



Datenbank Administration neu denken 2.0

Andre Lünsmann und Sebastian Pielawa

29.01.2025

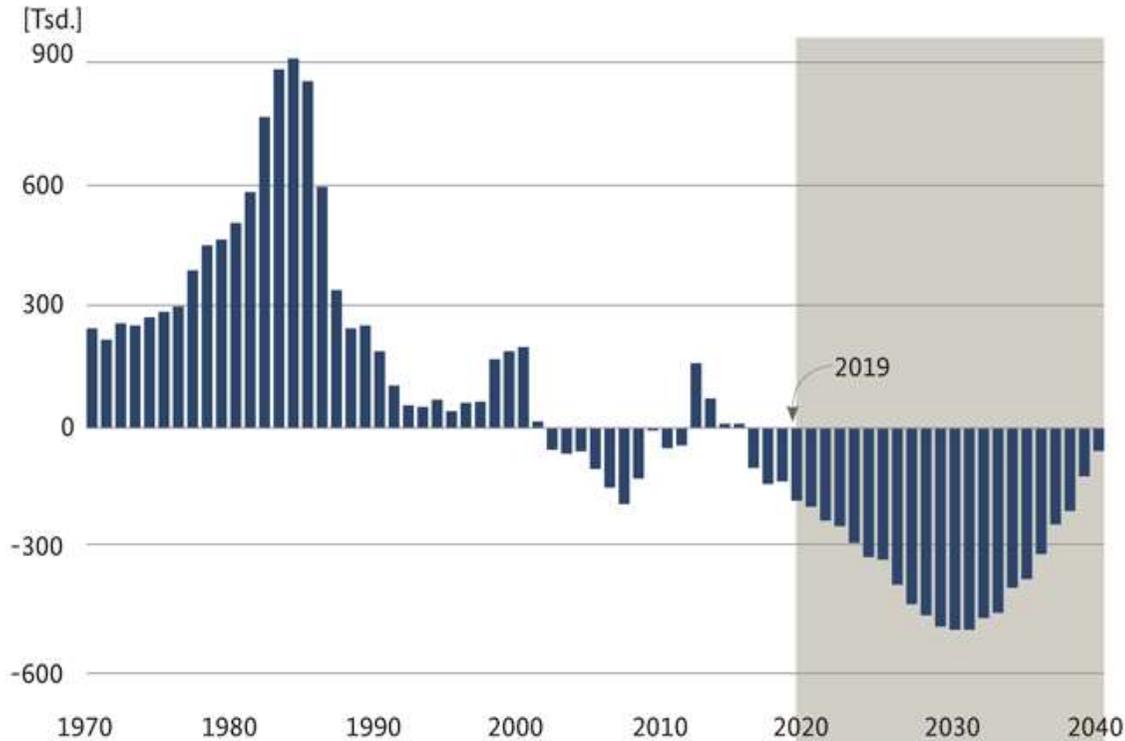
A person in a dark suit is seen from behind, interacting with a large, glowing blue digital interface. The interface displays various data points, charts, and text in a futuristic, slightly blurred font. The overall scene is set against a dark background with a strong blue light source, creating a high-tech, data-driven atmosphere.

Database meets AI

Agenda

- Warum wir uns mit dem Thema beschäftigen
- Überblick unserer Infrastruktur/Rahmenbedingungen
- Vorstellung der Idee zur KI gestützten Lösung
- Rückblick auf die DOAG Konferenz 2023 und den damaligen Ausblick
- Demo der entstandenen, KI gestützten Lösung
- Bestandteile der Anwendung
- Status Quo und Ausblick
- Fazit

Unsere Herausforderungen zu Zeiten des demografischen Wandels



Die Lücke der 18 Jährigen im Vergleich mit den 67 Jährigen ist seit 2019 negativ

Demografische „Lücke“ auf dem Arbeitsmarkt. Jährliche Differenz zwischen 18- und 67-Jährigen.

Daten: Statistisches Bundesamt, 14. Koordinierte Bevölkerungsvorausberechnung, Variante G2-L2-W2 (ab 2019)

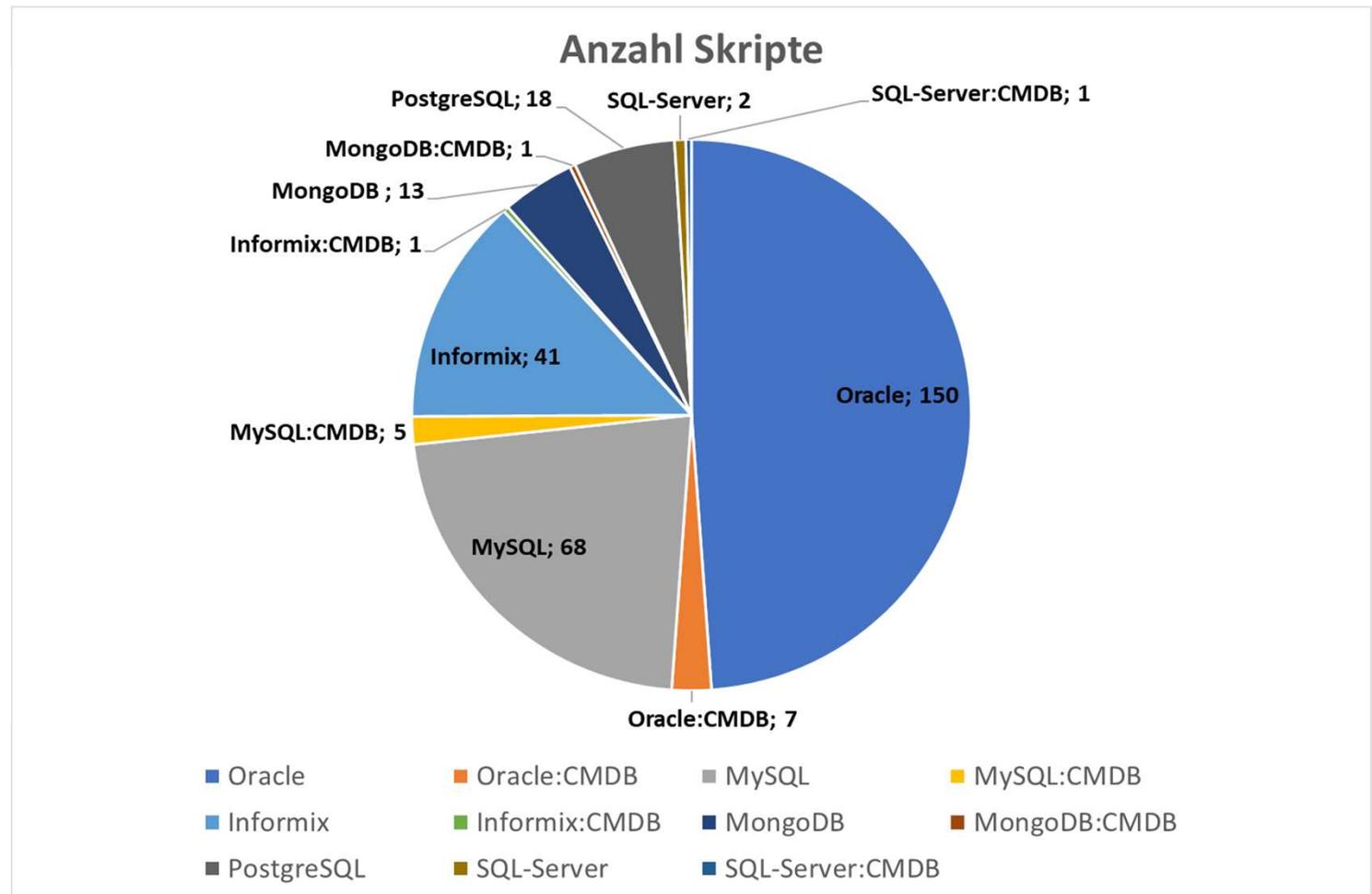
Unsere Herausforderungen zu Zeiten des demografischen Wandels

- Personalsituation
 - Von 5 DBAs sind 3 über 60 mit durchschnittlich 35 Jahren Datenbank Erfahrung.
 - Zum 1.4. konnte 1 neuer DBA eingestellt werden, leider ist zum 1.10. ein anderer DBA ausgeschieden.
- Bis zu den geplanten Renteneintritten verbleibt nur noch wenig Einarbeitungszeit.
- In 2025 beginnt die eigene Ausbildung von Fachinformatikern Systemintegrator.
- Besonders schwierig für die neuen DBAs ist die Erkennung der Ursachen bei auftretenden Fehlern.
- Auf einer KI Veranstaltung im Mai 2023 von mir und einem Kollegen, wurde die Idee einer KI Unterstützung geboren.

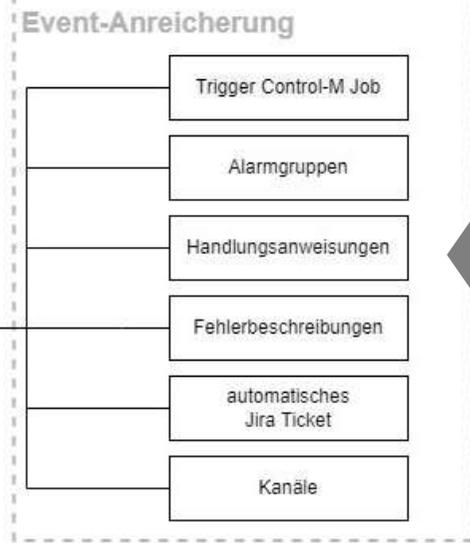
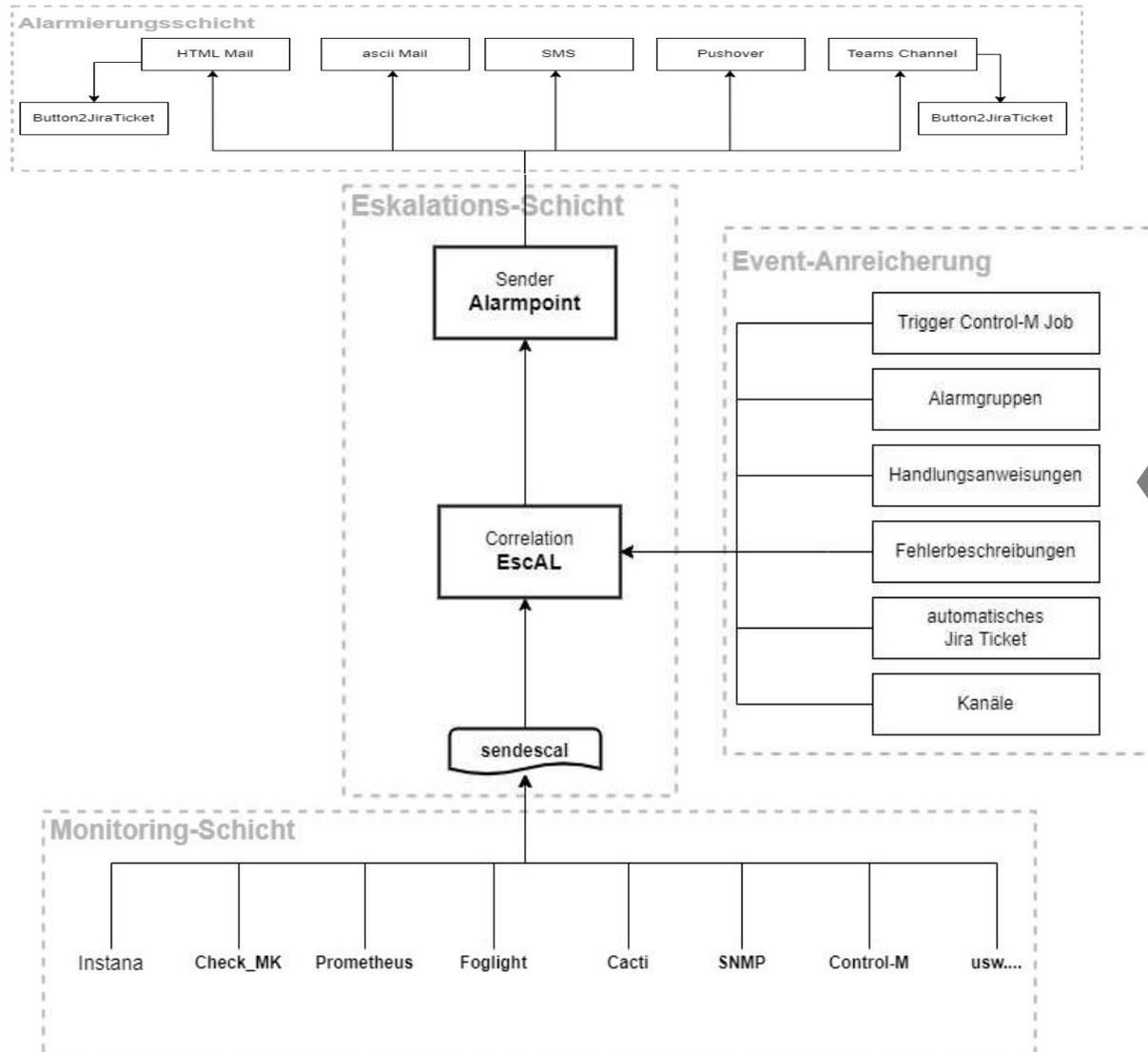
Barmenia Datenbank Administration (Skripte)

Insgesamt wird diese Infrastruktur mit über 300 Skripten automatisiert betrieben.

Für die SQL-Server-Umgebung sind das Powershell-Skripte, für die anderen DBMS Bash- und Perl-Skripte.



Idee einer KI gestützten Hilfestellung mit dem Startpunkt „Alarmierung“

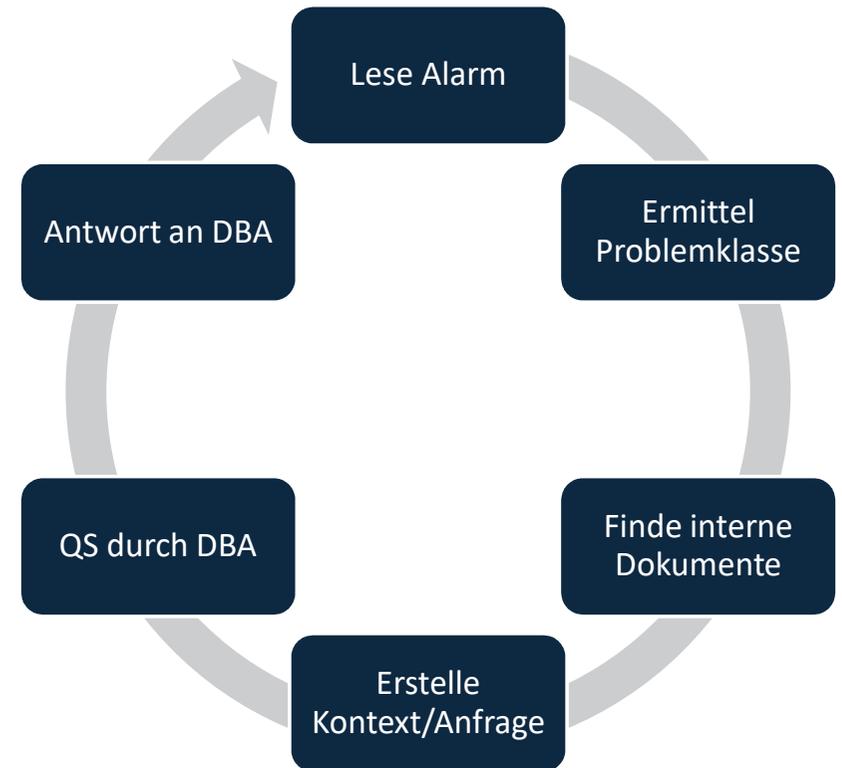


Erweiterung der Event-Anreicherung um Ergebnisse der KI Lösung

unser Ansatz „Auslesen vorhandener Infos, Qualitätscheck der DBA's, Anreicherung um interne Dokumentationen aus einer KI Lösung

• Quellen

- Monitoring Tools
- Confluence Infos
- KI gestützte Lösung
- Von uns präferierte KI Methode
- Überwachtes Lernen



Gruppieren von Alarmmeldungen in Problem-Klassen

- Erkennung aus den Alarmmeldungen
- Ergänzung von Klassifizierung aus den Monitoring Tools
 - Foglight bietet hier schon Lösungsvorschläge im Feld „INSTANCE SITUATION“
 - Auf dieser Basis erfolgte eine Ableitung von eigenen Alias Einträgen durch die DBAs
- Geprüfte Einträge der DBAs werden persistiert und für Anreicherung der Handlungsanweisungen im Event genutzt



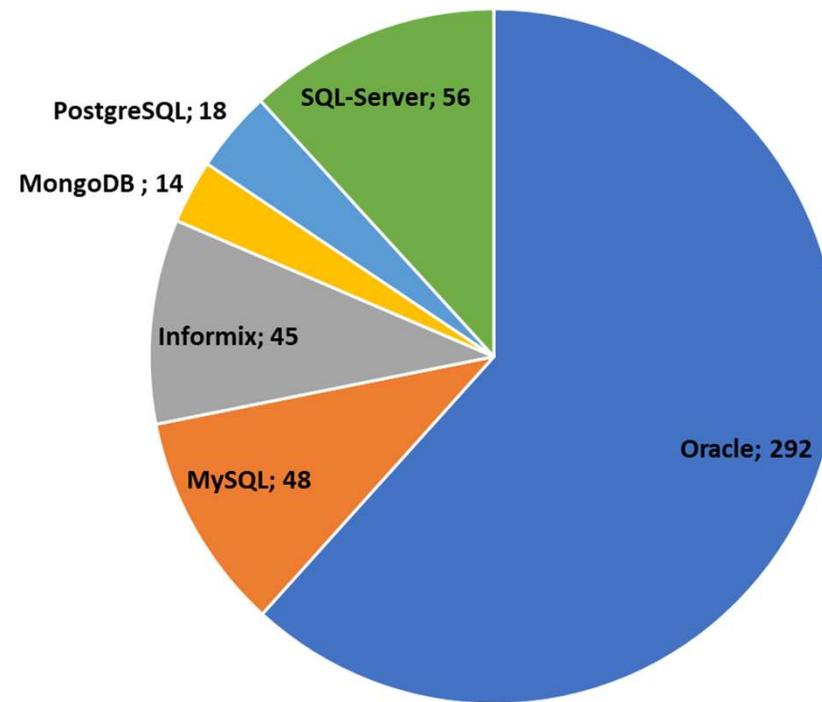
Dokumentationsquelle „Confluence“

Insgesamt haben wir ca. 770 Confluence Seiten mit Informationen zur Administration.

Den einzelnen Datenbank Systemen sind zwischen 14 und 292 Seiten konkret zugeordnet.

Darüber hinaus sind eine Vielzahl von Dateien verknüpft oder hochgeladen in Confluence.

Anzahl Confluenceseiten



■ Oracle ■ MySQL ■ Informix ■ MongoDB ■ PostgreSQL ■ SQL-Server

Mapping von Confluenceseiten zu den Alarmmeldungen

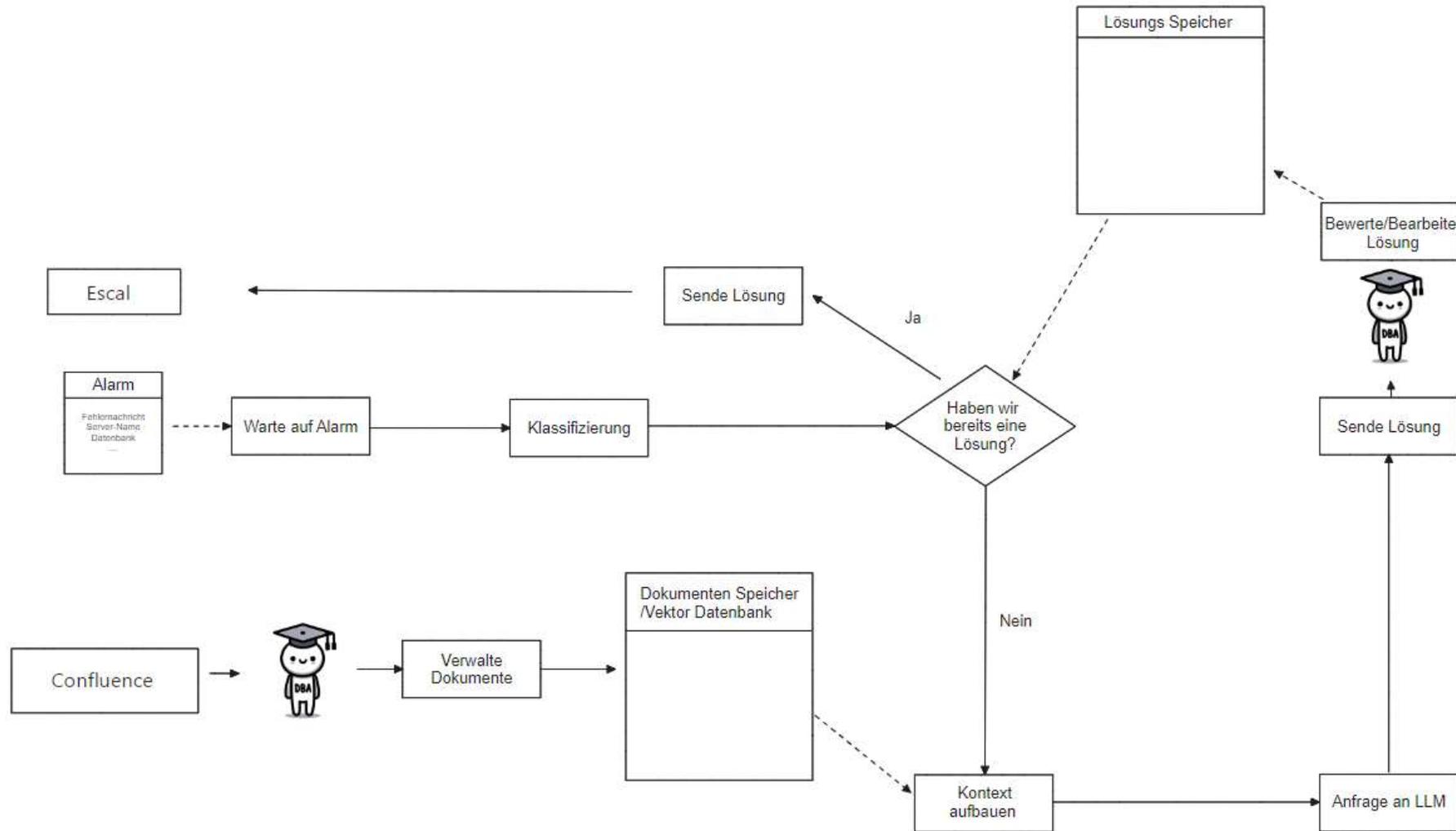
- Aktuell
 - Mapping von Alarmmeldungen auf unsere Confluenceseiten
 - Empfehlung der im Internet gefundenen ChatGPT Lösungsansätze
- Demnächst
 - Einbindung von externen Quellen (z.B. Oracle Doku, Blog Einträge o.ä.)



Rückblick DOAG Konferenz 2023

- Zunächst müssen wir alle Informationsquellen für die KI Lösung erschließen.
- Dann werden wir in Iterationen die Qualität der Auswertungsergebnisse sukzessive verbessern, bis wir gute Ergebnisse aus der KI Lösung bekommen.
- Anschließend soll die Eventanreicherung mit den Vorschlägen der KI in das Feld „Handlungsanweisung“ übertragen und bei Alarmmeldungen direkt mitgeschickt werden.
 - Sensible Daten dürfen nicht ins Internet übermittelt werden, daher sind solche Informationen vorher aus den Datenquellen zu extrahieren.
- Neben der Handlungsanweisung stellen wir uns die Erstellung einer interaktiven Lösung vor, die in einem Chat mit dem DBA über die KI Lösungsansätze finden soll.
- In einem Folgeschritt sollen automatisierte Lösungen ausgeführt werden um Probleme zu beheben.

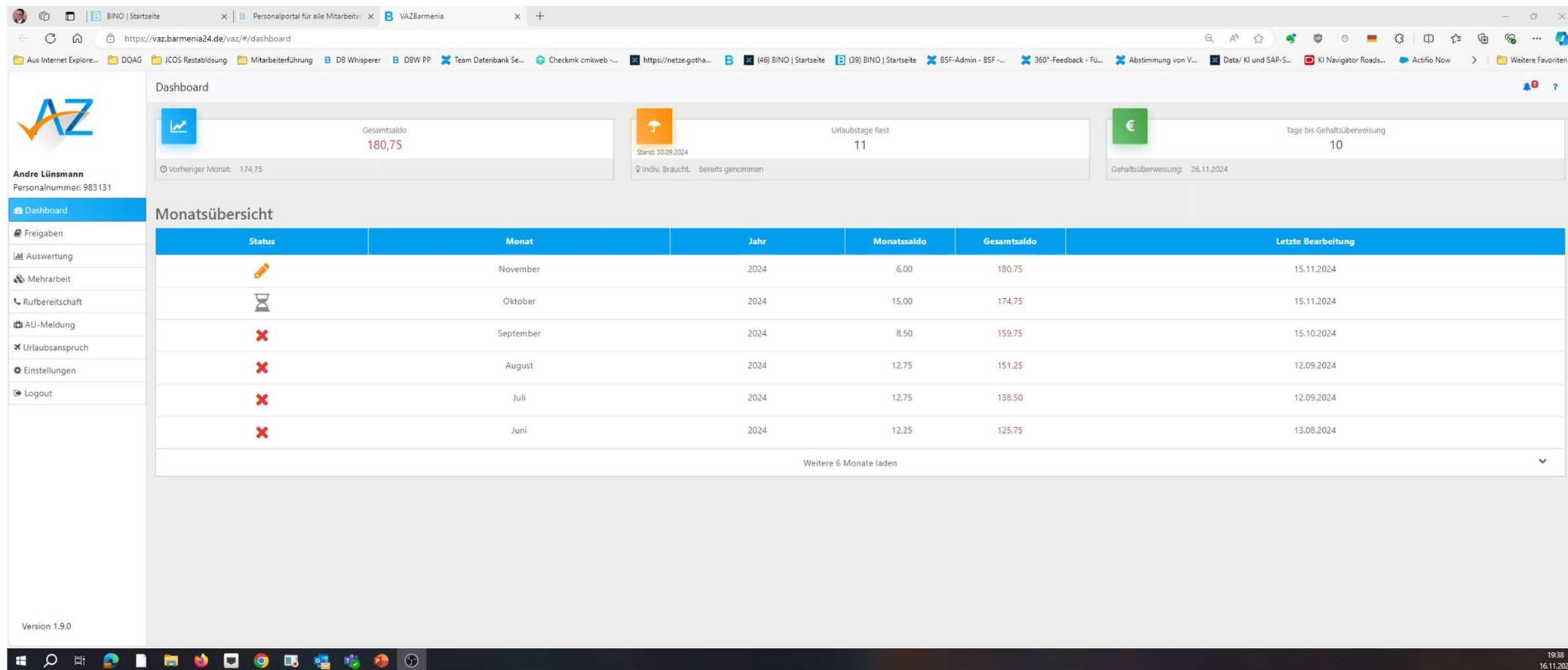
Prozessablauf



Database Whisperer

- Das Event Archive Datensenke mit den aufgetretenen Events.
- Die Queue Events, die im Archiv mit „Solve“ markiert wurden.
- Der Solution Store Lösungstexte, die in der Queue „geCheckt“ wurden.
- Documents eingelesene Dokumente aus der Vector Datenbank.
- Retrieve Abruf von Treffern aus der Vector Datenbank.
- Alarmmeldung angereichert um den Link auf die Solution des DBW.

Event Archive



Dashboard

Gesamtsaldo
180,75
Vorheriger Monat: 174,75

Urlaubstage Rest
11
Stand: 30.09.2024
Indiv. Braucht: bereits genommen

Tage bis Gehaltsüberweisung
10
Gehaltsüberweisung: 26.11.2024

Monatsübersicht

Status	Monat	Jahr	Monatssaldo	Gesamtsaldo	Letzte Bearbeitung
	November	2024	6,00	180,75	15.11.2024
	Oktober	2024	15,00	174,75	15.11.2024
	September	2024	8,50	159,75	15.10.2024
	August	2024	12,75	151,25	12.09.2024
	Juli	2024	12,75	138,50	12.09.2024
	Juni	2024	12,25	125,75	13.08.2024

Weitere 6 Monate laden

Version 1.9.0

Selektion und Prüfung von gemeldeten Events, die einer Solution zugeordnet werden können

Queue

DBW Event Archive Queue Solution Store Import Documents Documents Retrieve

Monitor	Timestamp	Criticality	Component-Type	Component	Sub-Component	Instance-Situation	Source	Instance-Id	QuickSolveable	KI-Alias	Confidence
EqoLight	16.11.2024 19:28:42	prod	Database	Oracle	DBO_Jobs_Queue_Db	DBO - Job Queue	ixsassa2p.barmenia.lan	ixsassa2p.barmenia.lan-CDWHHP_51	✗	ARCHIVE_DEST_STATUS	23.00%
Job Queue. The number of jobs that are waiting in the queue is 12.00.											
EqoLight	16.11.2024 16:38:32	prod	Database	Oracle	DBO_Jobs_Queue_Db	DBO - Job Queue	ixsassa2p.barmenia.lan	ixsassa2p.barmenia.lan-CDWHHP_51	✗	ARCHIVE_DEST_STATUS	23.00%
Job Queue. The number of jobs that are waiting in the queue is 13.00.											
EqoLight	16.11.2024 16:38:32	prod	Database	Oracle	DBO_Jobs_Queue_Db	DBO - Job Queue	ixsassa2p.barmenia.lan	ixsassa2p.barmenia.lan-CDWHHP_51	✗	ARCHIVE_DEST_STATUS	23.00%
Job Queue. The number of jobs that are waiting in the queue is 13.00.											
EqoLight	16.11.2024 12:23:19	prod	Database	Oracle	DBO_Jobs_Queue_Db	DBO - Job Queue	ixsassa2p.barmenia.lan	ixsassa2p.barmenia.lan-CDWHHP_51	✗	ARCHIVE_DEST_STATUS	23.00%
Job Queue. The number of jobs that are waiting in the queue is 13.00.											
EqoLight	16.11.2024 12:23:19	prod	Database	Oracle	DBO_Jobs_Queue_Db	DBO - Job Queue	ixsassa2p.barmenia.lan	ixsassa2p.barmenia.lan-CDWHHP_51	✗	ARCHIVE_DEST_STATUS	23.00%
Job Queue. The number of jobs that are waiting in the queue is 13.00.											
EqoLight	16.11.2024 09:23:04	prod	Database	Oracle	DBO_Jobs_Queue_Db	DBO - Job Queue	ixsassa2p.barmenia.lan	ixsassa2p.barmenia.lan-CDWHHP_51	✗	ARCHIVE_DEST_STATUS	23.00%
Job Queue. The number of jobs that are waiting in the queue is 12.00.											
EqoLight	16.11.2024 09:23:04	prod	Database	Oracle	DBO_Jobs_Queue_Db	DBO - Job Queue	ixsassa2p.barmenia.lan	ixsassa2p.barmenia.lan-CDWHHP_51	✗	ARCHIVE_DEST_STATUS	23.00%
Job Queue. The number of jobs that are waiting in the queue is 12.00.											
EqoLight	16.11.2024 08:04:54	prod	Database	Oracle	DBO_SGA_Components	DBO - Large Pool Full	ixsassa2p.barmenia.lan	ixsassa2p.barmenia.lan-CDWHHP_51	✗	LARGE_POOL_STATUS	100.00%
Large Pool Full. The utilization of the Large Pool is 95.78%.											
EqoLight	16.11.2024 08:04:54	prod	Database	Oracle	DBO_SGA_Components	DBO - Large Pool Full	ixsassa2p.barmenia.lan	ixsassa2p.barmenia.lan-CDWHHP_51	✗	LARGE_POOL_STATUS	100.00%
Large Pool Full. The utilization of the Large Pool is 95.78%.											
EqoLight	16.11.2024 06:09:22	prod	Database	Oracle	DBO_Archive_Destination_Misc_Data	DBO - Redo Archive Time to Fail	ixorau6a.barmenia.lan	ixorau6a.barmenia.lan-PSHP	✓	ARCHIVE_DEST_STATUS	100.00%

Events, die im Archiv mit „Solve“ markiert wurden

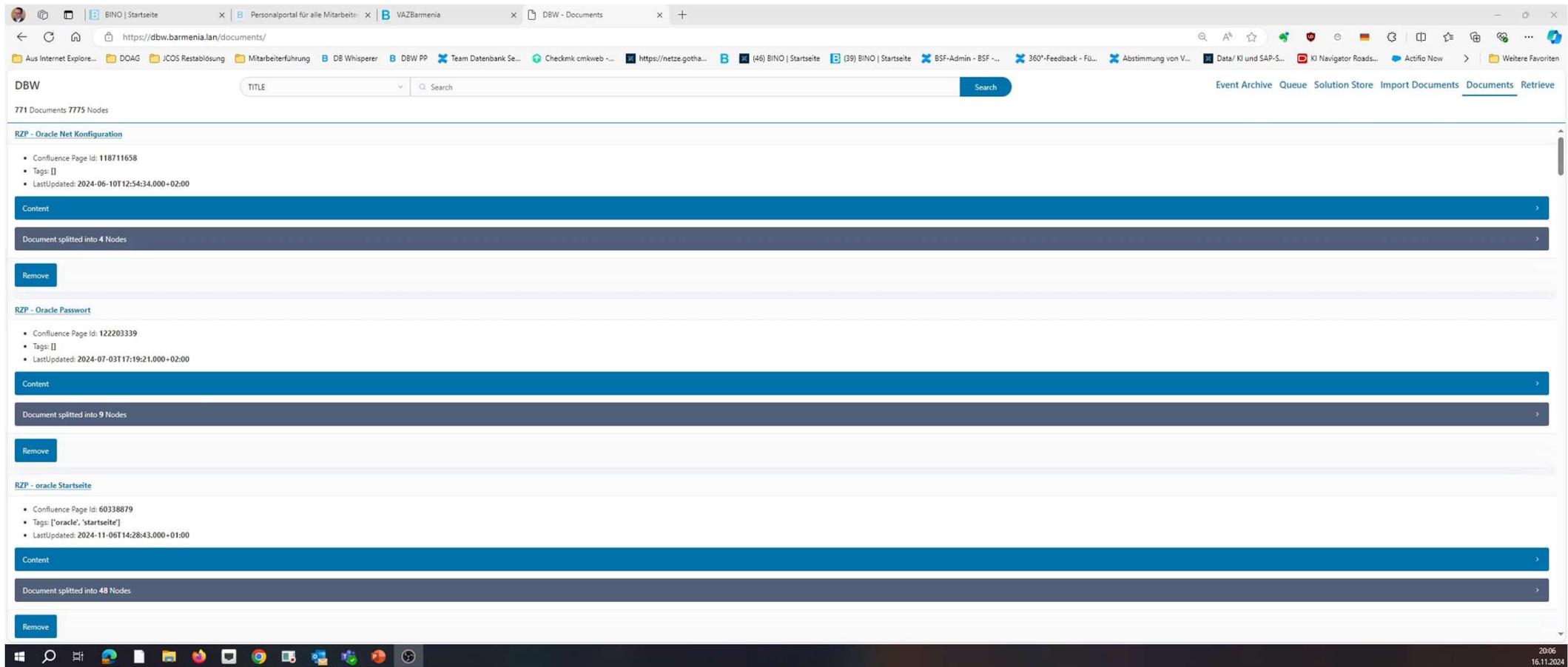
Solution Store



The screenshot displays the DBW Solution Store interface. The top navigation bar includes 'Event Archive', 'Queue', 'Solution Store', 'Import Documents', 'Documents', and 'Retrieve'. The main content area is a table of monitoring events. The columns are: Monitor, DbType, Status, Runtime, Checked, QuickSolved, Source, Ki-Alias, and Confidence. The 'Checked' and 'QuickSolved' columns contain red 'X' marks, indicating that the solutions have been applied. The 'Confidence' column shows a value of 26.30% for most events. The 'Source' column lists various server names like 'ixmncds2ua.barmenia.lan'. The 'Ki-Alias' column shows 'CONNECTION_FAILED' for most events. The 'Monitor' column contains detailed error messages, such as 'The server ixmncds2ua.barmenia.lan:27017 in replica set rscdsua0 has changed state from SECONDARY to PRIMARY.' and 'Ratio of page faults to total operations on ixmncds1ua.barmenia.lan:27017 is high: 2392.00 / 2156.00'. The interface also includes a search bar, a filter dropdown, and buttons for 'Check', 'Flag', and 'Details' for each event.

Lösungstexte, die in der Queue „geCheckt“ wurden

Documents



eingelene Dokumente aus der Vector Datenbank und deren Zerlegung

Retrieve



DBW

771 Documents 7775 Nodes

RZP - Oracle Net Konfiguration

- Confluence Page Id: 118711658
- Tags: []
- LastUpdated: 2024-06-10T12:54:34.000+02:00

Content

Document splitted into 4 Nodes

Remove

RZP - Oracle Passwort

- Confluence Page Id: 122203339
- Tags: []
- LastUpdated: 2024-07-03T17:19:21.000+02:00

Content

Document splitted into 9 Nodes

Remove

RZP - oracle Startseite

- Confluence Page Id: 60338879
- Tags: ['oracle', 'startseite']
- LastUpdated: 2024-11-06T14:28:43.000+01:00

Content

Document splitted into 48 Nodes

Remove

Abruf von Treffern aus der Vector Datenbank

Alarmmeldung



CRITICAL

Eventmeldung

DBW

SOLVE_TASK - 69334310-7123-4b3b-98a4-67f6cd59d2ec

Status: **DONE**
Monitor: **Check_MK**
DbType: **MYSQL**
Source: **lxmyodp1p.barmenia.lan**
Started: **2024-11-13 10:54:13.906379+01:00**
Ended: **2024-11-13 10:58:31.262112+01:00**
Runtime: **0:02:00.646258**

KI-Alias: **DISK_USAGE(100.0%)** Flag

Message: **fs/app Used: 52.95% - 208 GiB of 394 GiB (warn/crit at 300 GiB/329 GiB used), trend per 1 day 0 hours: 68.9 GiB, trend per 1 day 0 hours: 17.51%, Time left until disk full: 2 days 16 hours (warn/crit below 3 days 0 hours/1 day 0 hours)()**

Solution:

- The error message indicates that there is a disk space issue with the file system where the database is located. The available space is below the warning threshold of 300 GiB, and the trend is increasing at a rate of 17.51% per day, which will lead to the disk being full in 2 days 16 hours.
- The context data shows that the database is using approximately 208 GiB of the available disk space.
- To solve the issue, you need to free up some disk space on the file system, here are some steps you can take:
 - Identify large files or directories that can be deleted or moved to another storage location. You can use the "find" command with the "-size" option to locate large files. For example, to find files larger than 1 GiB, you can use the following command:

```
find /path/to/directory -type f -size +1G -exec ls -lh {} \;
```
 - Optimize the database by removing unnecessary data, such as old backups, transaction logs, or index statistics. You can use the database management system's built-in tools to perform these tasks. For example, in Oracle Database, you can use the DBMS_REDEFINE_OB3 package to drop unused objects, and the DBMS_MONITOR package to analyze the database and identify potential performance issues.
 - If the issue persists, you may need to consider upgrading the file system or adding more disk space to it.
- Once you have freed up some disk space, monitor the database and the file system regularly to prevent similar issues from occurring in the future. You can use tools such as the database management system's performance monitoring features or system monitoring tools like "top" or "htop" to keep track of disk usage and database performance.

Task Log:

2024-11-13 10:54:13.906379+01:00 - **START**

```
{
  "alertkey": "fs_/app",
  "component": "Linux",
  "componenttype": "Operatingsystem",
  "criticality": "prod",
  "escalstatus": 0,
  "eventcount": 1,
  "eventtime": "2024-11-13T09:54:12.447000Z",
  "extendedattr": "7/wato/linux/003/ auto-piggyback checkmk-agent cmk-agent ip-v4 ip-v4-only lan lx no-smpp opsdatabase_mysql p prod pss site:cmkweb tcp",
  "id": null,
  "instanceid": "/app",
  "instancesituation": "check_mk-df",
  "instancevalue": "",
  "lastmodified": "2024-11-13T09:54:12.447000Z",
  "message": "fs_/app Used: 52.95% - 208 GiB of 394 GiB (warn/crit at 300 GiB/329 GiB used), trend per 1 day 0 hours: 68.9 GiB, trend per 1 day 0 hours: 17.51%, Time left until disk full: 2 days 16 hours (warn/crit below 3 days 0 hours/1 day 0 hours)()",
  "messageid": null,
  "messagelevel": null,
  "messageurl": "https://cmk-web.unix.lan/cmweb/check_mk/index.py?start_url=view.py?view_name%3Dservice%26host%3Dlxmyodp1p.barmenia.lan%26service%3Dfs_%252fapp%26site%3Dcmkweb",
  "notifygroup": null
}
```

Bestandteile der Anwendung

- KI Modelle

- Embeddingmodell zum erstellen der Vektoren - Sentence-Transformer(all-minilm-l6-v2)
- Sprachmodell zum generieren der Texte - Mistral 7B Instruct 0.2
 - Llama.cpp um die Sprachmodelle zu benutzen

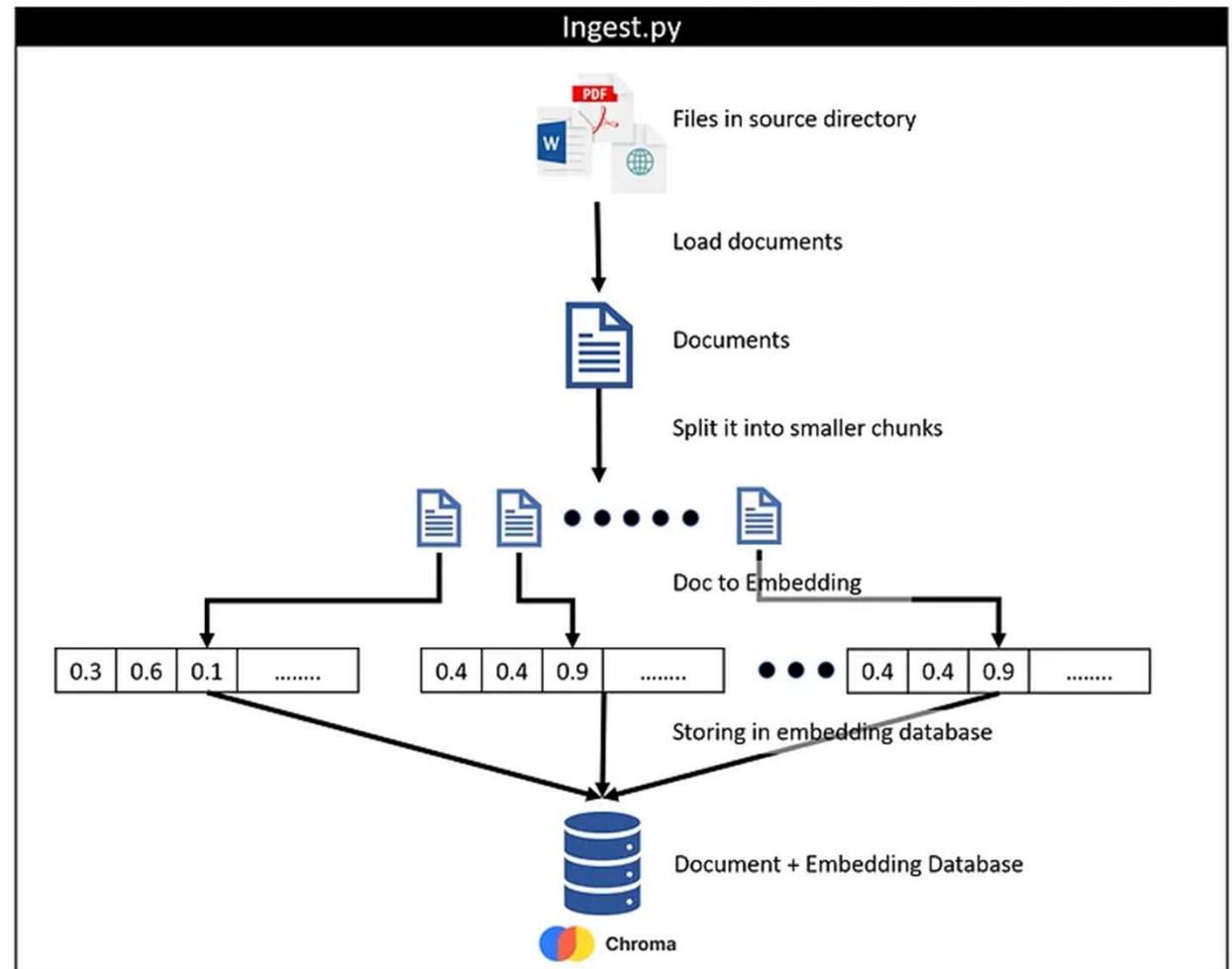
- Postgres Datenbank mit Vector Erweiterung

- CSS/Javascript Bibliotheken

- Pico
- HTMX
- Tabulator
- Luxon

Sentence Transformer

- Der Sentence Transfer erkennt Ähnlichkeiten der eingehenden Alarmmeldungen und clustert diese.
- Im Event Archive werden die geclusterten Events mit den gefundenen Lösungsansätzen aufgelistet und gefiltert.
- Die DBAs bewerten die Lösungsvorschläge und legen einen KI Alias und Solution fest.



- Large Language Mode (LLM) mit Mistral 7B

Mistral 7B ist Projekt und gleichnamiger Textgenerator des französischen KI Startups Mistral AI.

Die Besonderheit im Vergleich zur Konkurrenz ist, dass das Sprachmodell mit einer Größe von „nur“ knapp 7,3 Milliarden Parametern arbeitet und trotzdem sehr leistungsstark ist.

- Vorteil: Apache 2 Lizenz
- Llama.cpp ist eine C++ Bibliothek, welche es ermöglicht die Sprachmodelle relativ effizient auf CPUs laufen zu lassen.

Infos z.B. auf Youtube

[Mistral 7B -The Most Powerful 7B Model Yet](#)  

Postgres Datenbank mit Vector Erweiterung

Warum PostgreSQL mit der Extension pgvector verwendet

- Es ist Open source.
- Ist eine etablierte Datenbank.
- Die Extension ist weitverbreitet beim Arbeiten mit Vektor-Suche.
- Es ist keine extra Vektordatenbank notwendig
- Es können weiterhin alle Features eines RDBMS verwendet werden.
- Die Vektorfunktion von pgvector wird durch SQL bedient, heißt jeder Sql-client kann diese in Zukunft verwenden.

Infos findet man hier: <https://github.com/pgvector/pgvector>

CSS/Javascript Bibliotheken

- Pico
 - Minimales CSS-Framework für semantisches HTML
 - Ein minimalistisches und leichtes Starterkit, das die semantische Syntax priorisiert und jedes HTML-Element standardmäßig reaktionsfähig und elegant macht.
- HTMX
 - HTMX ermöglicht Ihnen den Zugriff auf AJAX, CSS-Übergänge, WebSockets und Server Sent Events direkt in HTML unter Verwendung von Attributen, sodass Sie moderne Benutzeroberflächen mit der Einfachheit und Leistungsfähigkeit von Hypertext erstellen können.
- Tabulator
 - Eine benutzerfreundliche JavaScript-Bibliothek zur interaktiven Tabellengenerierung.
- Luxon
 - Luxon ist eine Bibliothek zum Arbeiten mit Datum und Uhrzeit in JavaScript.

Berechtigungskonzept

Zuordnung der Rechte zu fachlichen Rollen

Die Rollen-Berechtigungsobjekt-Matrix zeigt die eindeutige Zuordnung von Berechtigungsobjekten zu den verwalteten fachlichen Rollen auf.

	DBW_Admin	DBW_DBA	DBW_EventSubmitter	DBW_ReadOnly
Anmeldung und Administration des DBW Servers	X			
Eventbearbeitung (Zuordnung von Events zu vorhandenen Solutions)	X	X		
Pflege der Solutions	X	X		
Pflege von Aktionen zur Fehlerbehebung per Skript bzw. Control-M Job	X	X		
Übermittlung von Events an den DBW (Escal)			X	
reine Leserechte mit dem Recht zur Ansicht der Daten in der DBW Oberfläche (z.B. für Revision)	X	X		X

Status Quo und Ausblick für 2025

- ✓ Historie der Probleme der einzelnen Server/Datenbanken persistieren um mehr Masse für die Lernmodelle vorzuhalten
- ✓ Fortlaufende Verfeinerung des Einsatzes des Sentence Transformers
- ✓ Tägliche Nutzung des Database Whisperers, die dem DBA die Möglichkeit zur Abfrage von Lösungsansätzen der KI bietet.
- Weitere Quellen (z.B. Jira) über API in die KI Modell laden und mit den Events und Solutions zusammen zu bringen (geplant in Q1 2025)
- Schaffung einer Schnittstelle für die revisionssichere Ausführung von Skripten für Fehlerbehebungen, die den Solutions zugeordnet werden (geplant für Q2 2025)
- Erweiterung auf weitere Bereiche wie z.B. OS, Storage, Anwendungen etc. in einer weiteren Ausbaustufe

Fazit

- Mit der neuen Anwendung und den KI gestützten Lösungsansätzen haben wir in Kombination mit der Qualitätssicherung durch die DBAs eine gute Basis für eine schnellere Fehlerbearbeitung/-behebung geschaffen.
- Die Anwendung unterstützt auch für allgemeine Fragen und der Suche nach Dokumentation in den angebundenen Quellen im Tagesgeschäft.
- Neue oder gefundene, alte Dokumentation kann adhoc in die Modelle nachgeladen werden.
- Über die Anwendung können wir sukzessive die Modellergebnisse ansehen, hinterfragen und ggfs. die Konfiguration der KI Modelle optimieren.



Nur Mut!
Starten Sie mit KI
in eine neue Zukunft
mit dem 12. Mann,
es ist mehr Arbeit da,
als wir bewältigen können!

Andre Lünsmann & Sebastian Pielawa

