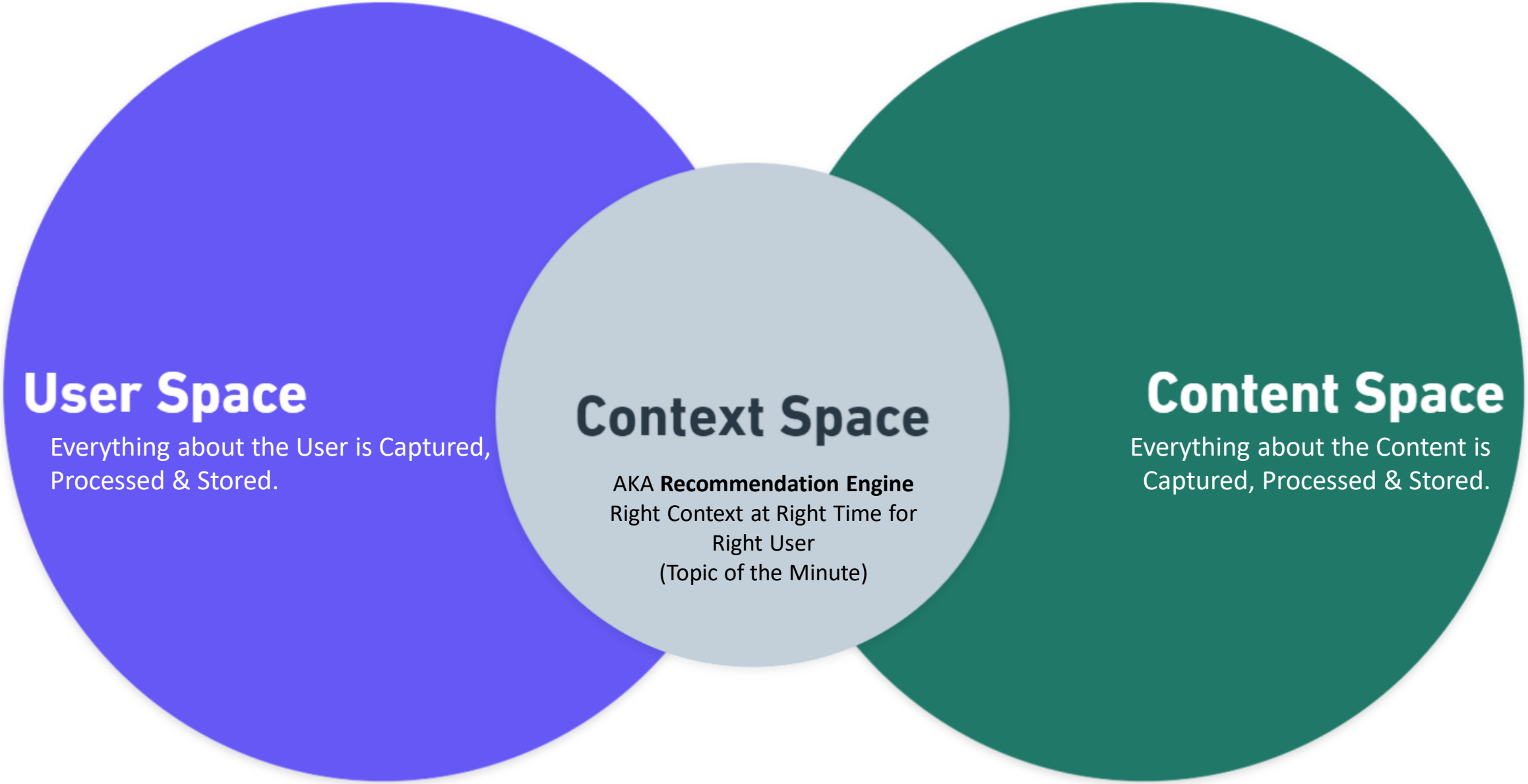
Content Recommendation Systems

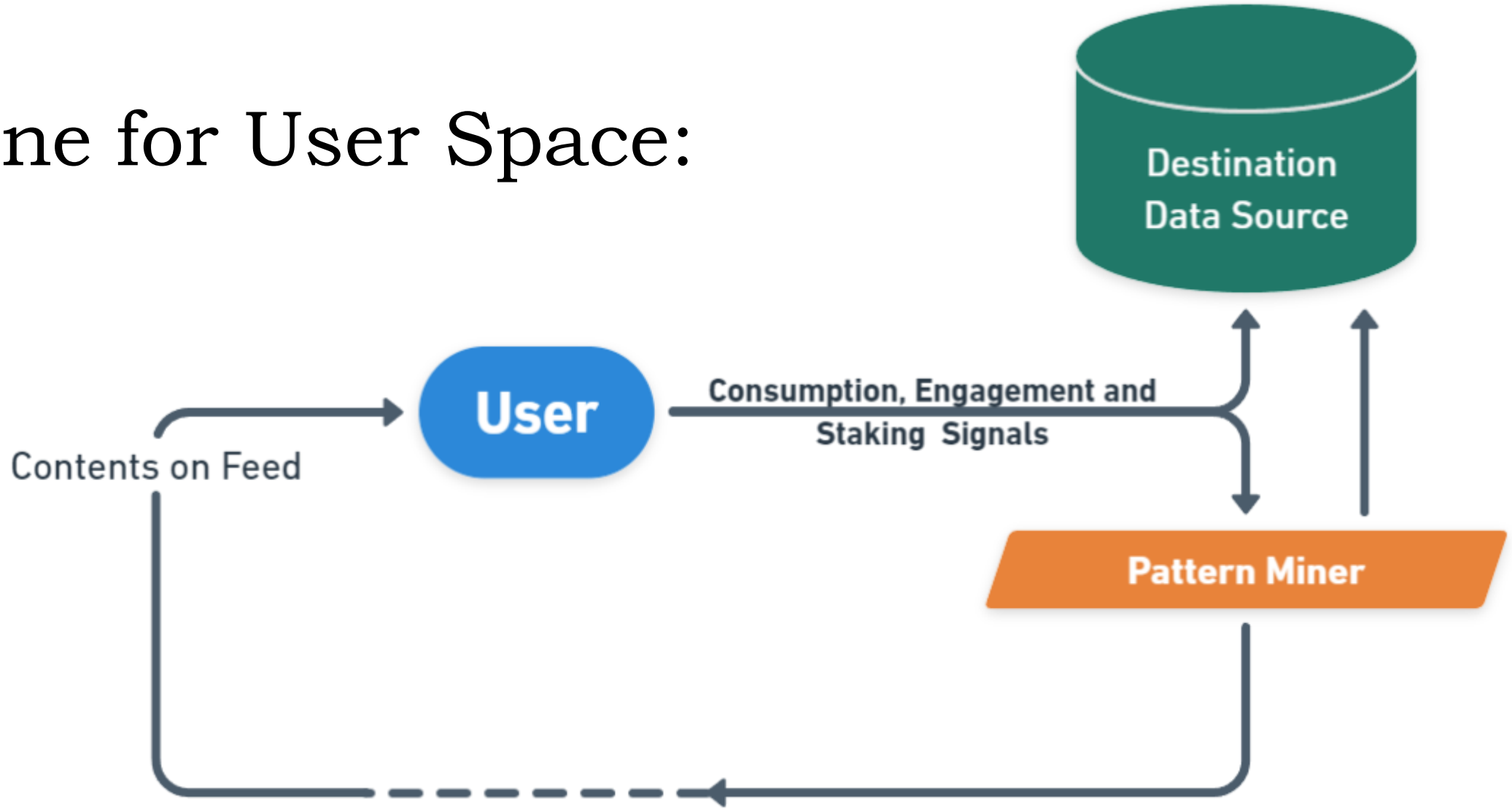
Abhishek Pawaskar

Recommendation Venn Diagram:

Diagram describing various spaces involved in the process of Recommending contents

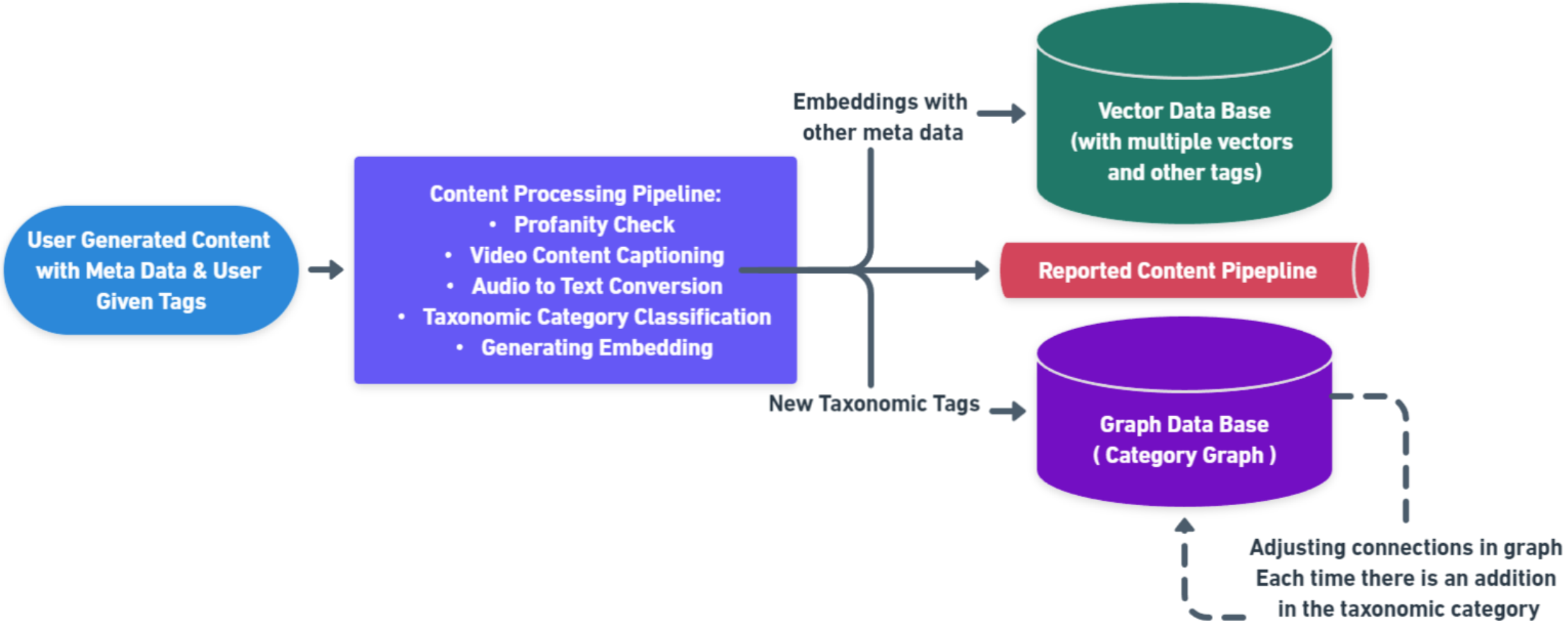


Pipeline for User Space:

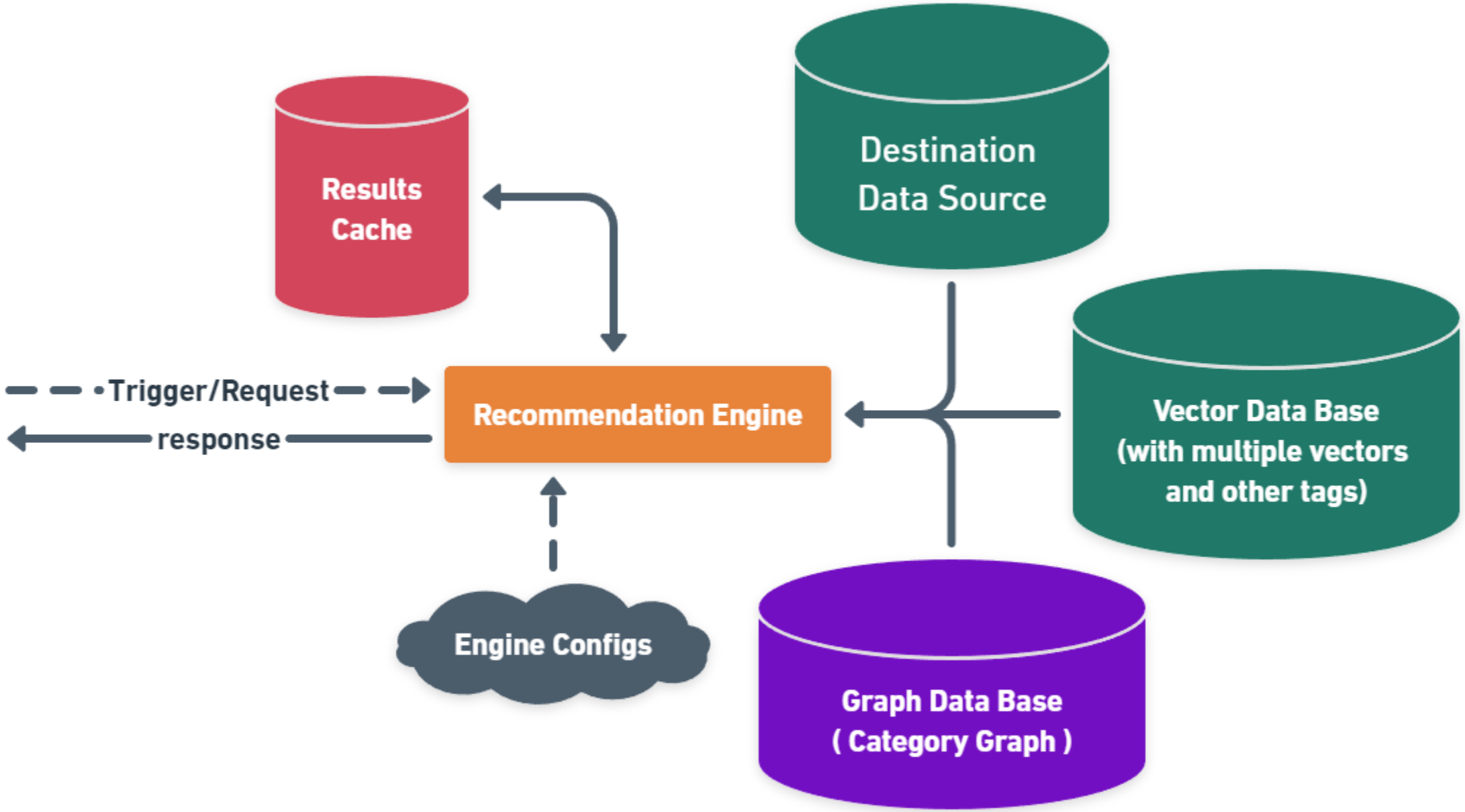


Note: Data Collection, Processing, Storage & Utilization depends on Data & Privacy Policies of the organization. Also the quality of the pattern discovered could be proportional to the amount of data/signals collected.

Pipeline for Content Space:



Pipeline for Context Space:



Suggested Metrics & Caveats:

North Star Metric for Recommendation Feature:

- **Engagement Ratio:** This metric is the ratio between the number of engagements for a content to the number of views garnered by the content. Generally an average of all the content's Engagement Ratio is considered in the industry. But this value could be misleading as the distribution could be severely skewed in most of the cases. Best approach would be an average of Engagement Ratio of contents sampled using stratified sampling.

Other Metrics:

- Daily Consumption Time (Average)
- Session Time (Average, can also be specific to part of the day)
- Number of Sessions per Day (Average)

Certain Caveats:

- If the Data Policy restricts collecting & processing user data, or enforces certain degree of anonymity then approaches like User Segmentation, User Persona Classification can be implemented. But approaches like these requires continuous evaluation & modelling due to gradual changes in the behavioural patterns and hence may not yield good recommendation results for a long time especially if the platform is still growing. Hence continuous model training/upgrades could be the efficient way for the same.
- For a better yield of recommendation results, it becomes very important for Creators to constantly create engaging contents with topics catering to wider audience (atleast during the initial years of the platform)