

RIKEN Wex (hereafter, Wex) is a tool to extract edit histories and/or comments in the MS Word format and convert them to XML/TSV. What makes Wex (and the MS Word format) different from most annotation tools is that it makes it possible to annotate arbitrary labels or contents to a given document; note that most existing annotation tools only allow the annotator to select one from pre-set labels. While the conventional tools are suitable for grammatical annotations including POS-tagging, syntactical annotation, NER, and semantic role labelling, they are insufficient for more flexible annotations. One of the major examples is grammatical error annotation where one can annotate a grammatical error with an arbitrary piece of text (correction). Another would be comment annotation such as feedback comments given to learners of a language.

RIKEN Wex（以下、Wex）は、MS Word ファイル形式で記録された編集履歴とコメントを抽出し、XML/TSV 形式に変更するツールです。他の多くのアノテーションツールと異なり、Wex を用いると任意の内容を簡便に付与することが可能となります（一般的なアノテーションツールは、事前に決められたラベルの集合から選択してアノテーションを行うことが多い）。従来のアノテーションツールは、品詞、構文、固有名詞、意味役割などの情報を付与するのには適していますが、ラベルが事前に決められないようなアノテーションタスクには向きません。そのような例に、文法誤りのアノテーションがあります（文法誤りに対する訂正情報は、事前に決めることが困難）。また、書き手へのコメントなどの情報付与も含まれます。