VAI Challenge

You have been tasked to build an API where a user can query the complexity of a text segment. In the first iteration your PM asked you to use lexical density. Please use NodeJs + Express (or any of the more well known frameworks) to develop your API.

Definitions

Lexical density is defined as the number of **lexical words** (or content words) divided by the total number of words. In the following sentence the green words are lexical words and the density is 67%.

Kim loves going to the cinema

For the sake of simplicity, we define a **lexical word** as all words not contained in the provided list of non lexical words in the Appendix. Case sensitivity should be ignored.

Requirements

- **Route:** /complexity
  - Description:
    Return the lexical density of the inputted text. The input text should be provided via the request.
  - output :

```
{  "data":{
     overall_ld: 0.42
     }
}
```

- **Route:** /complexity?mode=verbose
  - Description:
    Return the lexical density of the text broken down into sentences. The input text should be provided via the request.
  - output :

```
{  "data":{
     sentence_ld: [ 0.23, 0.1, 1.0, 0.0],
     overall_ld: 0.42
     }
}
```

- **Error case:** Only texts with up to 100 words or up to 1000 characters are valid input. Please cover these cases with tests using the framework of your choice.

- **Storage:** The provided **non-lexical words** should be stored in a Mongo DB. If time allows, please provide a protected endpoint where new words can be added over time.

Delivery:

- You should provide a link to a public github with the solution, incl. a self sufficient Readme.
- You have 2 hours to deliver your code after the challenge has started. If you feel you are running out of time, pseudo code is a good option.
- We are going to evaluate code standards, code structure and creativity of the solution.
- You are free to use any other additional libraries

Appendix: Non lexical words

to
got
is
have
and
although
or
that
when
while
a
either
more
much
neither
my
that
the
as
no
nor
not
at
between
in
of
without
I
you
he
she
it
we
they
anybody
one