

7. Ngram Cosine Similarity

```
====> On File: Libpng
Command Running: atarashi -a Ngram -s CosineSim TestFiles/Libpng.c
The Obtained result by agent is: Libpng
Total files scanned = 100
Successfully matched = 59

100%|████████████████████████████████████████████████████████████████████████████████| 100/100 [20:02<00:00, 12.02s/files]

+++++ Result +++++
---> Total time elapsed: 1202.27 Seconds <---
---> Accuracy: 59.0% <---
```

8. Ngram Dice Similarity

```
====> On File: Libpng
Command Running: atarashi -a Ngram -s DiceSim TestFiles/Libpng.c
The Obtained result by agent is: Libpng
Total files scanned = 100
Successfully matched = 58

100%|████████████████████████████████████████████████████████████████████████████████| 100/100 [19:23<00:00, 11.63s/files]

+++++ Result +++++
---> Total time elapsed: 1163.3 Seconds <---
---> Accuracy: 58.0% <---
```

9. Ngram Bigram Cosine Similarity

```
====> On File: Libpng
Command Running: atarashi -a Ngram -s BigramCosineSim TestFiles/Libpng.c
The Obtained result by agent is: Libpng
Total files scanned = 100
Successfully matched = 58

100%|████████████████████████████████████████████████████████████████████████████████| 100/100 [24:32<00:00, 14.73s/files]

+++++ Result +++++
---> Total time elapsed: 1472.71 Seconds <---
---> Accuracy: 58.0% <---
```

10. DLD

```
====> On File: Libpng
Command Running: atarashi -a DLD TestFiles/Libpng.c
The Obtained result by agent is: Libpng
Total files scanned = 100
Successfully matched = 49

100%|████████████████████████████████████████████████████████████████████████████████| 100/100 [51:52<00:00, 31.13s/files]

+++++++ Result ++++++
+++++++
--> Total time elapsed: 3112.86 Seconds <---
--> Accuracy: 49.0% <---
+++++++
+++++++
```

<u>Algorithms</u>	<u>Time (in Seconds)</u>	<u>Accuracy</u>
semanticTextSim	113	1%
wordFrequencySimilarity	129	51%
tfidf	1440	58%
tfidf CosineSim	1426	58%
tfidf ScoreSim	159	0%*
Ngram	1204	59%
Ngram CosineSim	1202	59%
Ngram DiceSim	1163	58%
Ngram BigramCosineSim	1473	58%
DLD	3113	49%

* tfidf ScoreSim returns multiple results and the first one with the highest score is always "0BSD" for the test files. That's why the algorithm accuracy is showing as 0%