



# **Text Similarity function plug-in**

**Version 1.0.3**

## **Table of Contents**

<b>Overview</b>	<b>3</b>
<b>Key Features &amp; Functionality</b>	<b>3</b>
<b>Sample usage</b>	<b>3</b>
<b>Notes &amp; Known Issues</b>	<b>4</b>
<b>Feedback</b>	<b>4</b>



## Overview

The Text Similarity function plug-in allows users to compute a similarity score between pairs of texts. By computing a similarity score, one can find related items, group similar items, detect duplicates, and more.

## Key Features & Functionality

Computes a similarity score between pairs of texts, returning a value between 0.0 and 1.0, where a higher score implies a higher similarity.

## Sample usage

After deploying your plug-in, create an expression rule and enter the following:

```
a!fromJson(fn!similarity({  
  "The beach is fun",  
  "I like going to the beach",  
  "She has a new car",  
  "Let's go to the mountains",  
  "Did she get a new car?",  
}))
```

Execute the rule and you will get a list of lists that will contain values like the following:

0	0.8818402	0.55072	0.6452242	0.5068361
0	0	0.5662659	0.6885484	0.5229815
0	0	0	0.5385293	0.9165229
0	0	0	0	0.5093196
0	0	0	0	0

Each row and column represents the index of the string. For example, row 0 represents the similarity of the string "The beach is fun" with every other string in the array. The second column represents the string "I like going to the beach", and so, the value 0.8818402 is the similarity score between "The beach is fun" and "I like going to the beach". A value of 0.8818402 is high and implies a higher similarity.

The value 0.5093196 is the similarity score between the strings "Let's go to the mountains" and "Did she get a new car?" implying there's not much similarity between them.

## Notes & Known Issues

- The computation of the similarity scores is CPU intensive. Computing the similarity of 300 strings can take between 10 to 40 seconds depending on the instance size. The plug-in will accept no more than 300 strings to avoid a large performance problem in the server.
- If the plug-in is undeployed, in order to deploy it again, the app server needs to be restarted. After removing the plug-in, restart the app server, and then deploy the plug-in again.

## Feedback

Please feel free to reach out with any feedback, questions, or comments to [carlos.aguayo@appian.com](mailto:carlos.aguayo@appian.com)