

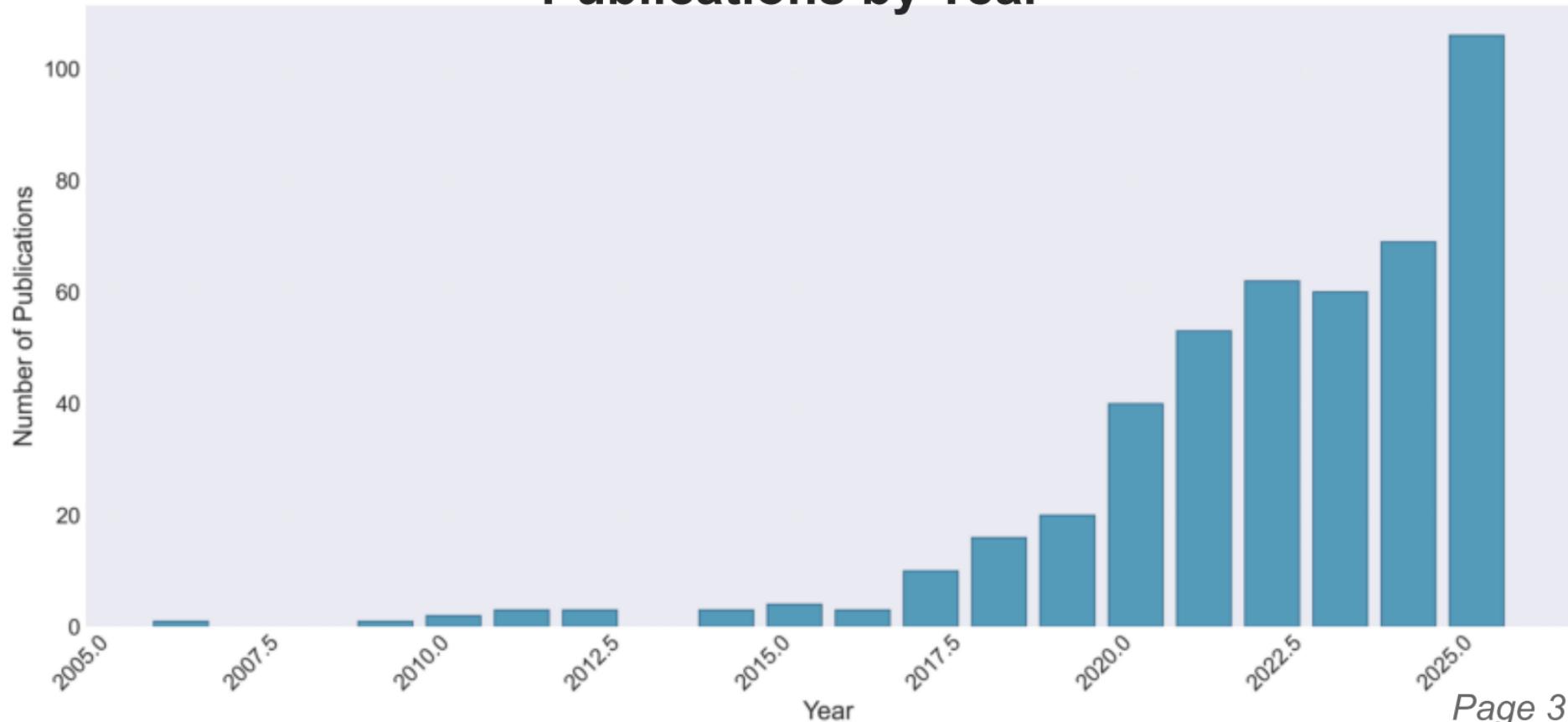
# **Meta-Analysis Graphical Abstract**

**Total Papers: 456**  
**Date: 2025-12-15**

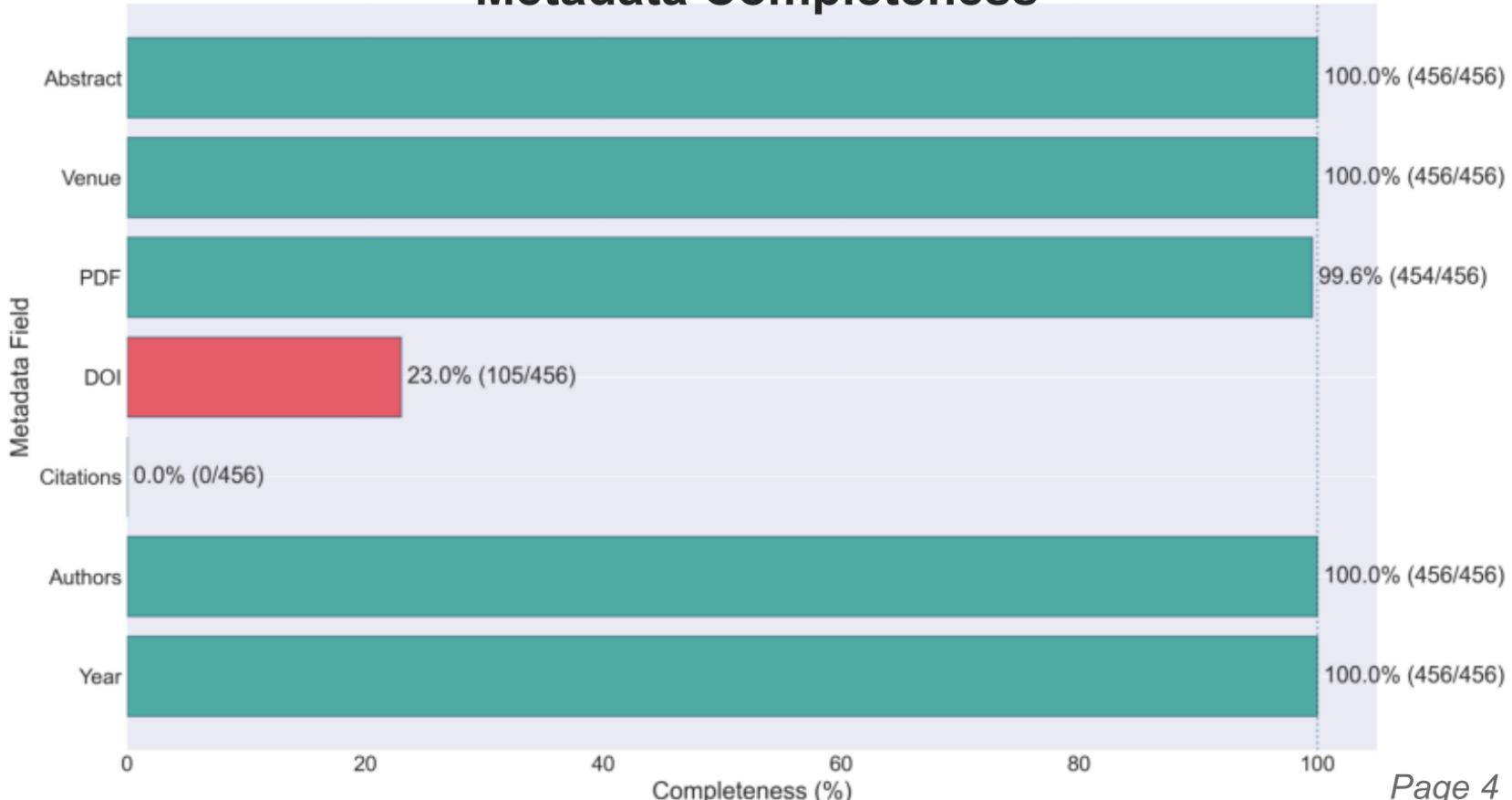
**(Includes Embedding Analysis)**

# **Overview**

# Publications by Year

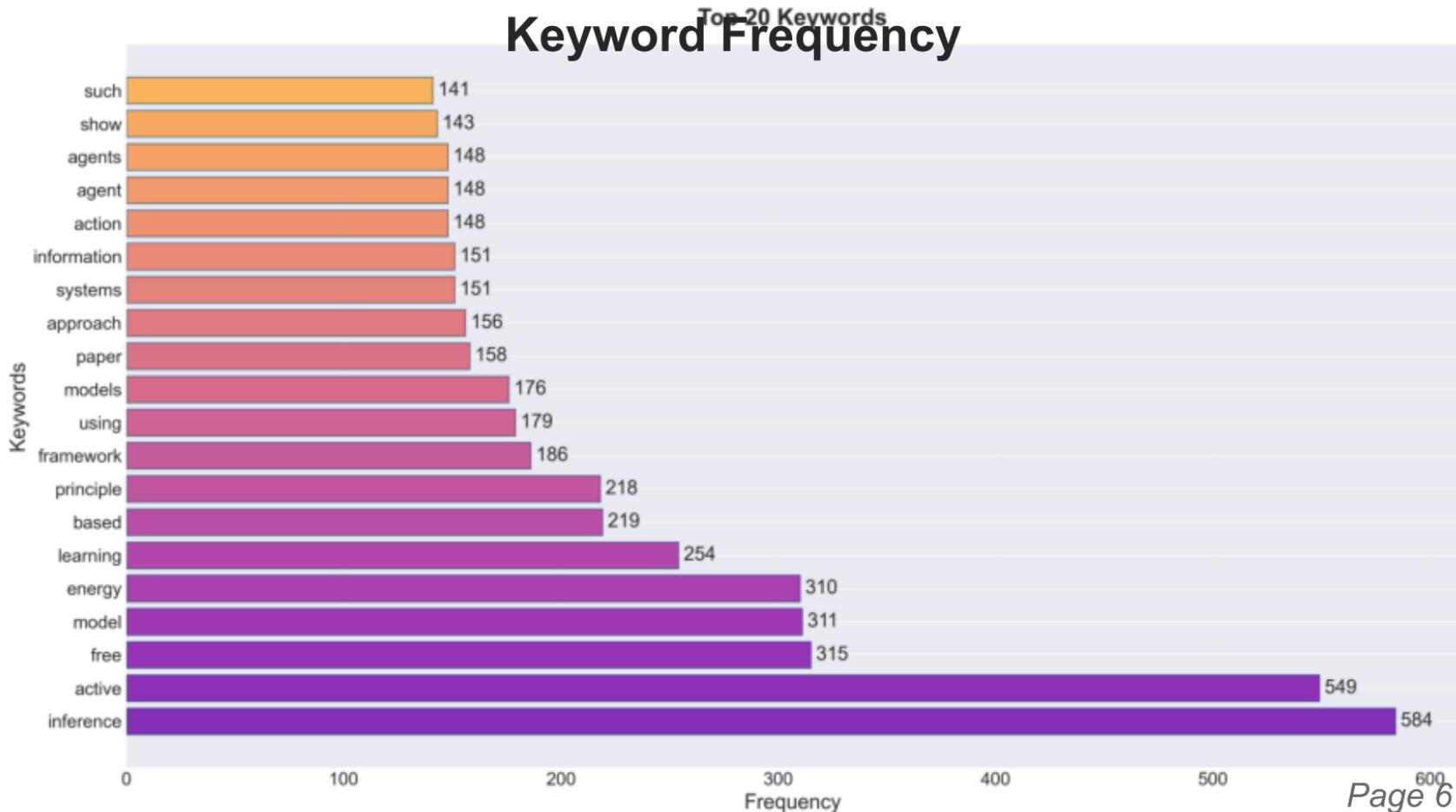


# Metadata Completeness

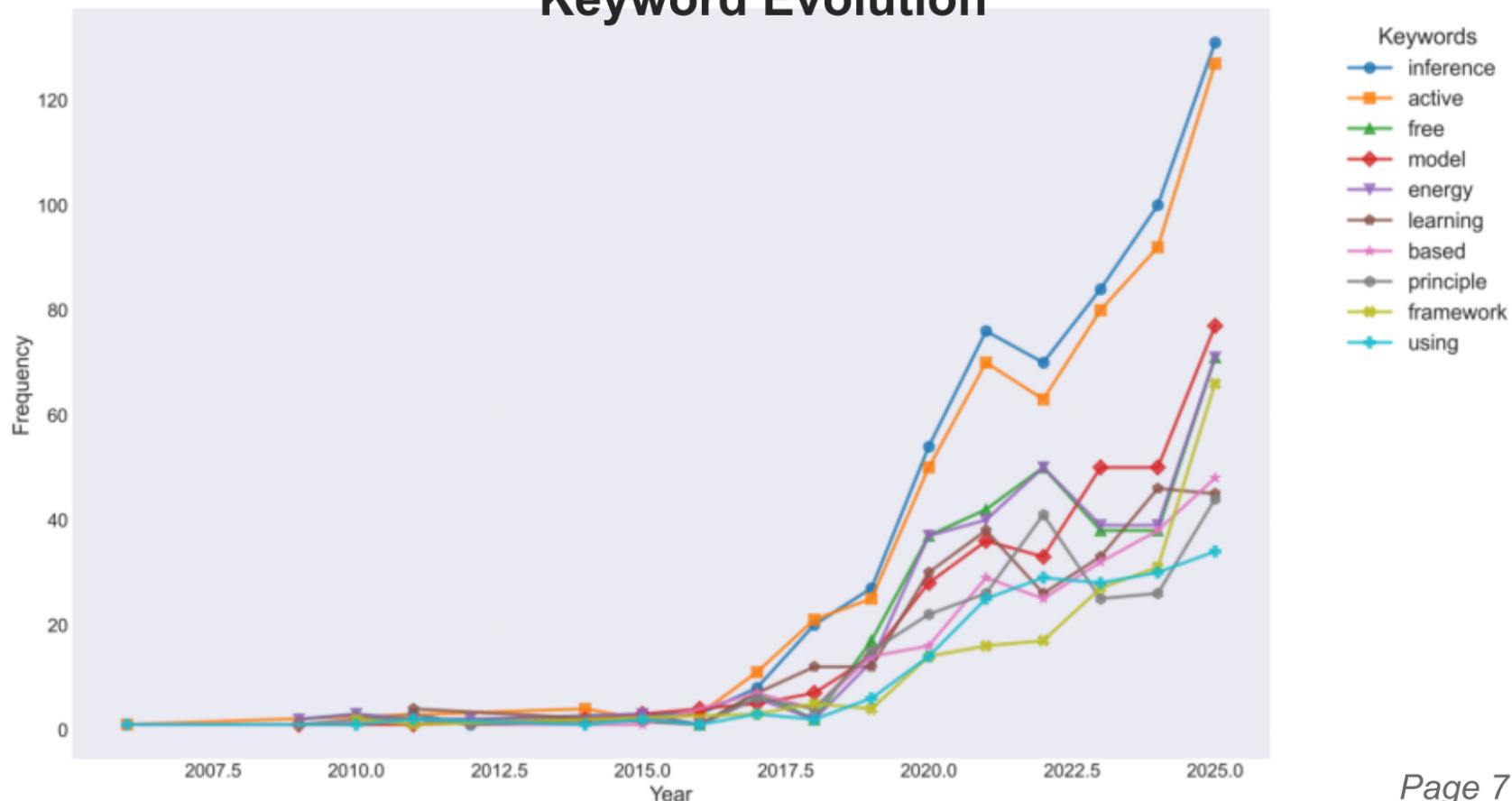


# **Keyword Analysis**

# Keyword Frequency

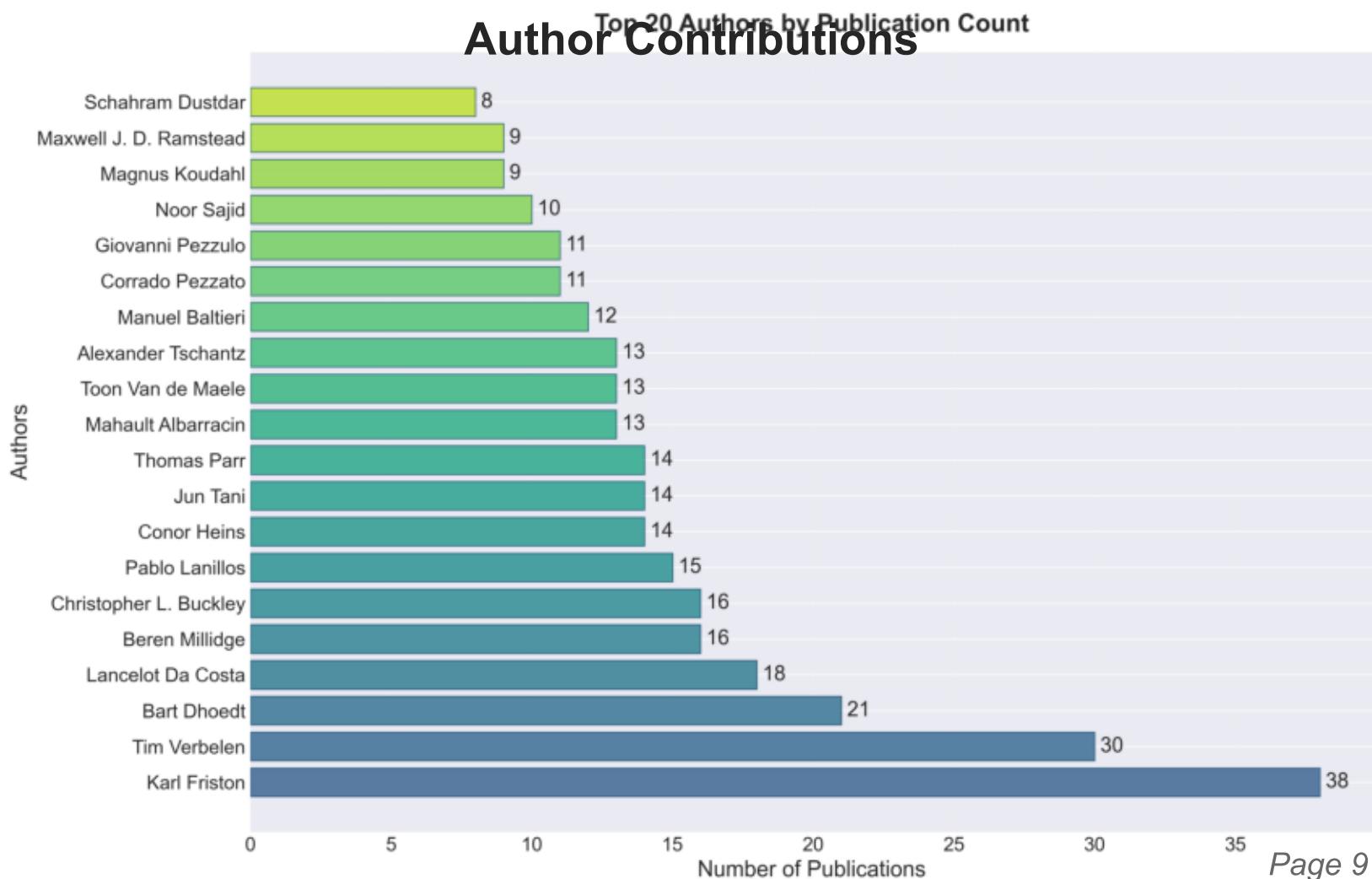


# Keyword Evolution



# **Author & Venue Analysis**

# Author Contributions



Top 15 Publication Venues

# Venue Distribution

Venues

arXiv

456



0

100

200

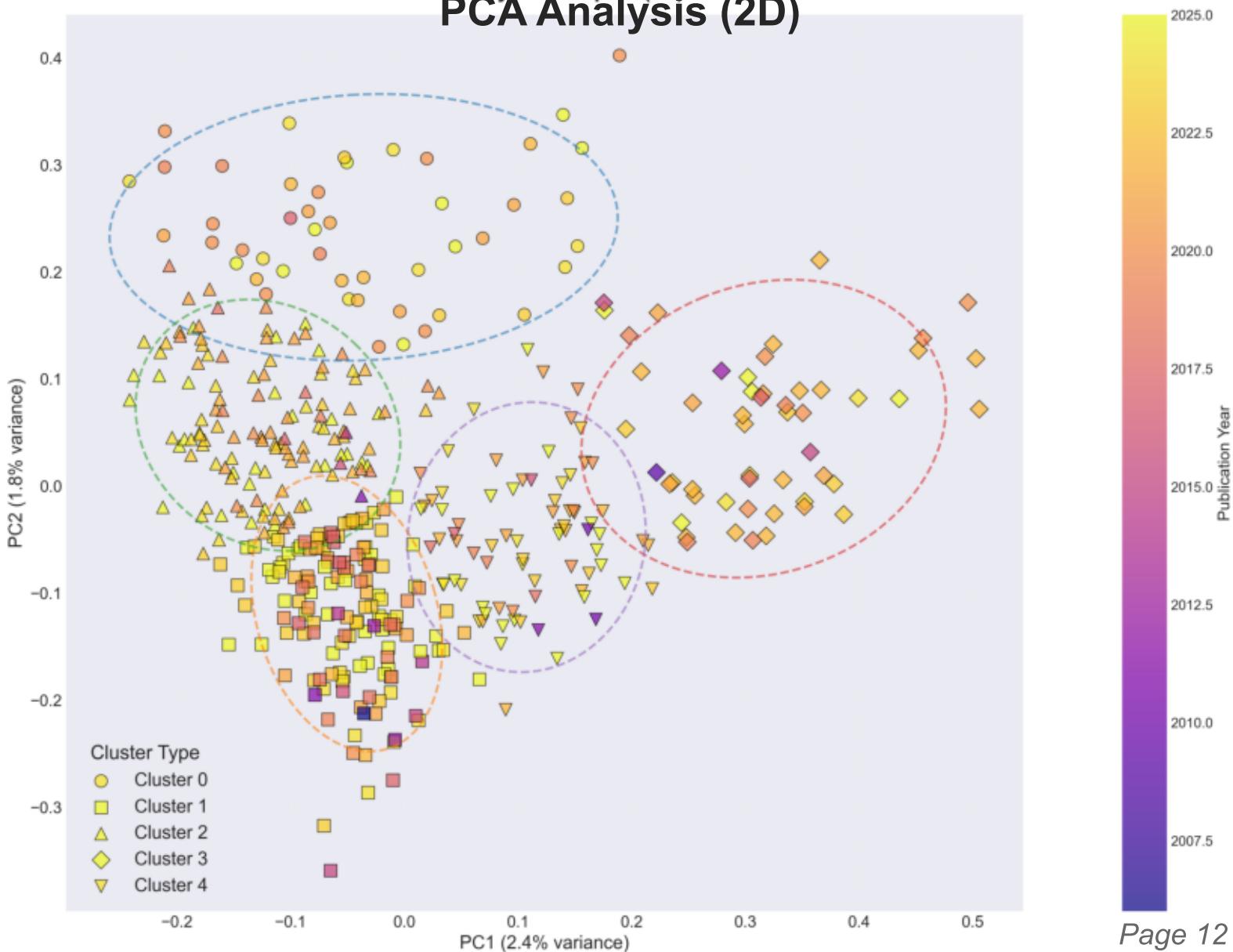
300

400

Number of Publications

# **PCA Analysis**

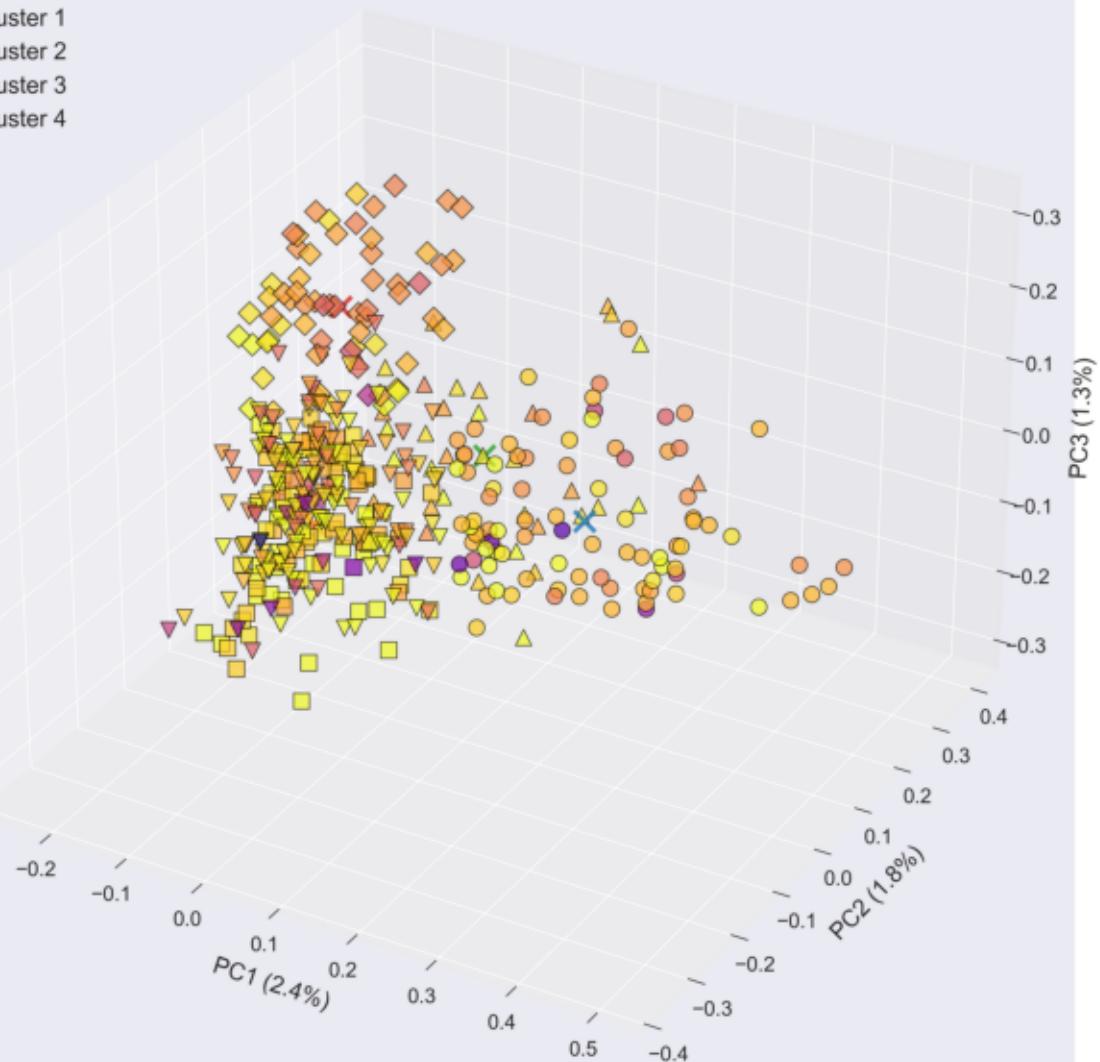
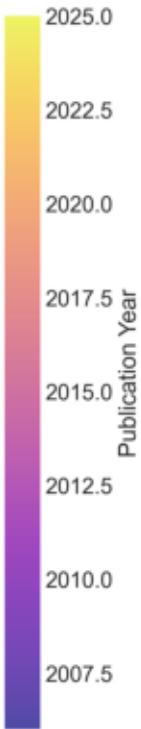
# PCA Analysis (2D)



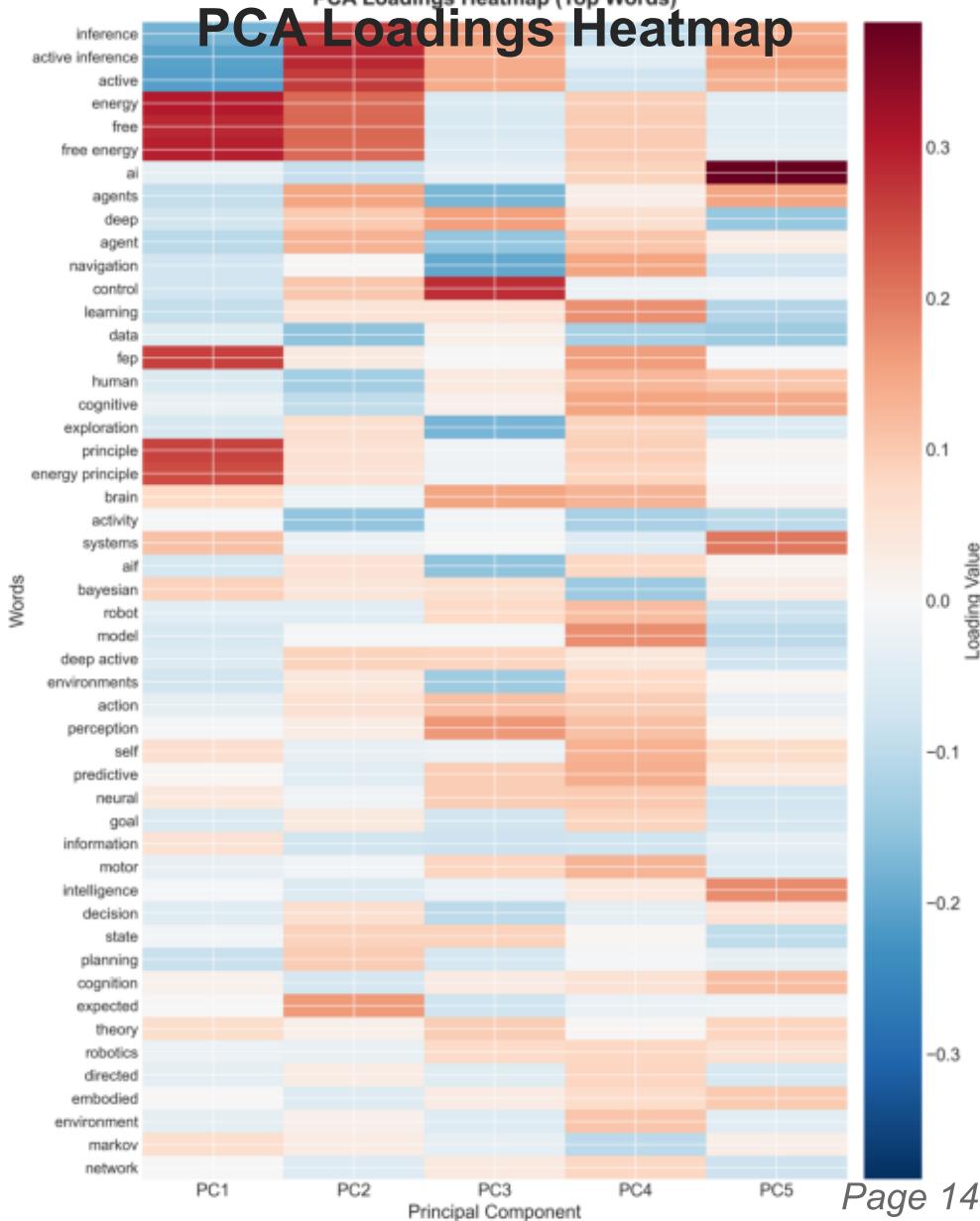
# PCA Analysis (3D)

Cluster Type

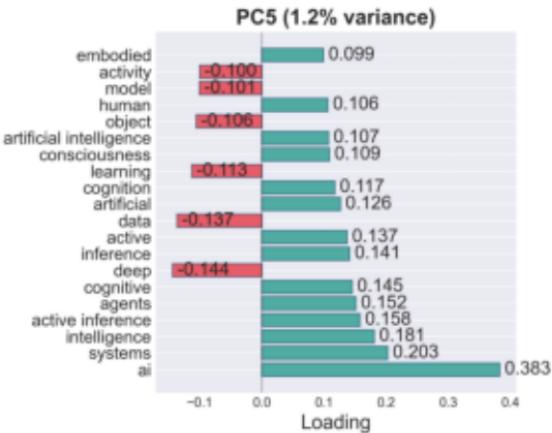
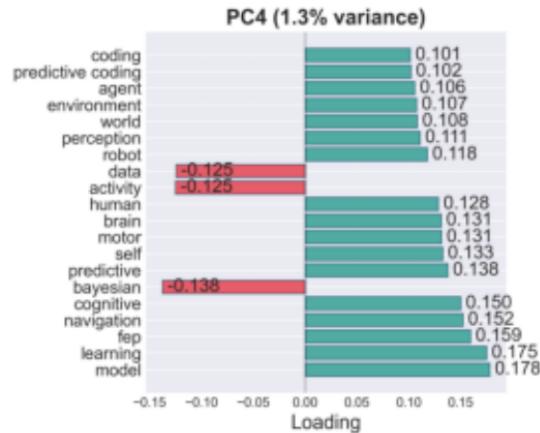
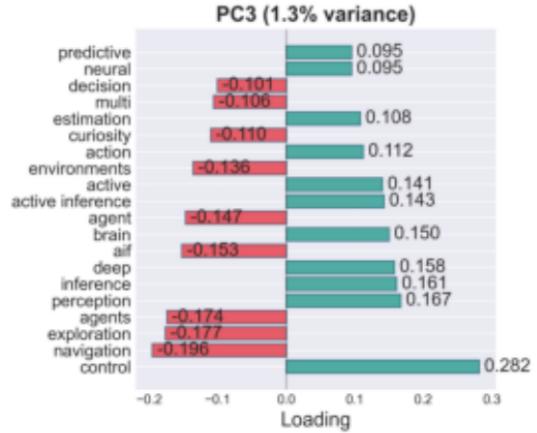
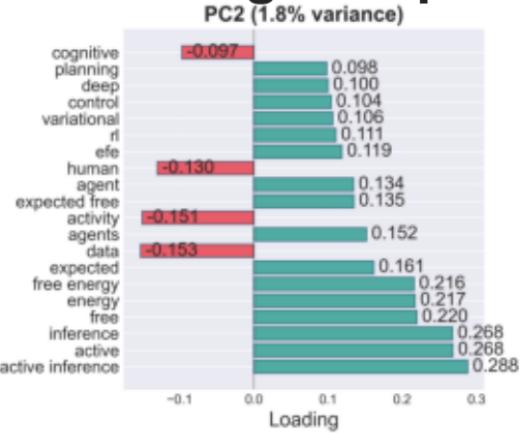
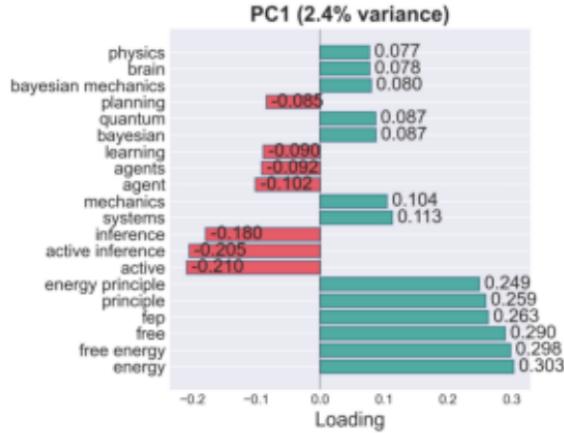
- Cluster 0
- Cluster 1
- ▲ Cluster 2
- ◆ Cluster 3
- ▼ Cluster 4



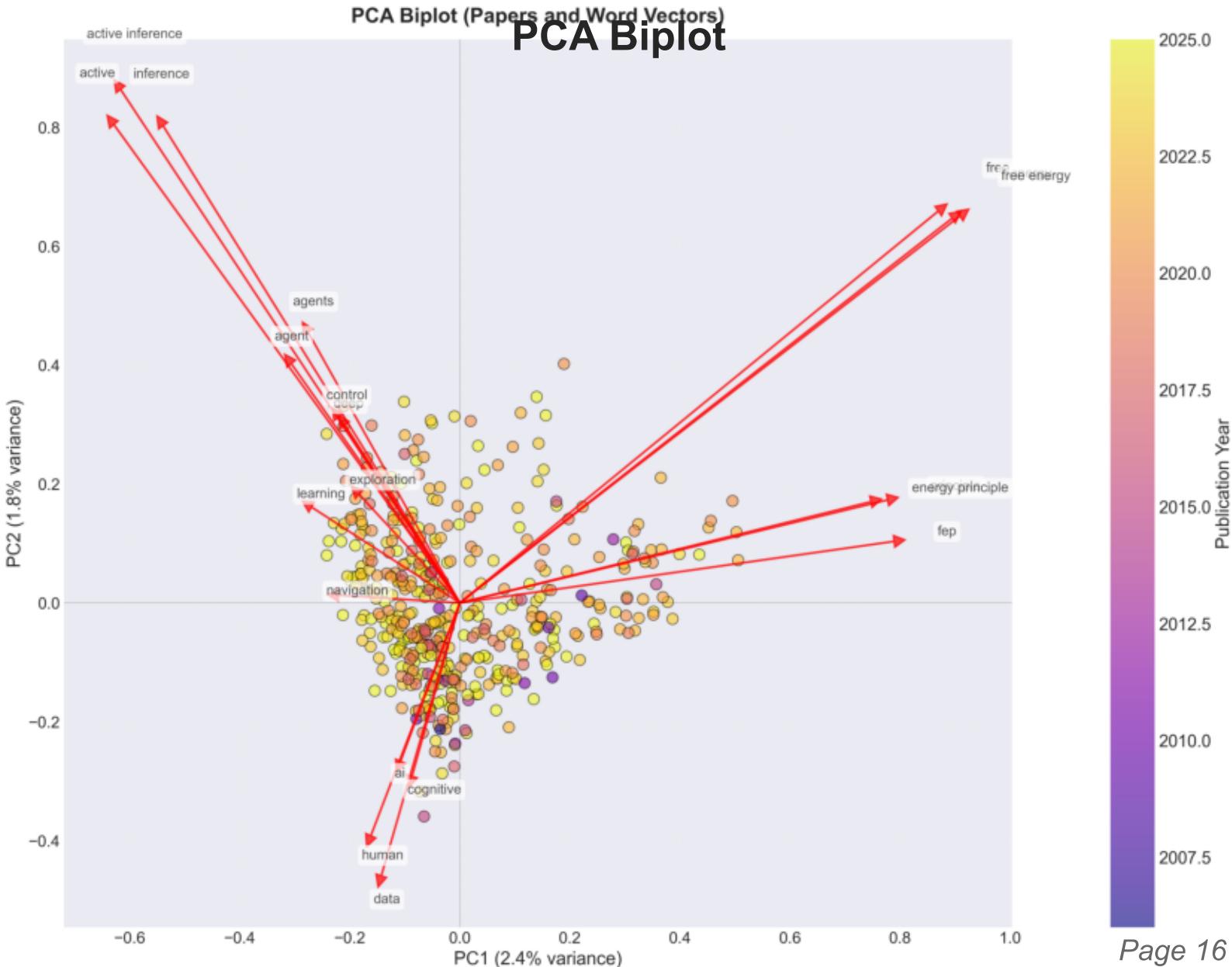
PCA Loadings Heatmap (Top Words)



# PCA Loadings Barplots

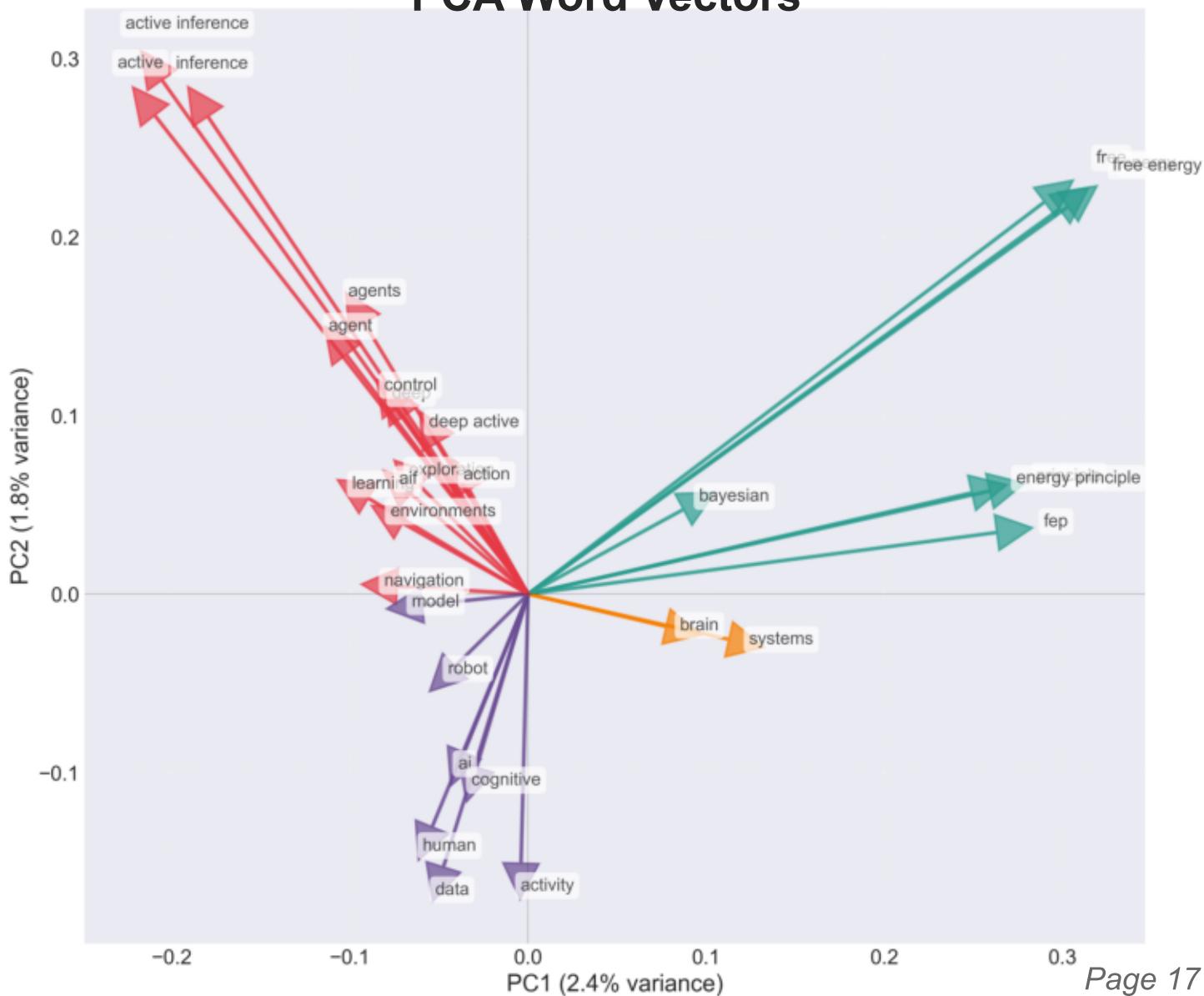


PCA Biplot (Papers and Word Vectors)



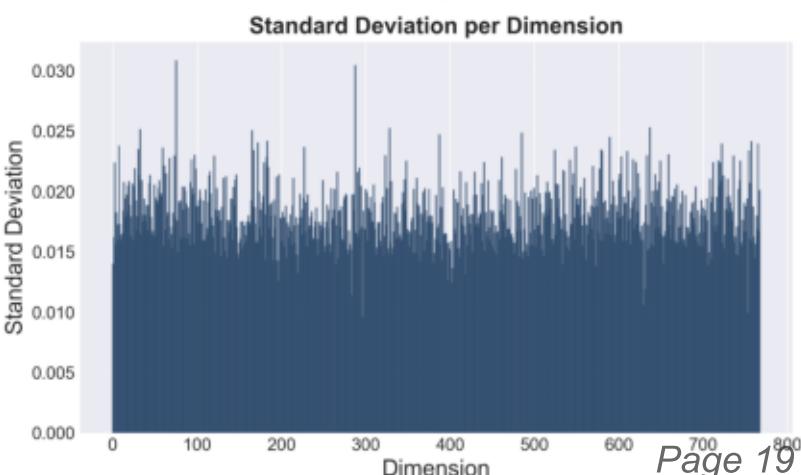
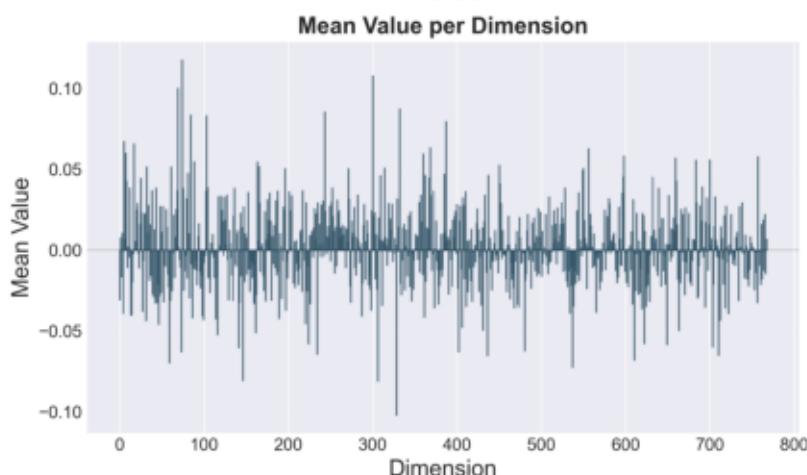
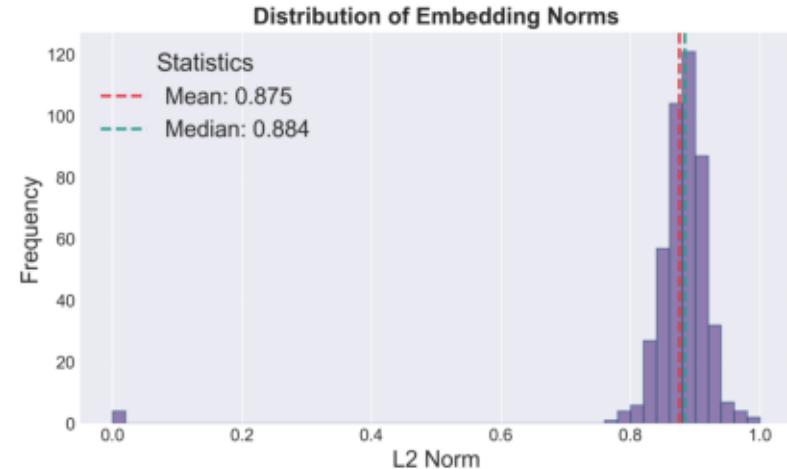
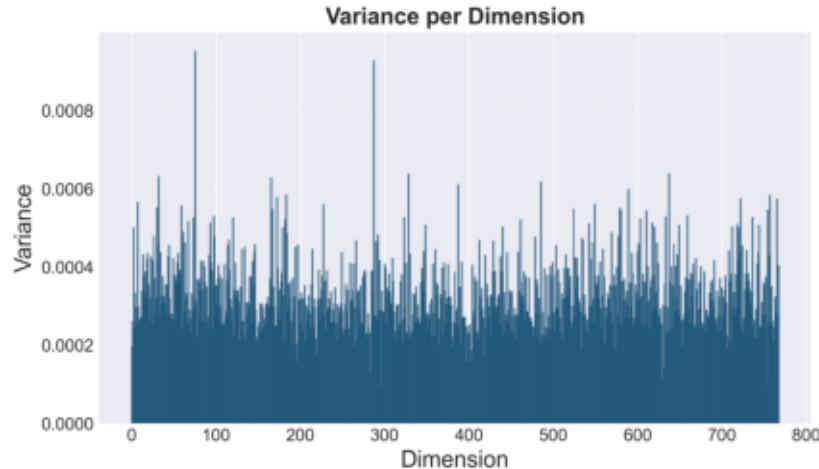
# Word Vectors in Principal Component Space

## PCA Word Vectors

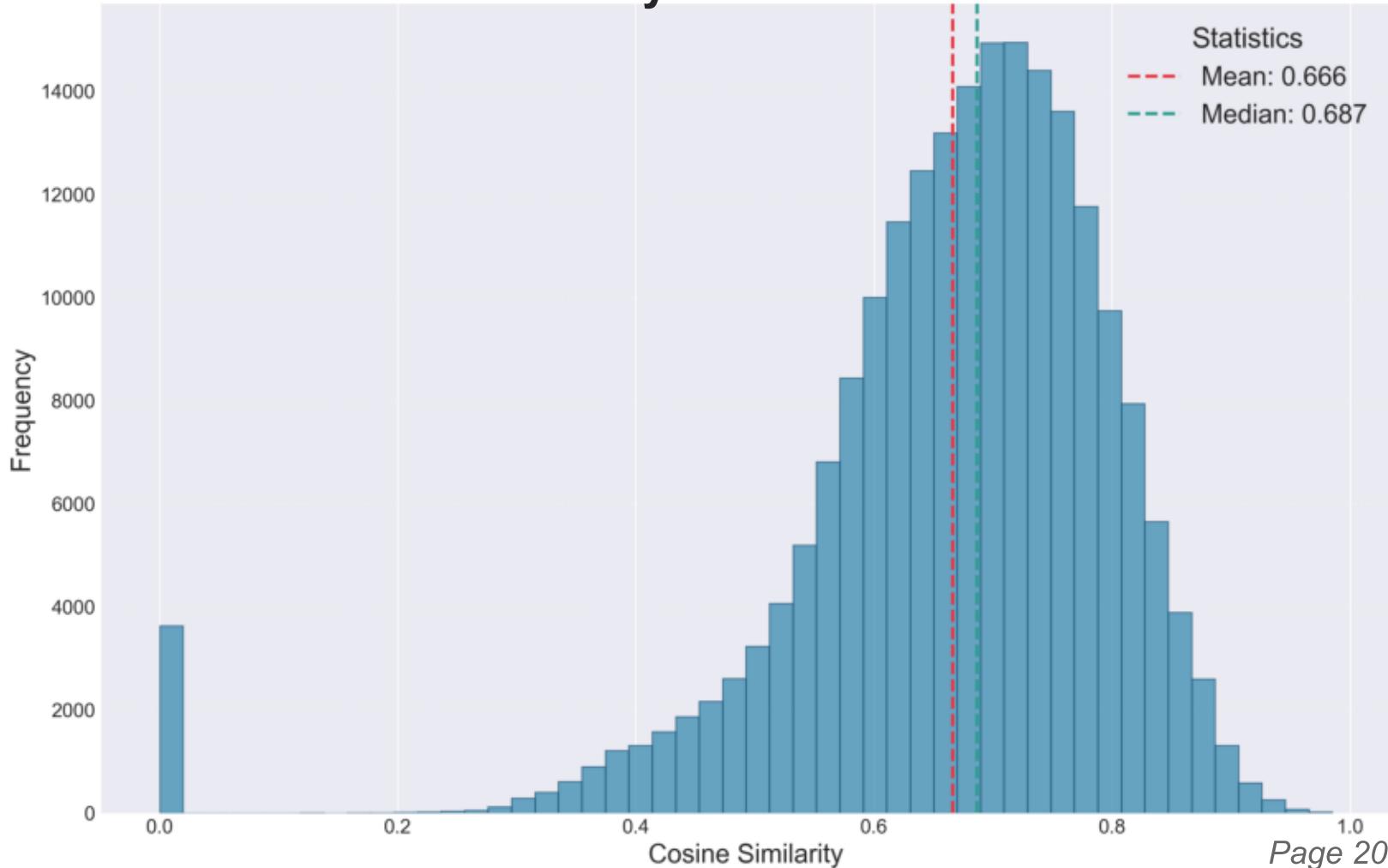


# **Embedding Analysis**

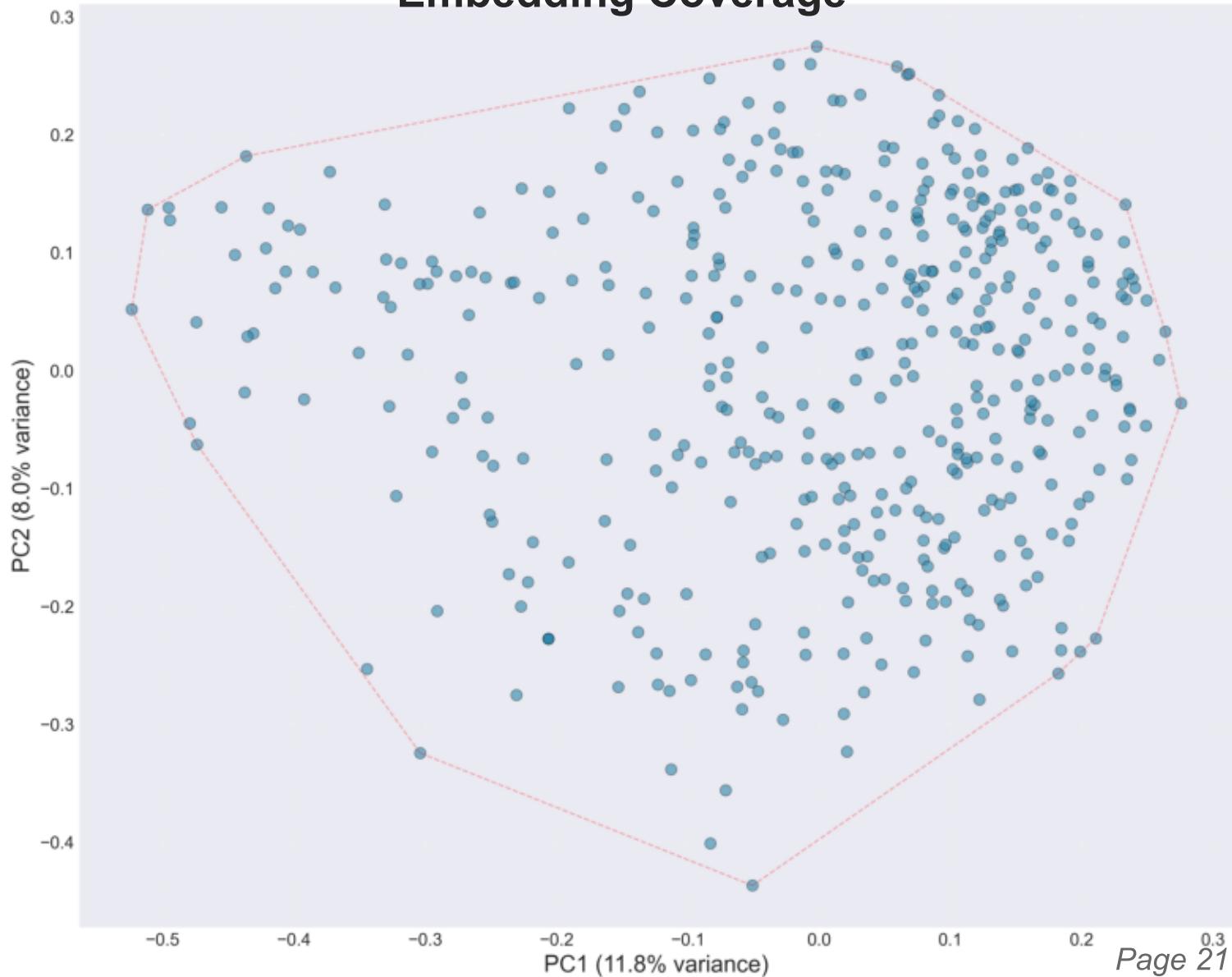
# Embedding Quality



# Similarity Distribution

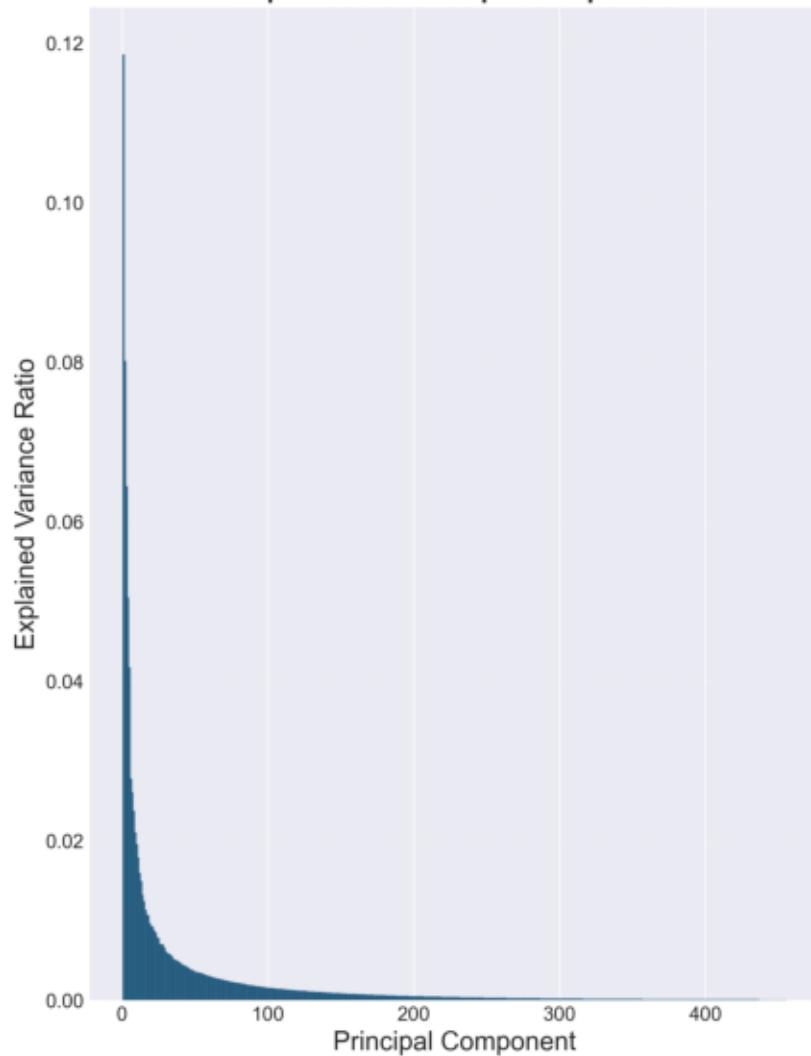


# Embedding Space Coverage

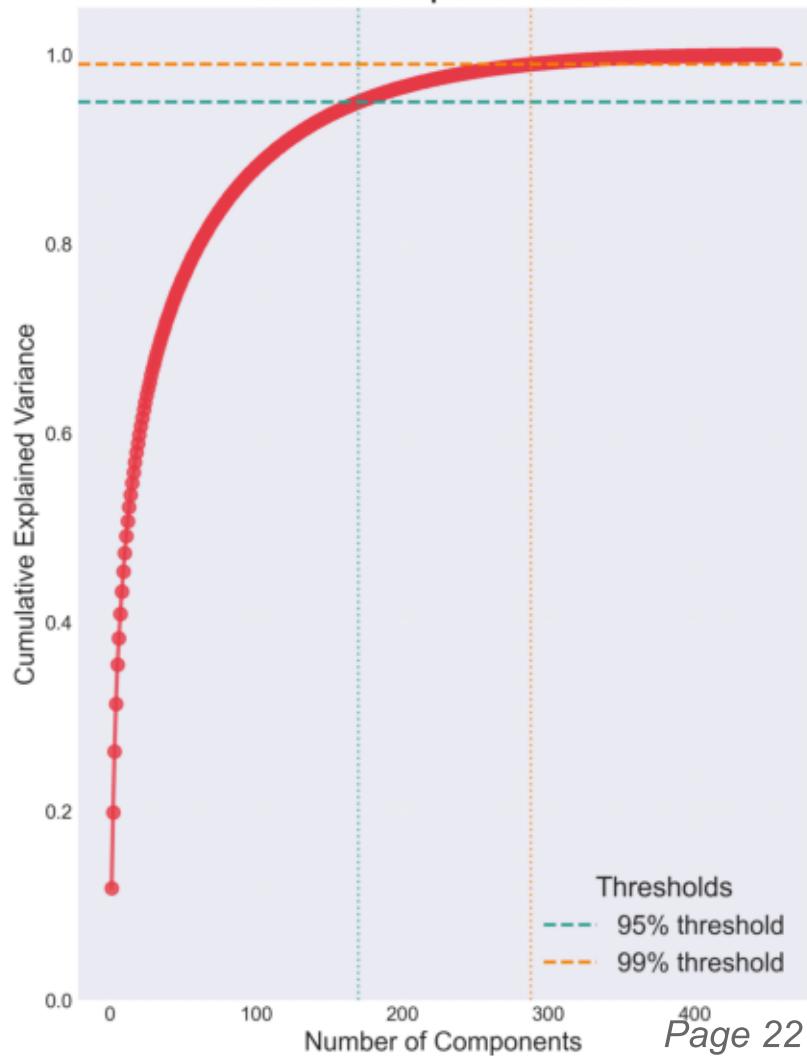


# Dimensionality Analysis

Explained Variance per Component



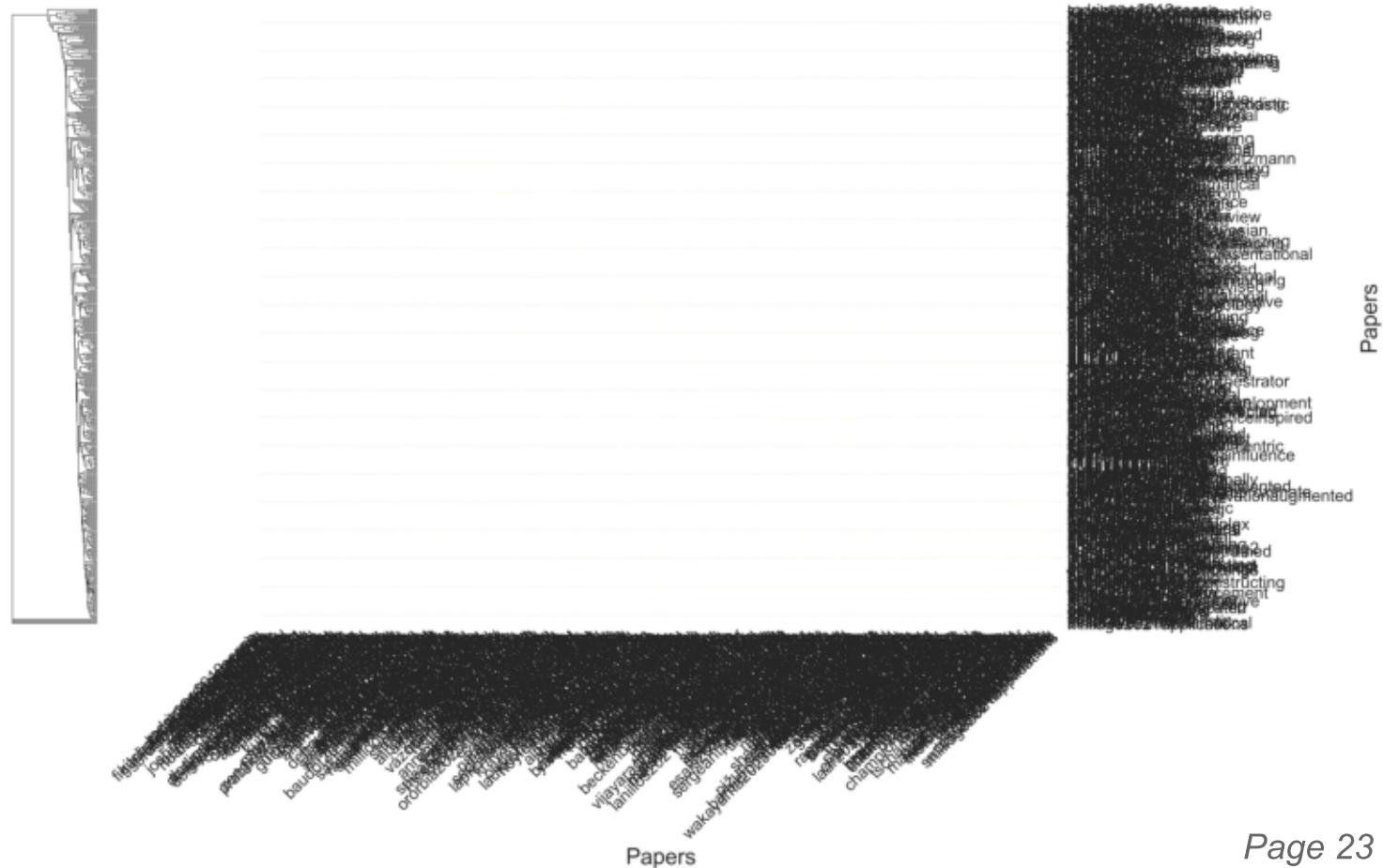
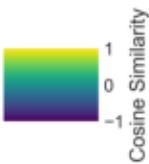
Cumulative Explained Variance



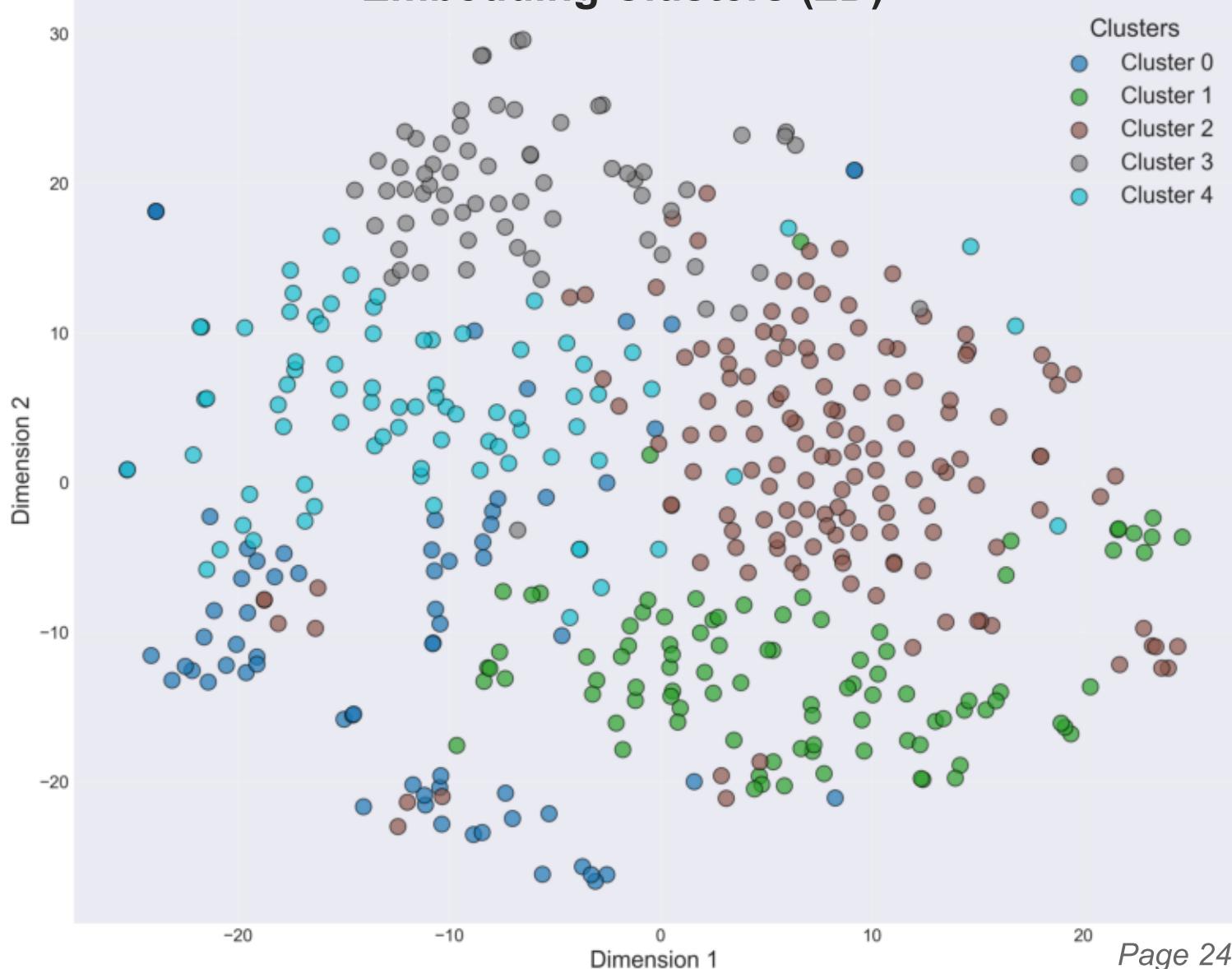
Thresholds

- 95% threshold
- 99% threshold

# Paper Similarity Matrix (Embeddings) Similarity Heatmap



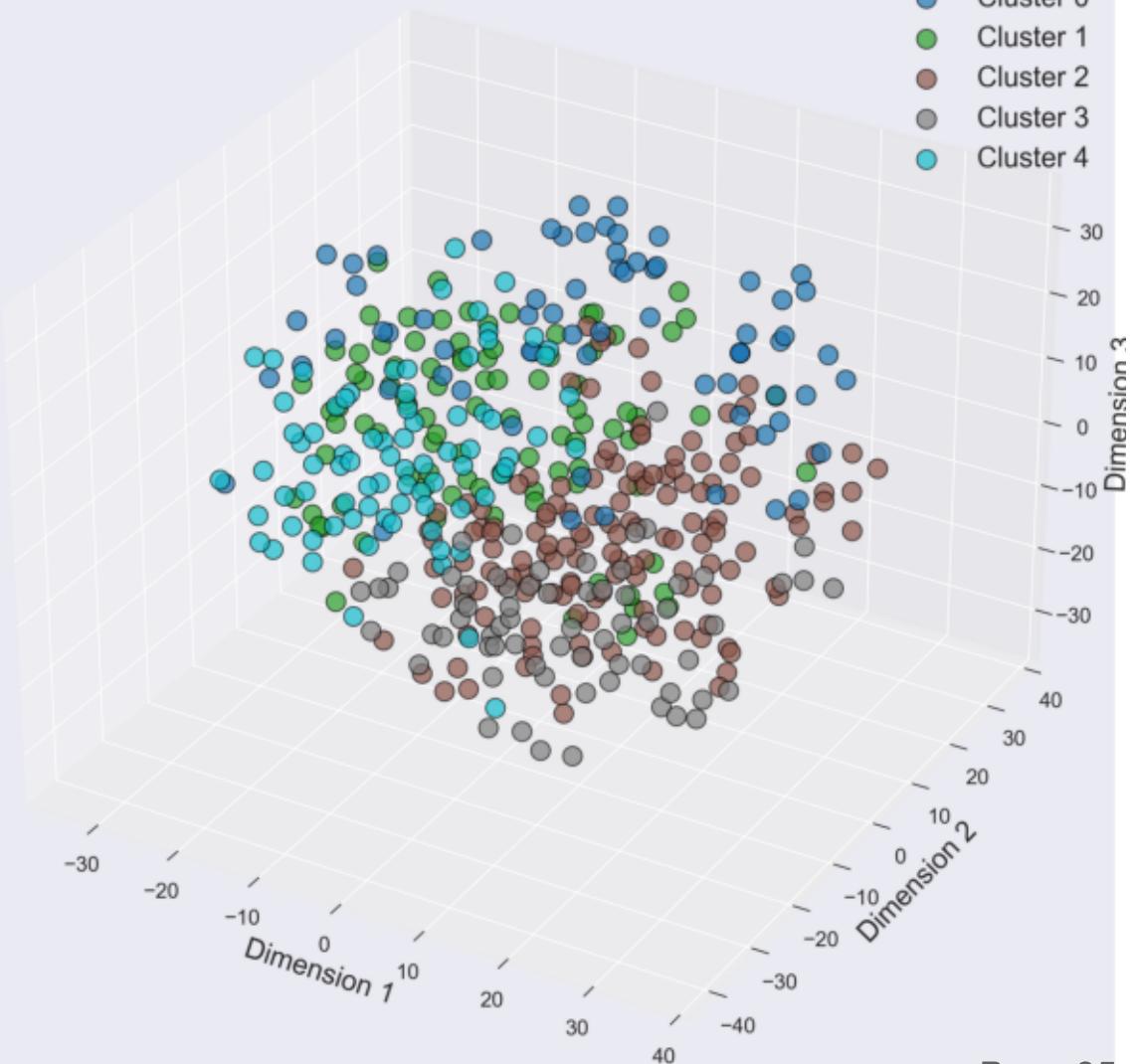
# Embedding Clusters (2D)



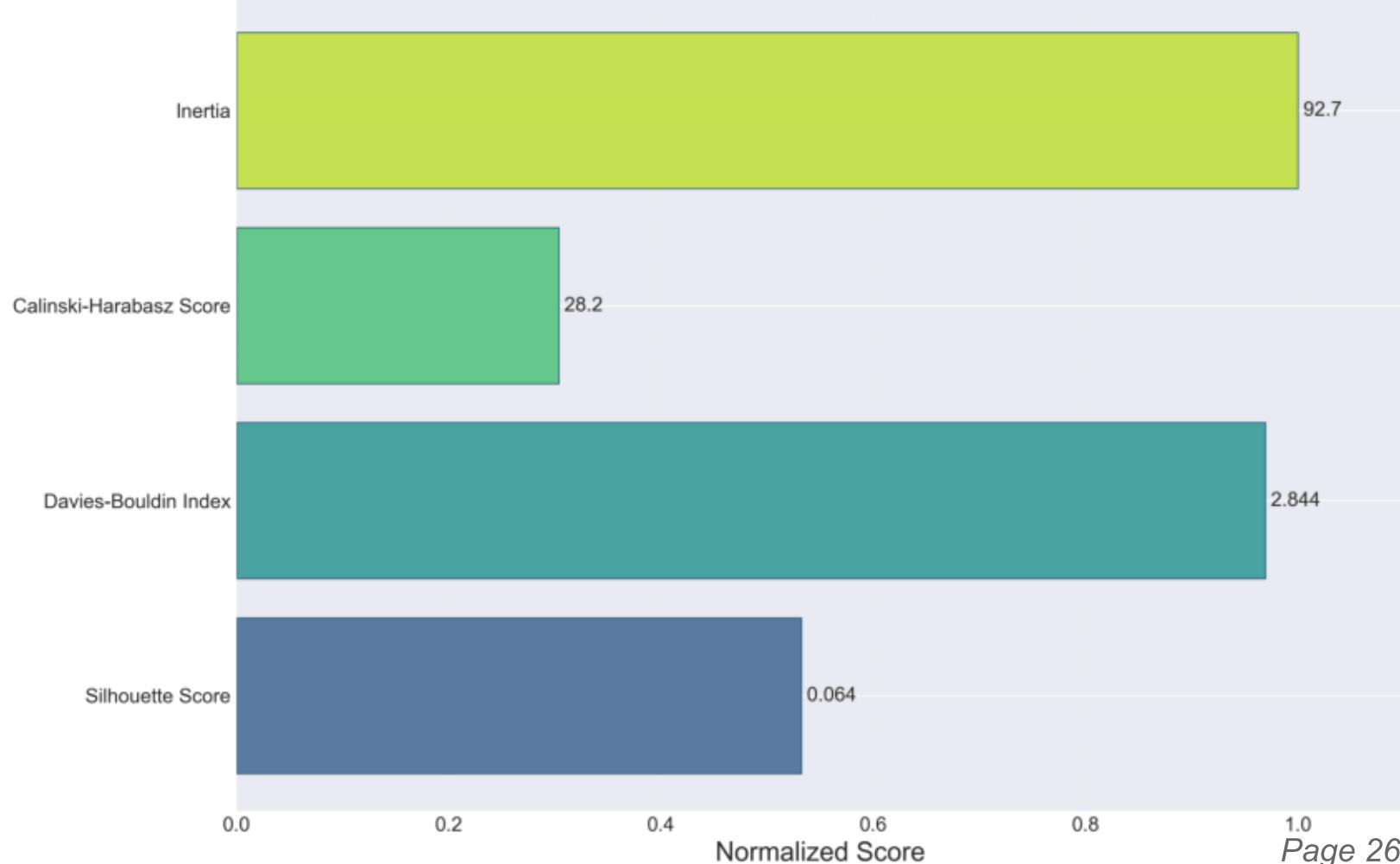
# Embedding Clusters (3D)

Clusters

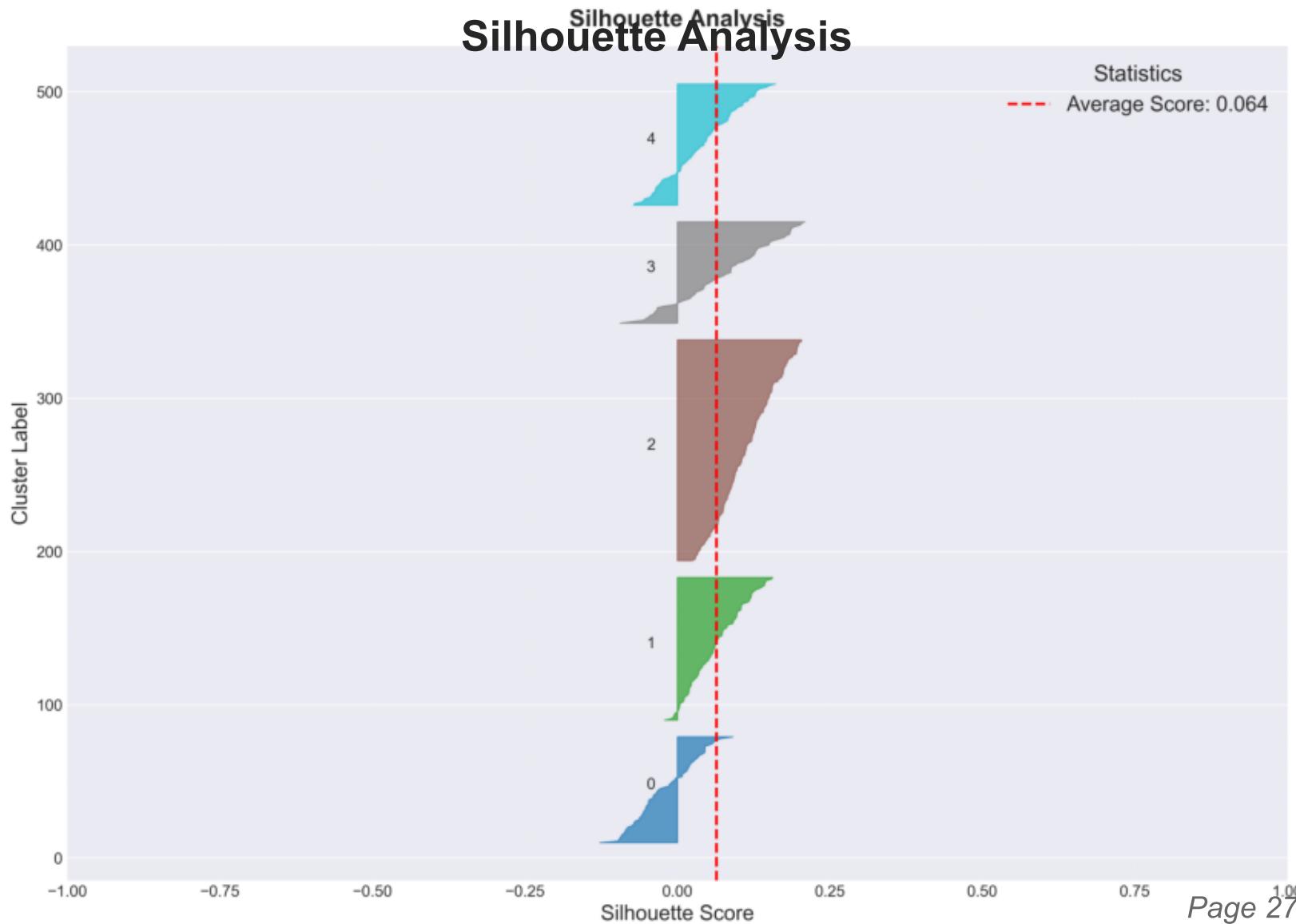
- Cluster 0
- Cluster 1
- Cluster 2
- Cluster 3
- Cluster 4



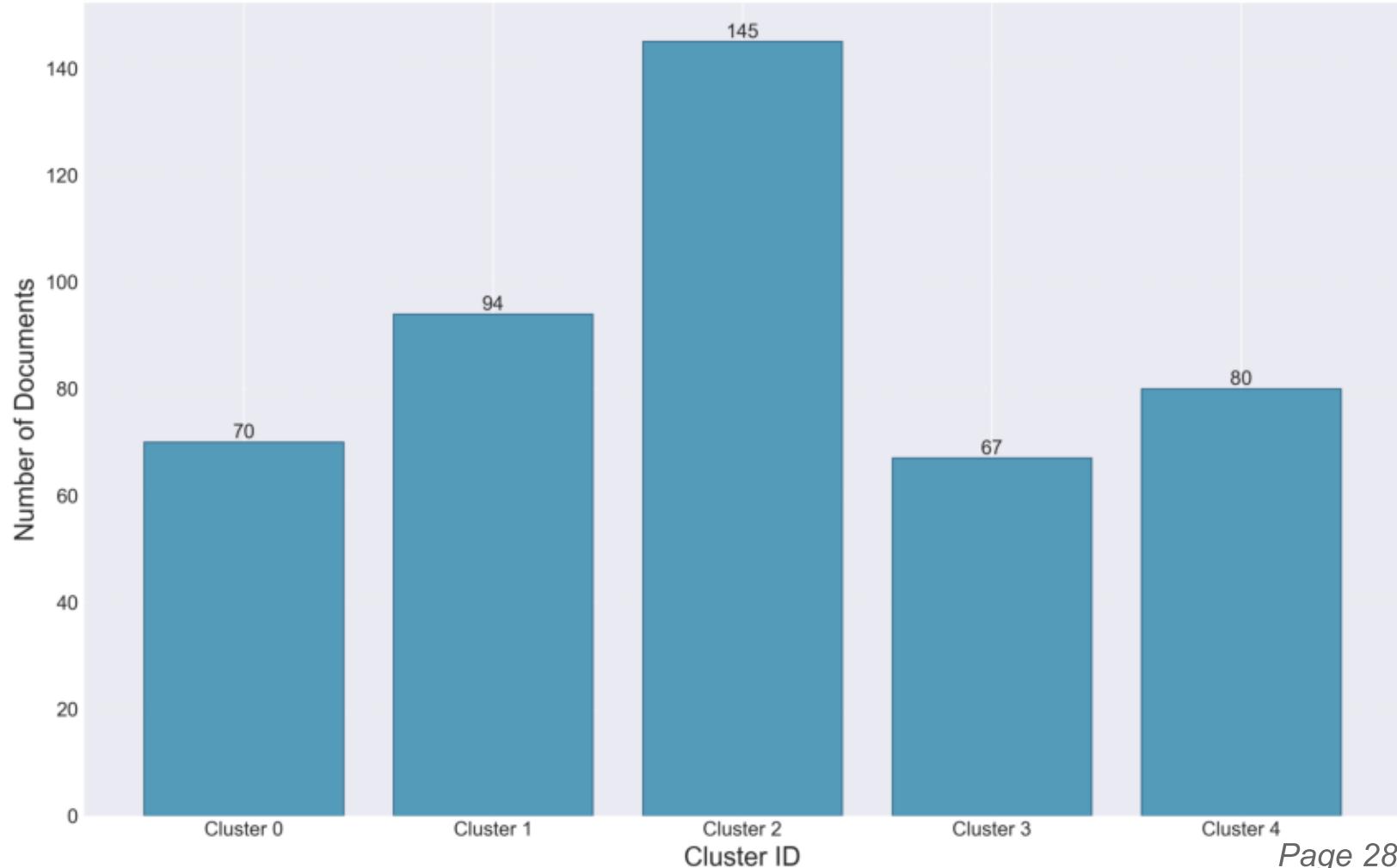
# Clustering Quality Metrics



# Silhouette Analysis



# Cluster Size Distribution



# Embedding Outliers

