1 INTRODUCTION

1.1   `Instance BETA` is a simple VSTO plug-in for Microsoft Word 365 which functions as a contract navigation and drafting assistant. It is designed for use by contract lawyers and anyone else who spends time editing contracts drafted using Microsoft Word. `Instance` enables rapid analysis and navigation of complex contracts where keeping track of defined terms can be difficult.

1.2   The principal function of `Instance` is to extract the text of the currently open document, perform some simple analysis on the extracted text, and use the returned data to populate a sorted, navigable, tree structure. The tree structure comprises the following principal nodes:

- An ordered list of the paragraphs of the contract, with the main body paragraphs collected together under numbered heading nodes, enabling the display of individual paragraphs or custom groups of (not necessarily contiguous) paragraphs.
- A sorted list of the contract's defined terms, together with sub-nodes corresponding to each paragraph of the contract in which one or more instances of each defined term occurs, all with optional defined term highlighting.
- A list of paragraphs corresponding to words and phrases indicative of contractual risk, initially confined to indemnity and governing law / jurisdiction risks, with key word highlighting.

1.3   The emphasis is on speed. `Instance` typically completes the analysis of the contract text and populates the tree in well under 50ms, often less than 1ms. `Instance` does not feature any AI, machine learning, or other "smart" technologies. Its functions are based on pure text search, inference and heuristics, rendered in low-level native binary code. The author is committed to improving the accuracy of these functions. However, where any candidate improvement that provides a marginal increase in accuracy requires any substantial increase in the time of response, it is unlikely to be implemented.

2 GETTING STARTED

2.1   The latest version of `Instance` is always hosted at the following link:

https://github.com/CambridgeStateMachines/Instance/blob/main/Instance.zip

2.2   The latest version of the documentation for **Instance** is always hosted at the following link:

https://github.com/CambridgeStateMachines/Instance/blob/main/Instance.pdf

2.3   The downloaded file, **Instance.zip**, contains a standard VSTO installer application and associated program files. Simply unzip **Instance.zip** and double click the **setup.exe** installer application. Once **Instance** is installed, open Microsoft Word. The main **Instance** window looks like this, note the main button is coloured red:
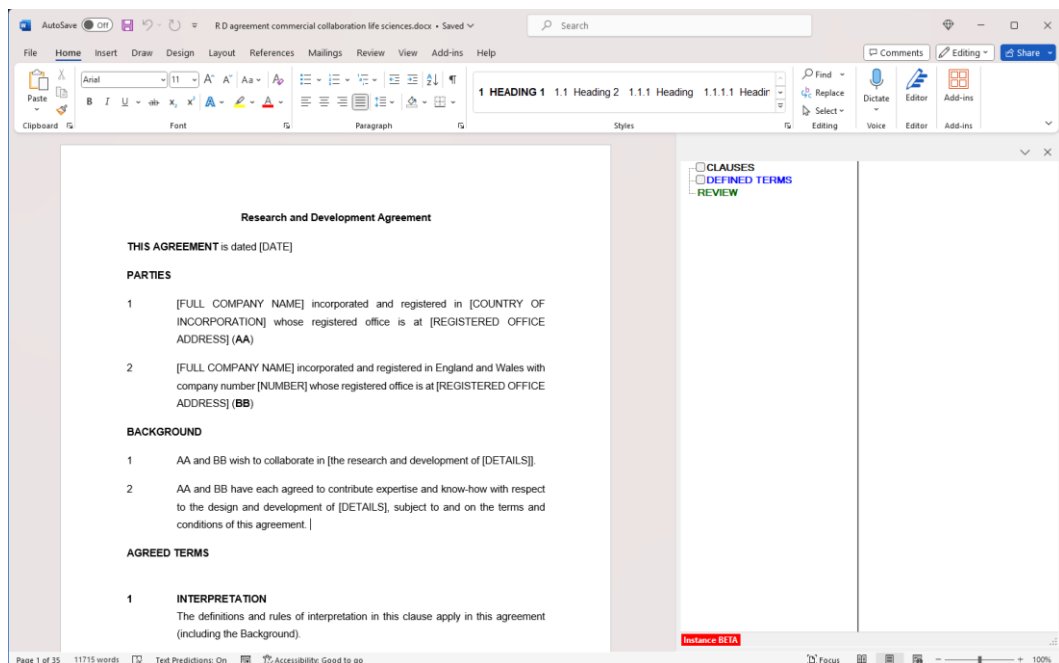

2.4   Initially, the tree has three main nodes labelled "CLAUSES", "DEFINED TERMS" and "REVIEW", with no sub-nodes, and the text display is clear. You can resize the main window by dragging the left-hand edge of the window. You can move the divider between the tree and the text display by dragging it with the mouse, or by pressing the "tab" key until the divider is selected and using the left and right keyboard arrows.

3   USING INSTANCE

3.1   When you are ready to process the document, click the button labelled "Instance BETA". After a short time, the display will update with "+" symbols at the root of

each main node indicating that they are now populated with sub-nodes. Immediately after processing is complete, the display should look like this:
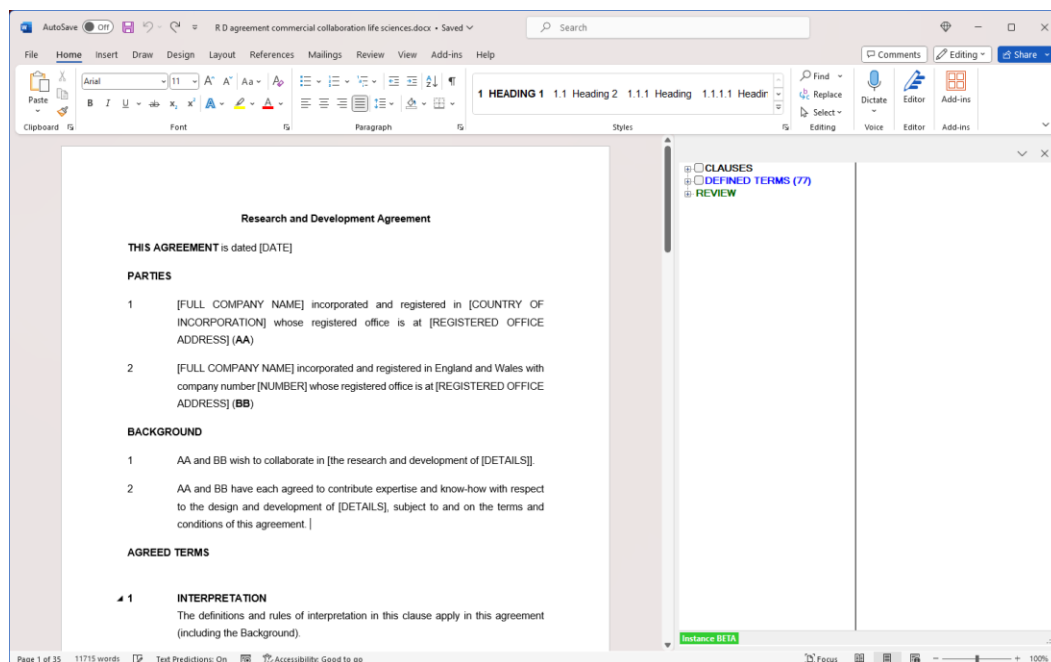
3.2   The main button turns green, indicating that the current version of the document's text has been processed. If the text of the document changes, the button will turn red again, indicating that the analysis should be rerun.

3.3   Clicking the "+" symbol at the root of a node (or selecting the node and pressing the right arrow key on the keyboard) will expand the node to reveal its sub-nodes.

3.4   The nodes labelled "<<" and ">>" contain sub-nodes corresponding to the paragraphs that appear in the document, respectively, before and after the paragraphs that **Instance** has identified as making up the main body of the contract.

3.5   Instance identifies defined terms two ways: (a) by searching for strings located between paired quotation marks; and (b) by searching for strings that are formatted as bold text. If you prefer **Instance** to only include quoted strings in the analysis, simply Shift-click the main button. This will speed-up the analysis slightly as well as reducing the number of false positives e.g. numbered headings interpreted as defined terms.

3.6   Where a defined term is used only once in the document, within its own definition, the corresponding node colour is red. The total number of defined terms is

displayed in brackets after the main DEFINED TERMS node's title. The total number of instances of each defined term is shown in brackets after the defined term's text.

3.7   You can navigate the tree by clicking nodes with the mouse or by using the keyboard arrow keys. The up and down arrow keys move the selection between adjacent nodes. The left and right arrow keys respectively expand and contract nodes that have sub-nodes (indicated by the presence of the "+" symbol at the node's root).

3.8   The sub-nodes under the CLAUSES node can be navigated by typing numbers corresponding to the clause numbers. Likewise, the sub-nodes under the DEFINED TERMS node can be navigated by typing letters corresponding to the initial letters of the defined terms. Repeated typing of the same letter or number causes successive nodes to be selected.

3.9   Selecting any single node that is associated with a contract paragraph will cause the paragraph's text to be displayed in the text window. Selecting any node whose checkbox is checked will cause not only the corresponding paragraph to be displayed, but also the text of any other paragraphs whose corresponding node's checkbox is also checked:
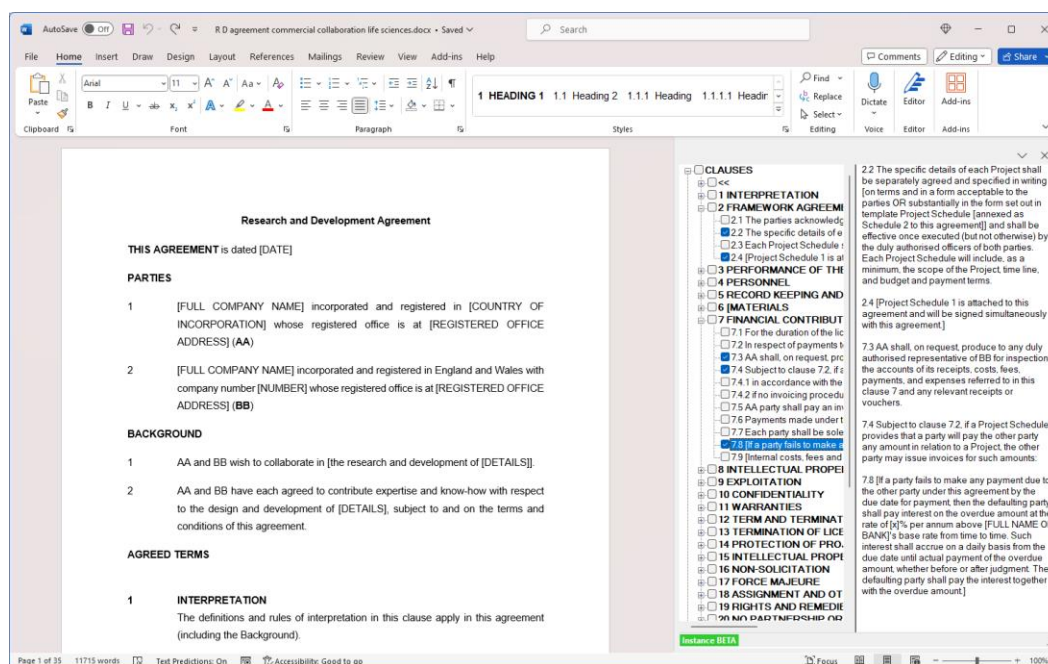

This feature enables customised collections of individual paragraphs to be displayed together.

3.10 Checking (or unchecking) the checkbox of an individual node will cause the text display to be updated accordingly. Checking (or unchecking) the checkbox of any node that has sub-nodes will cause the checkbox of all such sub-nodes (and their sub-nodes) to also be checked (or unchecked). This feature enables rapid checking and unchecking of nodes.

3.11 Highlighting of defined term instances in the displayed text is turned on and off by right-clicking the node corresponding to the defined term (this does not change the currently selected node). The currently displayed text is updated with each instance of the relevant highlighted defined term highlighted in yellow:
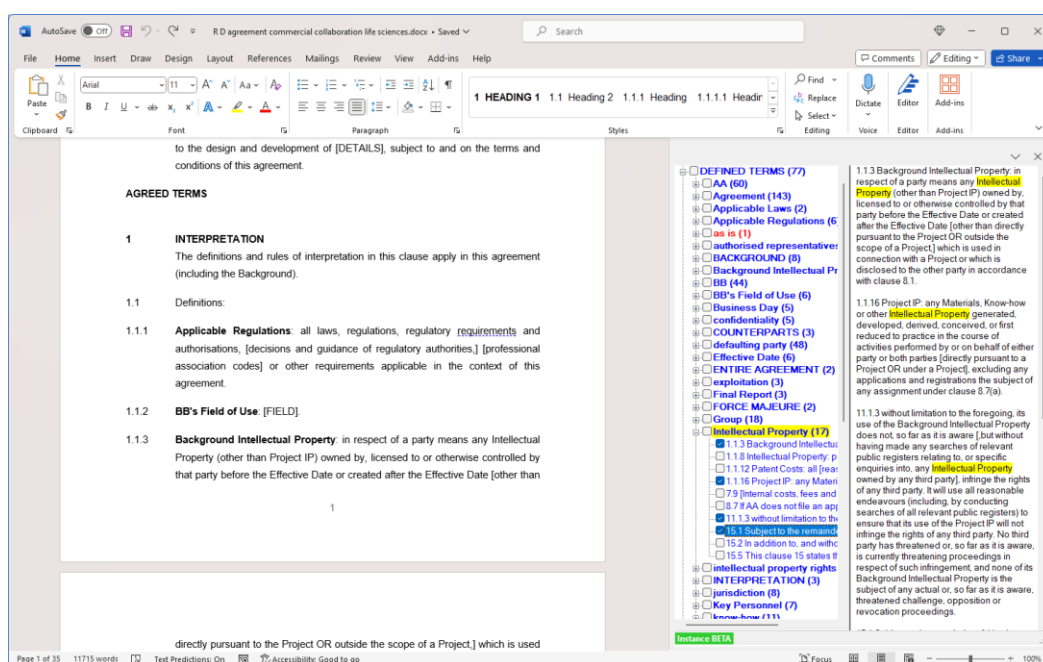
Note that, where a large number of paragraphs are shown and/or a large number of defined terms are highlighted, the display can take a long time to update. A busy cursor will be displayed while the display updates. You can highlight (or unhighlight) all of the defined terms by right clicking on the DEFINED TERMS node. This feature enables rapid highlighting and unhighlighting of defined terms.

3.12 When you click the sub-nodes of risk items appearing under the REVIEW node the text of the relevant paragraphs, as identified by **Instance** based on various key word search heuristics, is displayed:
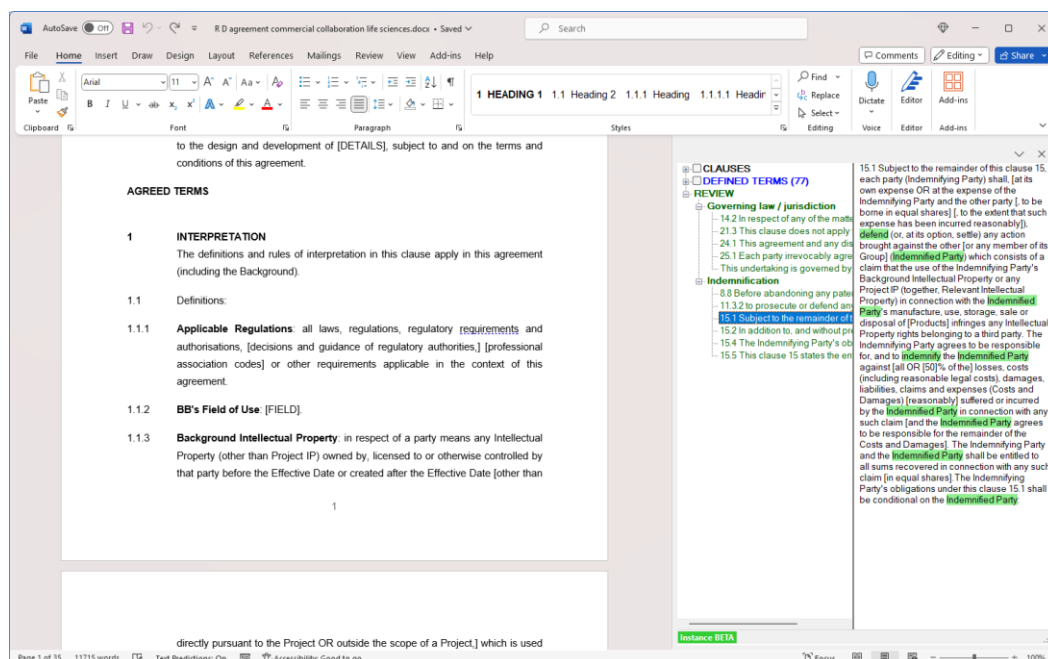
The specific text on which the assessment of relevance was based is highlighted in pale green. This allows for a rapid assessment of whether the assessment is correct or a false positive.

3.13   If you shift-click on the sub-node of a clause, a defined term, or a risk item, the corresponding paragraph of the Word document will be scrolled into view.

4   TROUBLESHOOTING AND KNOWN ISSUES

4.1   **Instance** is not intelligent. It relies on basic plain text search, inference and heuristics. **Instance** will not recognise non-standard words, such as company names, as defined terms unless they are formally defined between paired quotation marks or appear in bold text. If **Instance** fails to recognise ABC as a party name, change the definition text in the Microsoft Word document to "ABC" or **ABC**.

4.2   You may find that **Instance** mis-identifies the paragraphs that it considers constitute the main body of the contract. **Instance** uses basic heuristics to identify the main body of the contract. One such heuristic is to identify the main body of the contract as comprising the set of contiguously numbered paragraphs (starting with 1) containing the most characters. If, in the true main body of the contract, there is a discontinuity in clause numbering, **Instance** will designate the next longest contiguously numbered set of paragraphs as the main body, resulting in an erroneous identification. The solution is to carefully check the numbering in the true

main body of the contract and correct any discontinuities in the top two levels of clause number, i.e. heading levels 1 and 2.

5   FURTHER INFORMATION

5.1   `Instance` is a standard VSTO add-in for Microsoft Word written in C#. The add-in calls some optimised text-processing functions, located in a separate DLL written in C. No third party code is incorporated in `Instance`. The C# code was written and compiled using Visual Studio 2022 Community. The C code was written using Visual Studio Code and compiled with GCC via the MinGW-w64 toolchain. Bounds-checking versions of standard C functions are used in the DLL and the C source code has been extensively reviewed in order to eliminate memory leak risks

5.2   If you have any questions, suggestions, bug reports or general comments, please feel free to email the author: **info@cambridgestatemachines.co.uk**