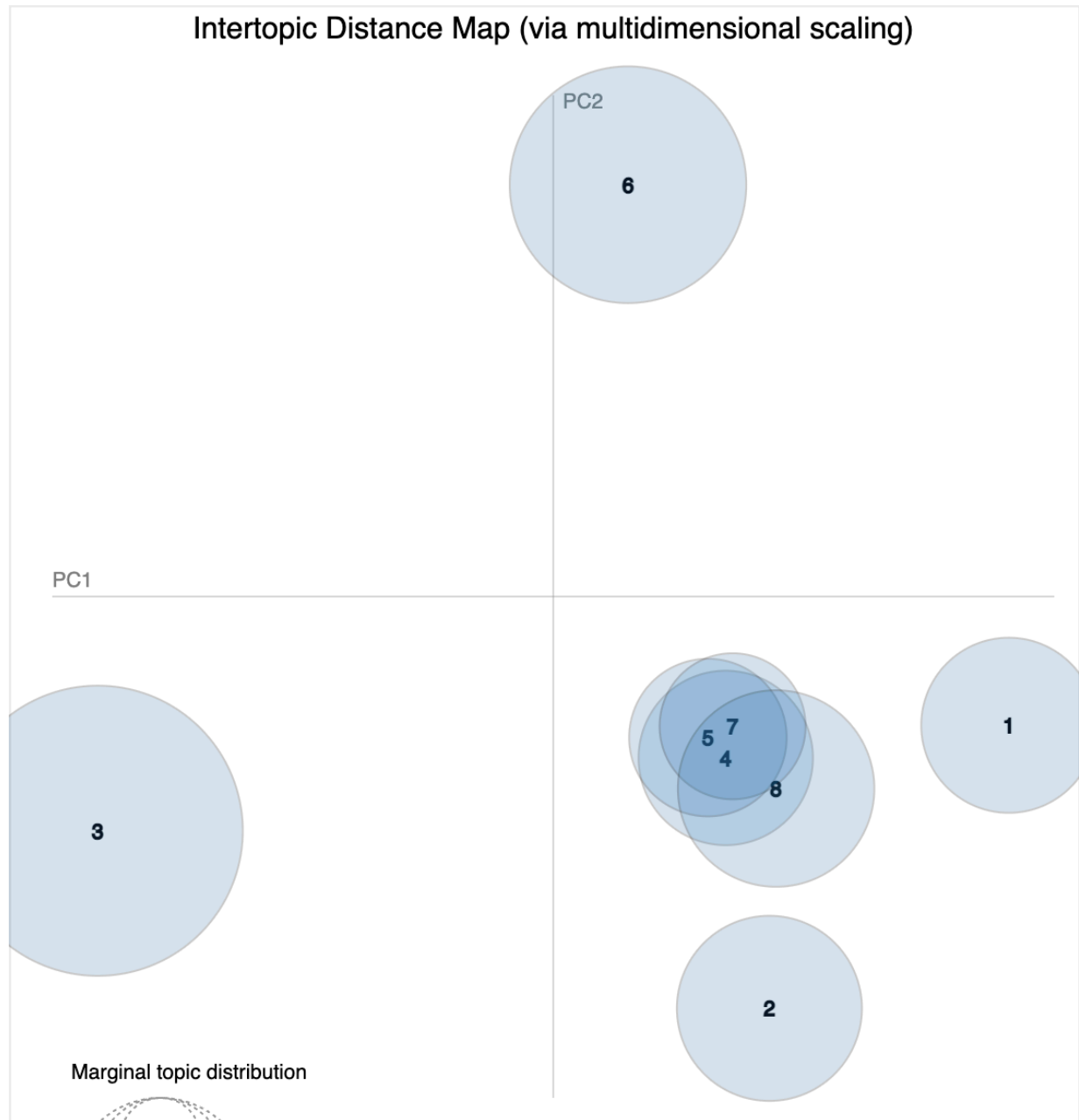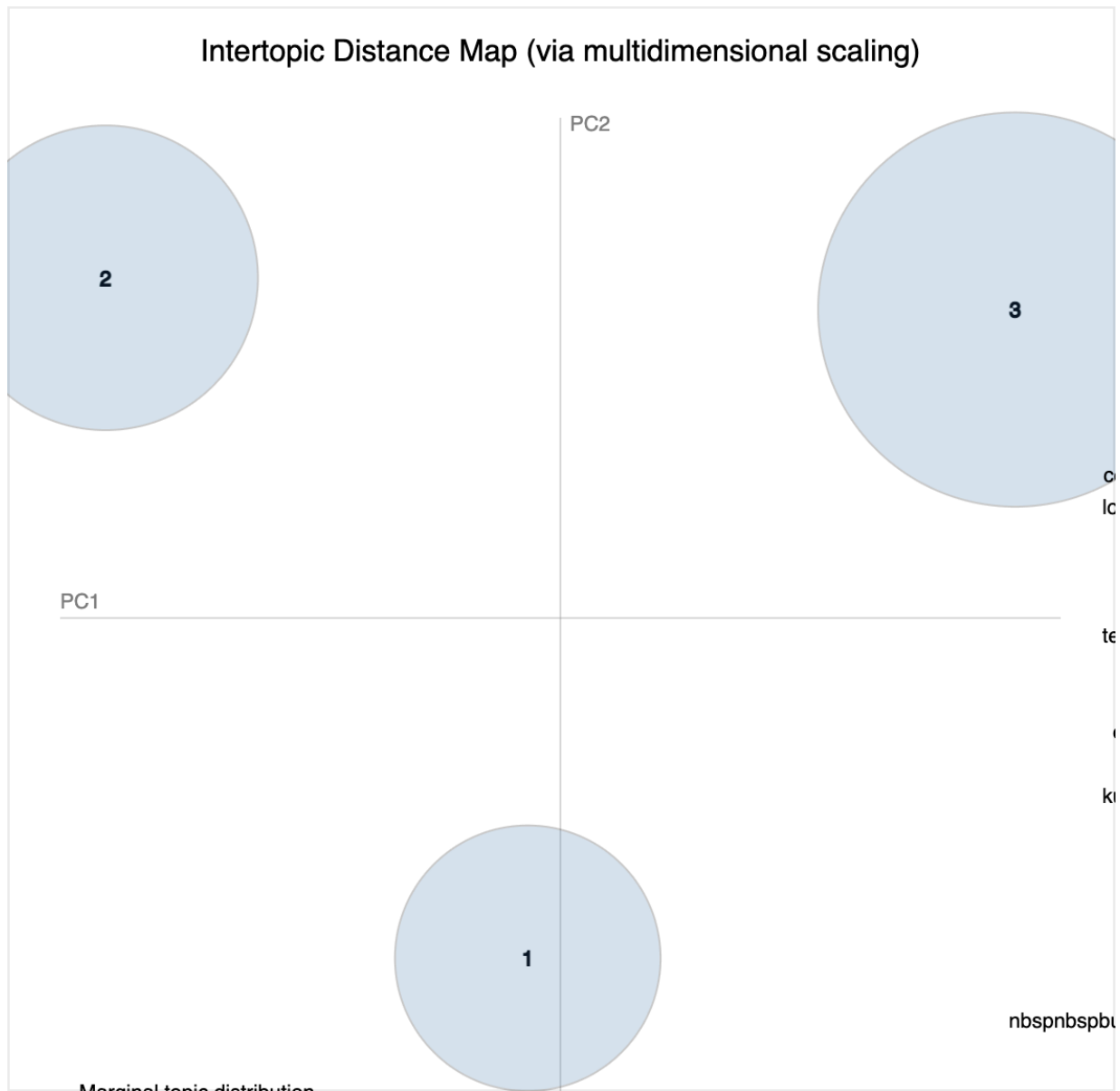# Analysis - Mahaveer

By Choosing 8 topics the Intertopic distance map displayed the following as follows

Intertopic Distance Map (via multidimensional scaling)

It can be inferred that the topics should only be 3 from above graph , so the resulting inter topic distance map is

Intertopic Distance Map (via multidimensional scaling)

PC2

PC1

2

3

1

c
lc

te

ku

nbspnbspbu

Marginal topic distribution

With the visualizer we can also check what words are more relevant to each topic

Topics without including the custom stop words
Topic 1: plugins, stability, located, containerd, tests, kubescape, specification, plugin, component, opengitops
Topic 2: branch, edge, registry, charts, distribution, keycloak, ingress, video, import, content
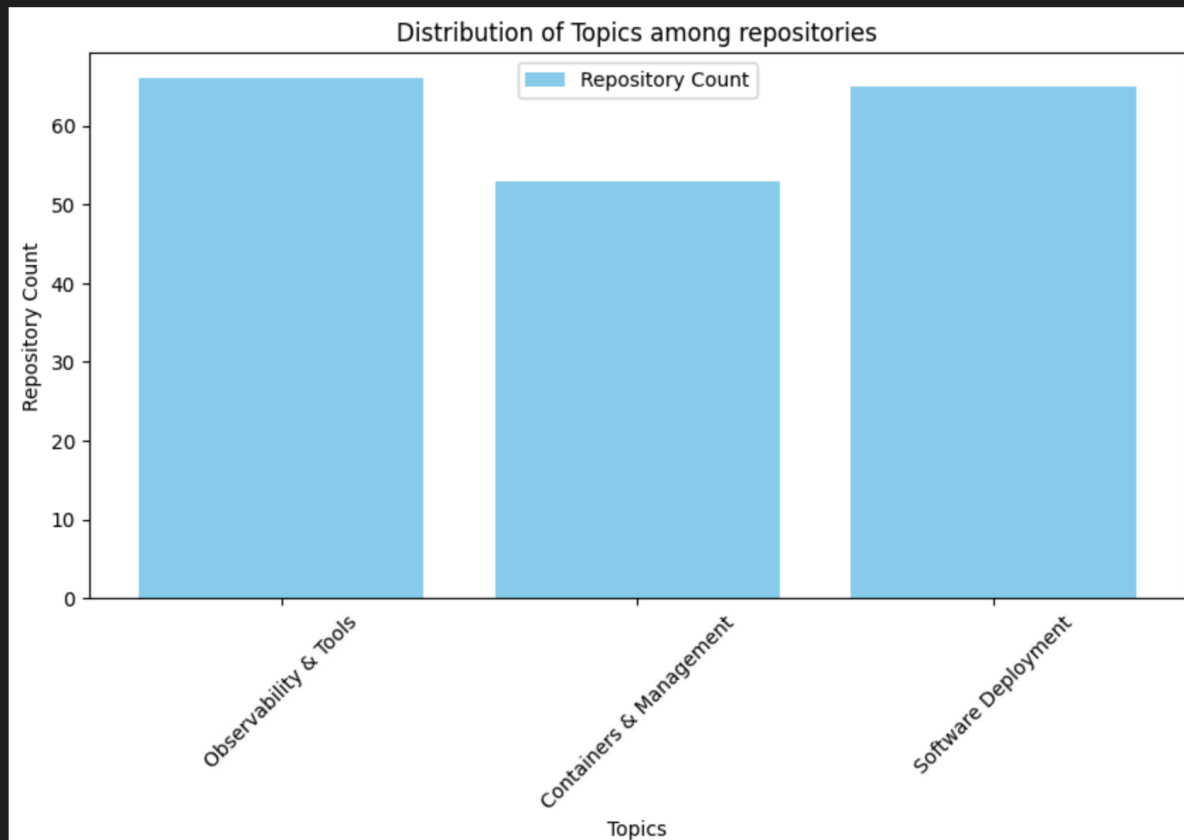Topic 3: opentelemetry, specification, proxy, exporter, gt, devstream, volume, gitops, traffic, pushgateway

So according to the top terms in the topics the topics are renamed as following and their counts in repositories are

Topic 1 : Observability and tools
Topic 2: containers and management
Topic 3: Software Deployment

```
Observability & Tools  :  66
Containers & Management  :  53
Software Deployment  :  65
```


Distribution of Topics among repositories

Here we only represented number of topics to be 3, for even greater granularity, We can simply change the number of topics, and number of topics assigned for each document.

THE CODE IN SUBMISSION SUPPORTS THIS BY CHANGING VARIABLES