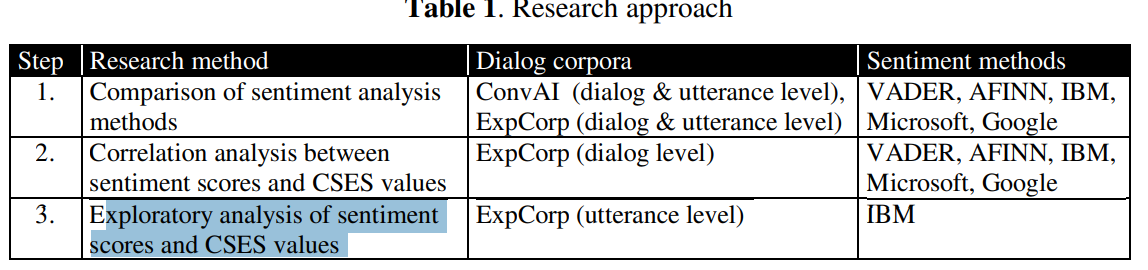
* + investigate the application of automated sentiment analysis methods as a proxy to measure CSES. Therefore,
    - we first compare different sentiment analysis methods.
    - Second, we investigate the relationship between objectively computed sentiment scores of dialogs and subjectively measured CSES values.
    - Third, we evaluate whether this relationship also exists for utterance sequences throughout the dialog. The paper contributes by proposing and applying an automatic and objective approach to use sentiment scores as a proxy to measure CSES

1. Intro
   * Users who are less happy with a chatbot use less assent, fewer positive, and more anger-related words and thus, express more negative sentiments [16].
2. Sentiment Analysis Method
   1. semantic
      1. Lexical (Best VADER and AFINN)
      2. morpho-syntactic text analyses (i.e., wordform, lemma, part of speech tags)
   2. statistical (support vector machines and Bayes classifiers) : require labeled data for training purposes (Best IBM Watson, Google Cloud, and Microsoft Azure )
      1. unsupervised ML
      2. supervised ML
3. Research
   1. benchmark analyses reveal that there is no superior sentiment analysis method because all tools perform differently depending on the specific context they are applied on or depending on the corresponding data source on which they were trained
   2. both benchmarks reveal several suitable methods depending on the respective context and the training data
4. Research Approach



* 1. Step 1: Sentiment analysis result on large text of data showed correlation among different methods as compared to short sentences. Different methods interpreted shorter sentences differently.
  2. Step 2: By doing this, we aimed to reveal whether sentiment scores are a valid proxy for CSES values
  3. Step 3: investigated the minimum number of utterances required to show a correlation between sentiment scores and CSES values
     1. Seq of 5 utterances have shown good correlation with Doc CSE