymization option (configurable); vector search and bag-of-words search in the "Context Search" function; standard output device can be selected in the transcriber ("Dev"); Freestyle now also processes binary objects in the clipboard; the bubble prompt in Word has been optimized; new parameters for configuring API calls (double calls and the GET method are now also supported); automatic function for generating "Response" templates; various bug fixes
9.6.2025 W=0.1.1.33 E=0.1.1.28 O=0.1.1.15	The pane now remembers its width; bold, underline, and italics are preserved when no markup (or at most DiffW) is used (can also be switched off); the Word Helper "Clip to Text" added; the automatic updater has been rewritten; sleep is disabled for the transcriber and text-to-speech generation; text-to-speech now also supports OpenAI (OpenAI only, not Azure); various stability measures and bug fixes (e.g., in the display of formatted texts, in the transcriber for service-specific merge prompts; separate OAuth2 tokens for the transcriber and text-to-speech)
9.6.2025 W=0.1.1.34	Added cap for Markdown conversion to prevent hangs
9.6.2025 W=0.1.1.35	Bugfix regarding the conversion of markdown formatting
10.6.2025 W=0.1.1.36 E=0.1.1.29 O=0.1.1.16	Bugfix re processing of clipboard data (that could lead to a crash) and context search; improved bubbles
13.6.2025 W=0.1.1.37 O=0.1.1.17	More stable integration of Edge-Connector; improved initialization when starting Word
15.6.2025 W=0.1.1.38 O=0.1.1.18	Red Ink can now automatically implement Word comments in the document; improved text marking for Bubbles and Context Search; Freestyle (Word) supports "Add: "; Sum-up also available in a separate window when email is open; various improvements
15.6.2025 W=0.1.1.39	Amendment of the text marking for Bubbles and Context Search; add markup choice for Apply comment
15.6.2025 W=0.1.1.40	Amendment of the text marking for Bubbles and Context Search
17.6.2025 W=0.1.1.41	Amendment of the text marking for Bubbles and Context Search
19.6.2025 W=0.1.1.42	Amendment of the text marking for Bubbles and Context Search; bug fixes
22.6.2025	Support for longer texts; bug fixes for Edge integration; bug fix error messages Freestyle; new transcript prompt; count-

VISCHER

W=0.1.1.43 O=0.1.1.19 E=0.1.1.30	down while waiting for LLM; added Query Assistant for Special Services; more Markdown support in Outlook (e.g., lists, tables)
23.6.2025  W=0.1.1.44	Bug fix re HTML Insert
27.6.2025  W=0.1.1.45	Bug fix re HTML Insert, added Debugging Help (output of LLM response)
30.6.2025  W=0.1.1.46 O=0.1.1.20 E=0.1.1.31	Improved HTML Insert, added "(iterate)" feature in Freestyle; improved formatting preserving feature incl. bug fix
7.6.2025  W=0.1.1.47 O=0.1.1.21	Support also for editing texts in footnotes, headers and footers; bug fixes; extension of parameter functionality for special services; support for editing texts in table columns only; translation in Outlook now works also on closed mails
8.7.2025  W=0.1.1.48	Bug fix when maintaining formatting (Word)
13.7.2025  W=0.1.1.49 O=0.1.1.22 E=0.1.1.31	Models can now also be configured for sending multipart requests (e.g. OpenAI image editing); Gemini search grounding citations are now supported; the diff markup function has been significantly improved and is faster (Word); various improvements and bug fixes (including for Iterate); the last Freestyle command can be repeated with a single click (Word)
14.7.2025 15.7.2025  W=0.1.1.51 E=0.1.1.32	Chatbot now also always sees the selection (Word, Excel); Bug Fixes
16.7.2025  W=0.1.1.52 O=0.1.1.23 E=0.1.1.33	Outlook now also supports alternative models; Bug Fix (Pane)
20.7.2025  W=0.1.1.53 O=0.1.1.24 E=0.1.1.34	New Markdown Parser for Word (more support, bug fixes); improved translate/sumup in Outlook with several windows open; new search config for OpenAI; bug fixes
23.7.2025 24.7.2025  W=0.1.1.54 W=0.1.1.55 O=0.1.1.25 E=0.1.1.35	Direct AI output to a new document ("Newdoc:"); Bug Fixes; Expansion of support for special services (for FragdenOK.ch, Entscheidsuche.ch)
27.7.2025  W=0.1.1.56 O=0.1.1.26 E=0.1.1.36	Clipboard-to-Text also in Outlook; Freestyle in Excel can now access multiple worksheets; improvement of the Copy-to-Clipboard functions; Bug Fixes
28.7.2025	Bug Fix; Improvement in Diff Markup (Word)

VISCHER

W=0.1.1.57	
1.8.2025 W=0.1.1.58 O=0.1.1.27 E=0.1.1.37	Output of formatted texts completely revised again and switched to a new system (Word, Outlook); Excel chatbot also supports "(addws)"; Slides function in Word; Token log
6.8.2025 W=0.1.1.59	Slides function extended (graphics, icons); Create Audio now also supports the dubbing of PowerPoint slides

**VII. ROADMAP**

288 Further planned developments:

- Option to select fast, inexpensive models for simple queries (translations and corrections)
- Integrating CustomGPTs
- Automatic Regex conversions of content marked out on websites when this is transferred to Word, including the option of installing Python scripts in Red Ink for this purpose
- Sample prompt for commenting on contract markups
- Sample prompt for extracting important events
- KeepFormat also for Freestyle in Outlook
- Internet access also for the chatbot

## **ANNEX 1: IDEAS TO GET TO KNOW RED INK**

(These ideas are written with lawyers in mind; of course, other texts can be used as contracts.)

- Open a blank Word document, write a few words of text and then right-click or go to the tile and select the translation function and see what happens.
- Highlight your text again and go to Freestyle. A text box for prompting opens. Type in: "Make that into a ten-line poem" and click OK.
- Open a contract and select a clause. Go to Freestyle again, but this time click OK without entering a prompt. A new window will open where you can select a predefined prompt. Select the first one, "Challenger", and see what happens.
- Select an entire contract then select Freestyle and enter the prompt: "Clipboard: What does the cancellation clause say?" and press Enter. The prefix "Clipboard:" means that the answer will appear in a window and in the clipboard instead of in the document.
- Now go to the top of the document and use Context Search and enter a search term related to a topic that has a clause but doesn't include the search term, e.g. "damages", when the clause only refers to "liability". Run a search for "indemnity" based on context – and should it show you the first text passage that is related to it. Try again but add "(m)" (for all matches = multiple) and "(all)" to search the entire document. The passages will be highlighted in yellow.
- Now select the entire text or nothing at all and open Freestyle. Enter "Bubbles: Which of the clauses contain unclear provisions and why?" After a certain processing time, the tool should have provided all clauses with a comment that answers the question.
- Adjust the content of one or two clauses in "Track Changes" mode, then select the entire text and enter the prompt "Clipboard: Explain the changes made to the text. (rev)" in Freestyle. It will summarise them for you.
- In the same document, highlight all of the text again and select Summarize. In the window that pops up, confirm the number or enter 200, for example, and press OK. A summary will appear at the end of the contract.

- Next, copy the content of an email or another document from Outlook or a passage of text from a new document that also contains names and company designations. Select the text and click Anonymize on the Red Ink tile.

ATTENTION: Depending on the settings, the Word Compare function will now be activated to create a markup. Various windows will open and close. An anonymised markup of your text should then appear.

- Now open a longer email chain in Outlook. Click on "Forward" or "Reply" and make sure that the email is open in a separate window for writing. The Red Ink logo should appear in the menu ribbon there, along with another tile with three buttons (if only a circle appears, then make the window wider).
- Do not select anything yet. Just click on the Sum-up button and wait. A summary of the email chain will appear shortly.
- Then delete this summary and write a text or two with two or three sentences with spelling mistakes or extra words. Select the text and choose Correct. Your text will appear in corrected form. Optionally, a markup can be added, but this takes some time (use Settings, then "Save Configuration"). If the markup with "Diff" takes too long, it can simply be cancelled with "Esc". For larger texts, the Word markup method is more suitable. Alternatively we recommend "DiffW".
- Finally, go back to Word and open a document and also the chatbot. Now ask the chatbot to change something in your document. To do this, "grant write access" must be selected, along with the box to the right of it, to give the chatbot access to the entire document. Example: "Please add a second paragraph with another line of thought that fits with the topic."

## ANNEX 2: PROGRAMMING RESPONSE TEMPLATES

*Note: Automatic programming of templates based on a sample JSON string and a description of the output is possible with the freestyle command "generateresponsetemplate".*

### Overview

Templates make it possible to convert JSON data into text documents. Placeholders and special instructions are used to display values from different paths in the JSON document, run through lists and format content.

### Basic placeholders

A placeholder is a text pattern in curly brackets that is replaced by a JSON value. Syntax: {path}

- Paths follow the JSONPath concept. Examples:
  - Simple field access: {user.name} accesses the 'name' field of the 'user' object.
  - Access to nested fields: {order.address.location} accesses 'location' in 'address' of 'order'.
- Paths can be specified with or without a leading '\$'. Both are equivalent:
  - {\$.user.age} or {user.age}

If the path is not found, the output remains empty.

### JSONPath basics

JSONPath is a way to select specific values from a JSON document. Some important rules:

- Objects are addressed by their name: 'objectName.fieldName'.
- Arrays are addressed by square brackets and index or filter: 'article[0]' for the first element.
- Placeholders in templates support simple paths and array passes (see chapter 4).

Examples:

- {products[0].price} shows the price of the first product.
- {category.name} shows the name of the category.

### Loops for lists

Loops can be used to repeatedly output the same template segment for all elements of a list.

Syntax:

```
{% for listenPath %}
```

Content with placeholders, e.g. {field1}, {field2} etc.

```
{% endfor %}
```

Explanation:

- 'listenPath' is a path that refers to an array or an object.
- Within a loop, the specified section is replaced for each element. Placeholders within refer to the current element.
- Example: `{% for article %}Article: {name} - price: {price}{% endfor %}` runs through the array under 'article' and outputs the name and price for each entry.
- If a single object is found under the path, it is also output once.

### Modifiers for placeholders

Placeholders can be customized using prefixes:

- `html:` Converts HTML content into Markdown. Example: `{html:description}`.
- `nocr:` Removes all line breaks and replaces them with spaces. Example: `{nocr:comment}`.
- `htmlnocr:` Combines both effects. Example: `{htmlnocr:description}`.

The order does not matter, upper/lower case is ignored.

### Set separators

For paths that return several values (e.g. several fields with the same name), a separate separator can be specified. Syntax: `{path|separator}`

Example:

- `{tags|; }` combines several keywords with a semicolon and space.
- `{products[?(@.available==true)].name|, }` combines all names of available products with commas and spaces.

By default, values are separated by a line break.

### Mapping definitions

A mapping can be used to replace certain output values. Syntax: `{path|key1=text1;key2=text2;...}`

Explanation:

- Each key = original value from JSON, which is replaced by Text1, Text2 etc.
- Example: `{status|0=Inactive;1=Active;2=Locked}` replaces 0 with 'Inactive', 1 with 'Active', 2 with 'Locked'.
- If no mapping rules are applied or the JSON value does not match any key, the original value is output.



## Line breaks and special characters

The following placeholders can be used to set fixed line breaks within the template:

- `\N` inserts a line break at this point.
- `\n` or `\r` or `\R` also works.
- `<cr>` (case-insensitive) is replaced by a line break.

Example:

Text before line break `\N` Text after line break.

## Examples

### Example 1: Simple placeholder

If the JSON object contains a 'user' key with a 'name' field, this is added to the template:

```
{user.name}
```

and only receives the name.

### Example 2: List of entries with loop and separator

```
{% for article %}
```

```
Item: {name} - Price: {price}
```

```
{% endfor %}
```

If there are several articles, each is output in a new line. If a comma is to be used as a separator between articles, this is used, for example:

```
{% for article %}{name} - {price}{% if not for_last %}, {% endif %}{% endfor %}
```

(There is an implicit placeholder 'for\_last', which is true for the last iteration).

### Example 3: Mapping with user-defined texts

Assuming that the JSON field 'status' can have values 0, 1 or 2, this is included in the template:

```
{status|0=Inactive;1=Active;2=Locked}
```

and receives the corresponding text depending on the status.

