

AutoSE: Automatisierung der Inhaltserschließung mit Machine-Learning-Methoden an der ZBW

*Impuls zu Software – Betrieb und Administration:
Erstellung und Betrieb eines Kubernetes-Clusters
zur Ausführung von AutoSE-Core*

Christopher Bartz

ZBW – Leibniz-Informationszentrum Wirtschaft

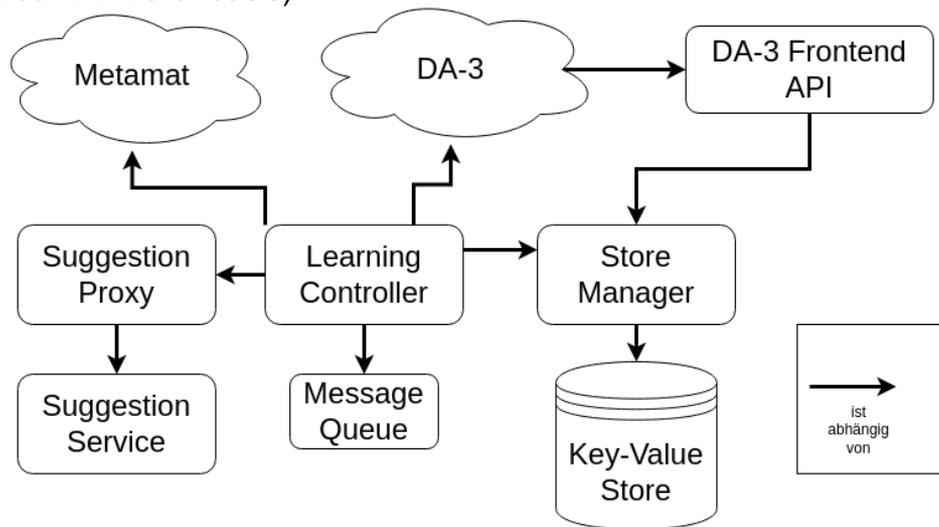
Workshop „Einsatz von KI und DH an Bibliotheken“, Frankfurt am Main, 03. und 04.11.2022



Meilenstein „Aufbau der Grundarchitektur“:



(EconBiz-Datenbasis)



Software für das Produktivsystem:

- **Kubernetes-Cluster** mit 5 Knoten (~ virtuellen Maschinen) 
- wird kontinuierlich weiterentwickelt – inkl. Lösungen für **Monitoring** (*prometheus, grafana*),   **Deployment** (*helm*), **Continuous Integration** (*GitLab*), etc. 

Vorbedingungen

- Inhouse-Betrieb
- Administration durch uns
- Deployment in Containern
- Skalierung und Ausfallsicherheit

Hardware

- Kubernetes-Cluster
 - 4 Worker-Knoten
 - 24 GB RAM
 - 4 CPU
 - 1 Master-Knoten
 - 12 GB Ram
 - 2 CPU



Kubernetes-Installation

- <https://kubernetes.io/docs/setup/production-environment/tools/kubeadm/create-cluster-kubeadm/>
- Alternativen: kommerzielle Produkte wie Redhat OpenShift
- erst Master-Knoten, dann die Worker-Knoten
- Network mit Weave
(<https://www.weave.works/docs/net/latest/kubernetes/kube-addon/>)

Speicher

- Ceph-Cluster
 - 600 GB
 - CephFS mit 3x *replication*
- Rook
 - Ceph wird innerhalb vom Kubernetes-Cluster betrieben
 - direkter, einfacher Zugriff aus den Containern



kubectl

- zeige alle Pods

kubectl -n autose-prod get pods

- restarte ein Deployment

kubectl -n autose-test rollout restart deploy/lcontroller-assessment

- zeige Logs

kubectl logs -n autose-test deploy/lcontroller-suggestion

- skaliere

kubectl -n autose-test scale deploy/lcontroller-suggestion --replicas=0

Deployment

Helm-Paketmanager

- Deployment

```
helm upgrade -n autose-test autose-core /path/to/autose-core/ -f /path/to/values.yml--set
```

```
release_tag=8add5392 --description "Set Qualle Threshold to 0.1"
```

- Rollback

```
helm -n autose-test rollback autose-core 367
```



Monitoring

- Prometheus für Alarme und perspektivisches Monitoring
- Grafana zur Visualisierung
- Elasticsearch – Stack für das Logging
- verschiedene Cronjobs

