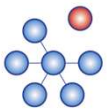
Proaktives Monitoring KI-Anomalieerkennung & GenAI/RAG für die Ursachenanalyse

DOAG Datenbank, 15. Mai 2025
Raum Versailles / Europapark Rust



aprevis
EXPLAINABLE AI

aprevis
EXPLAINABLE AI

Über mich

Felix Castillo Sanchez

Email: felix.castillo@aprevis.ai
Web: https://aprevis.ai

- Oracle Consultant seit 1991
- Freelancer seit 1997
- Tätig als: Consultant / DBA / Entwickler / Trainer

- Startup aprevis ab 2026

Über mich

- DOAG 2011
  - Oracle Performance Analyse mit statistischen Methoden
- DOAG Regio 2012
  - Oracle Performance Analyse
    Die Wiederentdeckung der Oracle Statistiken
- DOAG 2012
  - Oracle Performance Analyse
    Erweiterte Möglichkeiten mit Statistiken und Wartezeiten
- DOAG SIG 2013
  - Oracle Performance Analyse
    Quo Vadis, AWR?
- Hotsos Symposium, März 2013 in Dallas, USA
  - The Oracle Performance Signature
- Frankfurter Datenbanktage, März 2013
  - Die Oracle Performance Signatur
- DOAG 2013
  - Die Oracle Performance Signatur
- DOAG 2022
  - Von der Performance Signatur zu AIOPS
- **DOAG Datenbank 2025**
  - **Proaktives Monitoring – KI-Anomalieerkennung GenAI/RAG für die Ursachenanalyse**

aprevis
EXPLAINABLE AI

Herausforderungen im IT-Monitoring

- Zunehmende Komplexität moderner IT-Infrastrukturen
- Hoher manueller Aufwand bei der Problemidentifikation
- Limitierte Werkzeuge und eingeschränkte Metriken
- Notwendiges tiefes Fachwissen (z.B. Datenbank- und Oracle-Know-how)

Die zunehmende Komplexität moderner IT-Infrastrukturen erfordert einen Wandel im Monitoring: Herkömmliche Lösungen basieren oft auf starren Schwellwerten und einer begrenzten Auswahl an Metriken. Das führt zu reaktiven Maßnahmen, Fehlalarmen und unbeachtet kritischen Zuständen. Gleichzeitig ist häufig spezialisiertes Expertenwissen nötig, um Probleme schnell zu erkennen und zu beheben. Moderne KI-Methoden wie Deep Learning und Large Language Models können hier Abhilfe schaffen, indem sie aus großen Mengen von Zeitreihendaten automatisiert Anomalien identifizieren und über textbasierte Interaktionen (z. B. mittels Retrieval-Augmented Generation) eine fundierte Ursachenanalyse ermöglichen. Auf diese Weise lassen sich verschiedene Systemkomponenten ganzheitlich betrachten und ein proaktives Monitoring etablieren, das selbst in dynamischen, verteilten IT-Umgebungen stabile und verlässliche Ergebnisse liefert.

Externe Quellen (Auswahl):

- Gartner. (2021). *Magic Quadrant for AIOps Platforms*
- Google. (2018). *Site Reliability Engineering (SRE) Book*. O'Reilly
- Forrester. (2022). *The Forrester Wave™: Artificial Intelligence for IT Operations*

Nachteile bestehender Monitoring-Lösungen

- Starre Alarmschwellen (häufig “One-Size-Fits-All”-Ansätze)
- Fehlende Integration domänenspezifischen Wissens
- Oft isolierte Betrachtung weniger Metriken
- Reaktives statt proaktives Vorgehen bei Anomalien

Viele herkömmliche IT-Monitoring-Lösungen sind in ihrer Herangehensweise an die Fehlersuche und -erkennung recht begrenzt. Ein häufig anzutreffendes Problem ist die Beschränkung auf vordefinierte Kennzahlen und starre Alert-Mechanismen. Solche Systeme lösen Warnungen anhand festgelegter Grenzwerte aus, die oft nicht an veränderte Rahmenbedingungen angepasst werden. Dadurch kann es zu einer großen Anzahl von Fehlalarmen kommen, während tatsächlich kritische Situationen übersehen werden.
Zudem bleiben viele Monitoring-Lösungen auf wenige Datenquellen beschränkt, ohne die Möglichkeit, neue oder zusätzliche Metriken einzubinden. Dies erschwert eine umfassende Ursachenanalyse, insbesondere wenn komplexe Wechselwirkungen zwischen verschiedenen Komponenten in einer Hybrid- oder Multi-Cloud-Umgebung auftreten. Ein weiteres Manko besteht darin, dass bestehende Systeme selten integrierte KI- oder Machine-Learning-Komponenten aufweisen und sich kaum mit domänenspezifischen Daten kombinieren lassen. So wird eine tiefergehende Analyse, etwa im Bereich Datenbanken, Netzwerke oder Container-Orchestrierung, nur unzureichend unterstützt.
Ein typischer Nachteil traditioneller Monitoring-Lösungen ist auch die mangelnde Unterstützung bei der Kollaboration. IT-Teams müssen oftmals auf unterschiedlichen Plattformen agieren, was zu Informationssilos führt und die Zusammenarbeit behindert. Hinzu kommt der hohe manuelle Aufwand für Auswertung und Dokumentation, da Daten meist nicht automatisch ausgewertet oder kontextbezogen aufbereitet werden. In vielen Fällen entsteht somit ein reaktiver „Feuerwehrmodus“: Probleme werden erst dann ersichtlich, wenn sie bereits zu spürbaren Beeinträchtigungen geführt haben, was Ausfallzeiten und Kosten in die Höhe treibt.

Externe Quellen (Auswahl):

- Gartner. (2021). *Market Guide for IT Infrastructure Monitoring Tools*
- IDC. (2022). *Enterprise IT Infrastructure Strategies: Monitoring and Analytics*
- Forrester. (2022). *The Forrester Wave™: AIOps Platforms*

Motivation für KI-basierte Anomalieerkennung

- Automatisierte Erkennung von Nutzungsmustern in Zeitreihendaten
- Skalierbarkeit bei wachsender Zahl von Systemen und Metriken
- Kontinuierliche Verbesserung durch selbstlernende Modelle
- Minimierung menschlicher Fehler bei der Analyse

Die steigende Komplexität moderner IT-Landschaften erfordert innovative Ansätze, um Performance-Einbußen und potenzielle Risiken frühzeitig zu erkennen. In diesem Zusammenhang bietet KI-basierte Anomalieerkennung erhebliche Vorteile gegenüber konventionellen Methoden.
Machine-Learning-Modelle ermöglichen es, in großen Mengen von Metriken und Messwerten Muster zu identifizieren, die für menschliche Analysten nur schwer erkennbar sind. Indem das System kontinuierlich aus vorhandenen Daten lernt und sich an verändernde Bedingungen anpasst, lassen sich Abweichungen von normalen Betriebszuständen selbst in Echtzeit erkennen.
Ein zentraler Vorteil dieser Herangehensweise besteht in der hohen Flexibilität. Statt lediglich fixe Grenzwerte zu überwachen, kann ein KI-Modell dynamische Schwellwerte (Thresholds) automatisch an die tatsächliche Systemauslastung anpassen. Dies reduziert das Risiko von Fehlalarmen und erhöht zugleich die Wahrscheinlichkeit, unbekannte oder sehr seltene Fehlerbilder zu entdecken. Darüber hinaus ermöglicht die Automatisierung der Analysen eine schnellere Diagnose, was Ausfallzeiten minimiert und Ressourcen schont. Somit unterstützt KI-basierte Anomalieerkennung nicht nur die Stabilität des Betriebs, sondern liefert wertvolle Einblicke in die Ursachen von Störungen und trägt so zu einer nachhaltig verbesserten Systemarchitektur bei.

Externe Quellen (Auswahl):

- Gartner. (2021). *Magic Quadrant for AIOps Platforms*
- MIT Sloan Management Review. (2020). *The Strategic Benefits of Artificial Intelligence in Business*
- McKinsey & Company. (2022). *The AI Advantage: Leveraging Data and Analytics*

Grundlagen Zeitreihendaten am Beispiel Oracle

- Wie entstehen Zeitreihendaten (System Metriken, Performance-Indikatoren)?
- Beispiele: CPU-Auslastung, I/O, Netzwerk, Sessions, Tablespace-Statistiken, Wartezeiten
- Relevanz für Performance- und Kapazitätsplanung

Timestamp	Metric	Value
2025-05-10 10:33:15.334	Average Active Sessions	10.8
2025-05-10 10:33:45.228	Average Active Sessions	9.7

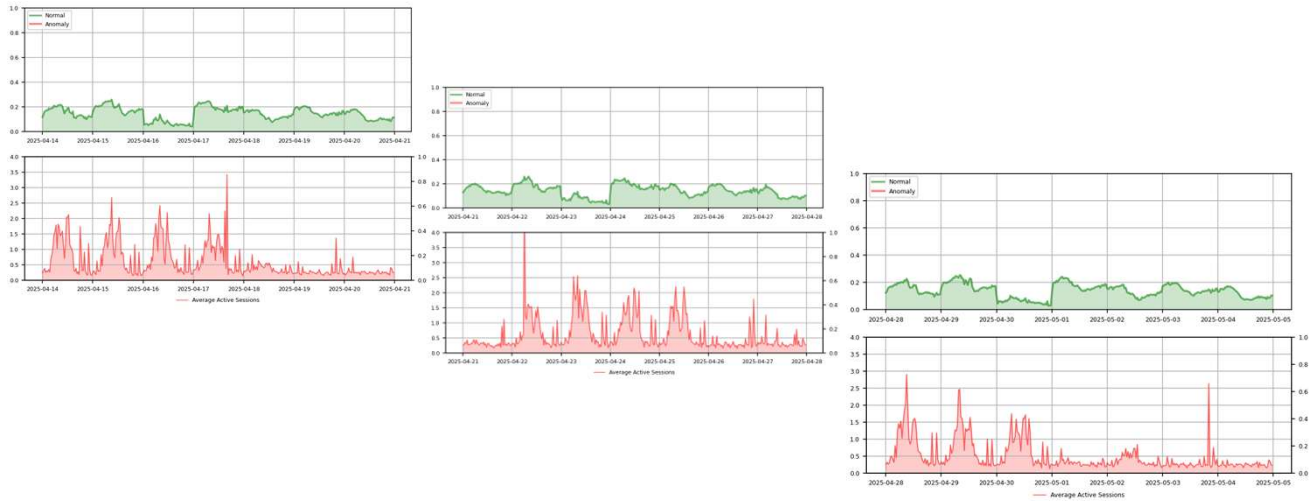
Timestamp	Metric	Time Waited	Wait Count
2025-05-10 10:33:15.334	log file sync	1.421.651	47.388.374

Zeitreihendaten sind Metriken, die in regelmäßigen Abständen erhoben und chronologisch abgespeichert werden. Im Kontext einer Oracle-Datenbank können das zum Beispiel Performance-Werte wie CPU-Auslastung, Speicherverbrauch, Anzahl von Sessions oder I/O-Operationen sein. Aber auch Wait-Events und bereits vorberechnete Metriken sind dabei hillfreich. Diese Daten ermöglichen es, Trends zu erkennen und potenzielle Engpässe frühzeitig zu identifizieren. In Oracle-Umgebungen werden relevante Messwerte häufig mithilfe des Automatic Workload Repository (AWR) oder vergleichbarer Tools gesammelt. Dabei entstehen umfangreiche historische Datensätze, die sich sowohl für die Fehlerdiagnose als auch für das proaktive Monitoring nutzen lassen.

Ein Vorteil der Zeitreihendatenspeicherung besteht darin, dass sich die Entwicklung von Kennzahlen über die Zeit hinweg analysieren lässt. Gerade bei kritischen Unternehmensanwendungen erlaubt eine präzise Verlaufskontrolle, Lastspitzen oder ungewöhnliche Schwankungen schnell zu erkennen. Gleichzeitig bildet dieses Vorgehen die Basis für weiterführende Analysen mit KI-Methoden, bei denen nicht nur einzelne Messwerte, sondern die Korrelationen zwischen verschiedenen Metriken in Betracht gezogen werden. Dadurch lässt sich ein genaueres Bild über den Zustand der Datenbank und ihrer Infrastruktur gewinnen, was in komplexen, verteilten Umgebungen besonders hilfreich ist.

Oracle AWR ist Teil des kostenpflichtigen Diagnostics Packs, was zu Mehrkosten führt, zudem werden die erfassten Daten nur für einen begrenzten Zeitraum aufbewahrt, wodurch langfristige Analysen erschwert werden. Das Sammeln und Speichern der Performance-Daten verursacht einen gewissen Ressourcenverbrauch und konzentriert sich weitgehend auf datenbankspezifische Metriken, was eine umfassende Integration externer Messwerte oder heterogener Umgebungen erschwert. Auch wenn die Daten aus AWR grundsätzlich exportierbar sind, gestaltet sich eine tiefe, automatisierte Einbindung in andere Monitoring-Lösungen meist aufwendig und bietet vor allem auch keine Echtzeit-Anbindung. Zudem fehlt eine ganzheitliche Perspektive, da AWR nur für Oracle-Umgebungen zur Verfügung steht und ohne entsprechendes Fachwissen die Interpretation von AWR-Reports schnell an Grenzen stoßen kann.

Grundlagen Zeitreihendaten am Beispiel Oracle



Warum universelle Anomalieerkennung statt Fokus auf einzelne Kennzahlen

- Entdeckung unbekannter Abweichungen ohne vordefinierte Grenzen
- Flexibilität bei neuen oder selten genutzten Metriken
- Unterstützung verschiedener Systeme und Technologien
- Mehrwert durch ganzheitliche Betrachtung aller relevanten Daten

Eine auf einzelne Metriken beschränkte Überwachung kann schnell zu Fehleinschätzungen führen, da sich Probleme oftmals aus dem Zusammenwirken mehrerer Faktoren ergeben. Durch die universelle Anomalieerkennung werden sämtliche verfügbaren Messwerte in die Analyse einbezogen, was eine flexible Anpassung an unterschiedliche Umgebungen und neu hinzukommende Datenquellen ermöglicht. Dieser ganzheitliche Ansatz erkennt nicht nur klar definierte Schwellwertüberschreitungen, sondern spürt auch Abweichungen auf, die sich erst durch die Kombination mehrerer Variablen zeigen. Damit können sowohl bislang unbekannte Problemmuster entdeckt als auch langfristige Entwicklungen besser im Auge behalten werden. Die universelle Anomalieerkennung trägt somit zu einer höheren Betriebssicherheit bei und macht den Einsatz komplexer, starrer Spezialregeln überflüssig, da das System automatisch lernt, wie „normales“ Verhalten für verschiedene Systeme und Zeiträume aussieht.

Externe Quellen (Auswahl):

- Gartner. (2021). *Magic Quadrant for AIOps Platforms*
- Forrester. (2022). *The Forrester Wave™: Artificial Intelligence for IT Operations*
- IEEE. (2021). *Transactions on Big Data: Anomaly Detection Techniques in Large-Scale Systems*

Deep-Learning-Ansatz im Monitoring

- Einsatz neuronaler Netze für Mustererkennung in Zeitreihen
- Vorteile: hohe Genauigkeit, Lernfähigkeit aus großen Datenmengen
- Beispiel-Architekturen: Autoencoder, LSTM, Transformer-basierte Modelle
- Skalierbarkeit in Cloud- und On-Premises-Umgebungen

Deep-Learning-Verfahren bieten die Möglichkeit, komplexe Muster in Zeitreihendaten zu erkennen und Abweichungen selbst dann zu identifizieren, wenn sie subtile Veränderungen betreffen. Anders als herkömmliche Machine-Learning-Methoden verwenden viele Deep-Learning-Modelle mehrschichtige neuronale Netze (etwa Autoencoder, LSTM oder Transformer-Modelle), die ein tieferes Verständnis für die Struktur der Daten entwickeln können. Diese Modelle sind in der Lage, eine Vielzahl von Eingaben zu verarbeiten und dabei sowohl kurzfristige als auch längerfristige Zusammenhänge zu erfassen.
Insbesondere im Anwendungsfall „Monitoring“ profitieren Deep-Learning-Verfahren von kontinuierlich gesammelten Messwerten, wie sie etwa in Datenbank- oder Infrastrukturumgebungen anfallen. Da die Modelle sich im Laufe der Zeit an verändernde Betriebszustände anpassen, wird das Monitoring zunehmend effektiver und kann gleichzeitig das Risiko von Fehlalarmen reduzieren. Eine gründliche Auswertung mehrerer korrelierter Metriken ermöglicht eine bessere Ursachenanalyse und schafft die Basis für eine engmaschige Fehlervorbeugung. Damit bildet Deep Learning eine zentrale Technologie, um in komplexen IT-Landschaften den Schritt vom rein reaktiven zum proaktiven Monitoring zu vollziehen.

Externe Quellen (Auswahl):

- Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep Learning*. MIT Press
- IEEE. (2021). *Transactions on Neural Networks and Learning Systems: Time-Series Anomaly Detection*
- Gartner. (2022). *Market Guide for AIOps Platforms*

GenAI/RAG für Ursachenanalyse

- Kombination von KI für Anomalieerkennung mit Large Language Models (LLMs)
- RAG (Retrieval-Augmented Generation) zur Einbindung domänenspezifischen Wissens
- Textbasierte Interaktion für tiefe Einblicke in Problemquellen
- Vorteil: Intuitives Dialogsystem statt rein technischer Fehlercodes

Die Kombination von Large Language Models (LLMs) mit Retrieval-Augmented Generation (RAG) bietet im Monitoring-Umfeld eine moderne Möglichkeit, komplexe Probleme schneller zu identifizieren und zu verstehen. Während LLMs auf umfangreichen Trainingsdatensätzen basieren und in der Lage sind, Zusammenhänge in natürlichsprachlichen Eingaben zu erkennen, erlaubt RAG den zielgerichteten Abruf zusätzlicher Informationen aus unternehmensinternen Wissensquellen oder externen Dokumentationen. Dadurch können Anwender mithilfe eines interaktiven Dialogs detaillierte Erklärungen und Handlungsempfehlungen zu möglichen Fehlerursachen erhalten, ohne hierfür einen Experten konsultieren zu müssen. Diese textbasierte Herangehensweise erleichtert die root-cause-Analyse deutlich, da nicht allein technische Kennzahlen ausgewertet, sondern auch kontextbezogene Faktoren in die Ursachenforschung einbezogen werden. Zugleich ermöglicht ein derartiges System eine kontinuierliche Erweiterung des Wissensbestands, indem neu erarbeitete Lösungen oder Best Practices dauerhaft hinterlegt werden.

Externe Quellen (Auswahl):

- OpenAI. (2020). *Language Models are Few-Shot Learners*
- *M*icrosoft Research. (2021). *Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks*
- Google. (2022). *LLM Applications in Enterprise Environments.*

Praktischer Ablauf einer KI-basierten Ursachenanalyse

- Anomalie erkennen (z.B. Performance-Abfall bei Oracle)
- Automatisierte Identifikation potenzieller Ursachen (Logs, Metriken, historische Daten)
- Interaktive Analyse via LLM: gezielte Fragen stellen und Expertenwissen abrufen
- Synthese der Ergebnisse und Handlungsempfehlungen

Bei einer KI-basierten Ursachenanalyse werden zunächst die anfallenden Messwerte kontinuierlich überwacht, um auffällige Abweichungen im Betriebsverhalten zu entdecken. Sobald eine potenzielle Anomalie auftritt, greifen KI-Methoden – zum Beispiel auf Basis von Deep-Learning-Modellen –, um das Problem einzugrenzen und mögliche Ursachen zu ermitteln. Dabei können sowohl historische Zeitreihendaten als auch Kontextinformationen, wie Konfigurationsänderungen oder Logdateien, herangezogen werden. In einem zweiten Schritt erfolgt durch textbasierte KI-Interaktion (etwa mit einem Large Language Model) eine weiterführende Analyse, bei der sich gezielt Fragen zu spezifischen Systembereichen stellen lassen. Dadurch wird eine tiefere Fehlerdiagnose ermöglicht, da die KI beispielsweise Hinweise zu selten auftretenden Fehlerquellen liefern oder Korrelationen zwischen mehreren Metriken aufzeigen kann. Im Ergebnis erhält das Monitoring-Team klare Anhaltspunkte für gezielte Gegenmaßnahmen, ohne sich ausschließlich auf manuelle Suche und Expertenwissen verlassen zu müssen.

Externe Quellen (Auswahl):

- Forrester. (2022). *The Forrester Wave™: Artificial Intelligence for IT Operations*
- IEEE. (2021). *Transactions on Neural Networks and Learning Systems: Time-Series Analysis for Anomaly Detection*
- MIT Sloan Management Review. (2020). *AI-Powered Decision Making in Complex Systems*

Proaktives Monitoring in der Praxis

- Kontinuierliche Überwachung aller relevanten Metriken
- Frühwarnsystem dank KI (Benachrichtigung vor kritischen Schwellen)
- Besseres Ressourcen- und Kapazitätsmanagement
- Vermeidung kostenintensiver Ausfälle durch rechtzeitige Gegenmaßnahmen

Proaktives Monitoring setzt nicht erst an, wenn ein Problem offensichtlich wird, sondern sucht bereits im Vorfeld nach ungewöhnlichen Mustern und möglichen Risikofaktoren. Statt sich auf manuelle Prüfungen oder vordefinierte Alarmregeln zu verlassen, erfolgt dabei eine kontinuierliche Analyse sämtlicher verfügbarer Messwerte. Mithilfe KI-gestützter Verfahren können neue und bislang unbekannte Anomalien erkannt werden, noch bevor sie spürbare Auswirkungen auf die Systemleistung haben. Auf diese Weise lassen sich zum Beispiel drohende Ressourcenengpässe, Sicherheitslücken oder Konfigurationsfehler frühzeitig identifizieren und beheben.
In der Praxis bedeutet das vor allem einen deutlich geringeren Zeitaufwand bei der Fehlersuche, da administrative Teams gezielt auf verdächtige Abweichungen hingewiesen werden. Darüber hinaus führt proaktives Monitoring zu mehr Stabilität und Planbarkeit im Tagesgeschäft, weil Ausfälle und Engpässe präventiv vermieden werden können. Im Idealfall werden relevante Informationen in Echtzeit auf einem Monitoring-Dashboard zusammengeführt, um Fachleuten jederzeit eine aktuelle Übersicht über den Zustand der gesamten IT-Landschaft zu ermöglichen.

Externe Quellen (Auswahl):

- Gartner. (2021). *Market Guide for IT Infrastructure Monitoring Tools*
- IDC. (2022). *Enterprise IT Infrastructure Strategies: Monitoring and Analytics*
- O'Reilly. (2020). *Practical Monitoring: Effective Strategies for the Real World*

Nutzen für DBAs und IT-Administratoren

- Reduktion des manuellen Aufwands bei der Fehlerdiagnose
- Bessere Ausnutzung der eigenen Expertise (Fokus auf Lösung statt Suche)
- Vereinfachte Zusammenarbeit zwischen Fachabteilungen und KI-Systemen
- Entlastung bei Routineaufgaben (z.B. Monitoring unzähliger Metriken)

Der Einsatz KI-gestützter Monitoring- und Analysetools bietet eine deutliche Entlastung für Datenbank- und IT-Administratoren. Durch automatisierte Fehlererkennung und gezielte Ursachenanalyse sinkt der manuelle Aufwand bei der Identifikation und Behebung von Problemen. Da das System dabei auf Mustererkennung und komplexe Abhängigkeiten in den Daten zugreift, können Engpässe frühzeitig erkannt werden, was Ausfallzeiten und damit verbundene Kosten reduziert. Gleichzeitig lassen sich wertvolle Ressourcen sparen, weil das Monitoring-Team weniger Zeit in reine Routineaufgaben investieren muss. Die Expertise der Fachleute wird somit gezielt dort eingesetzt, wo sie den größten Mehrwert bietet: bei der strategischen Optimierung der IT-Infrastruktur und der Koordination verschiedener Systeme und Datenbanken.
Ein weiterer Vorteil besteht in der verbesserten Verständlichkeit der Systemzustände, da KI-basierte Lösungen nicht nur Warnmeldungen ausgeben, sondern auch konkrete Hinweise auf mögliche Fehlerquellen liefern. Auf diese Weise lassen sich historische und aktuelle Daten in Echtzeit auswerten, um Trends und langfristige Entwicklungen besser einzuschätzen. Insgesamt erleichtern KI-gestützte Monitoring-Verfahren die Zusammenarbeit im IT-Team, da Daten und Ergebnisse zentral verfügbar sind und sich schnell mit anderen Abteilungen oder externen Dienstleistern teilen lassen.

Unterstützung unerfahrener Nutzer durch LLMs

- KI kann Fachwissen in verständlicher Form bereitstellen
- Schritt-für-Schritt-Erklärungen bei der Ursachenfindung
- Erhöhung der Problemlösungskompetenz ohne jahrelange Spezialausbildung
- Beispiele: “Was bedeutet ein hoher Wait-Event-Wert?” – direkte Antwort durch das LLM

Wenn großen Sprachmodellen (LLMs) die Rolle eines „digitalen Assistenten“ zukommt, wird der Zugang zu Fachwissen in vielen Bereichen deutlich vereinfacht. Auch Personen mit geringerer Erfahrung in komplexen IT-Systemen wie Oracle-Datenbanken oder anderen Infrastrukturkomponenten können durch direkte Fragen an das Modell rasch fundierte Informationen abrufen. Dabei können Fragen etwa nach der Bedeutung bestimmter Metriken oder Fehlermeldungen gestellt werden, ohne dass das Team ständig einen Fachexperten hinzuziehen muss. Die Vorteile liegen zum einen in der Zeitersparnis und zum anderen in der Tatsache, dass das bereits vorhandene Know-how in verschriftlichter Form (etwa in Dokumentationen, Knowledge Bases oder Handbüchern) über Retrieval-Augmented Generation (RAG) eingebunden werden kann.
Gerade in dynamischen Umgebungen, in denen die Lernkurve hoch sein kann, bieten LLMs eine Art „intuitives Interface“: Neue Mitarbeiter oder Administratoren, die nur gelegentlich mit bestimmten Systemteilen arbeiten, sind in der Lage, sich selbstständig Lösungsansätze oder Erklärungen zu erarbeiten. Dadurch wird das gesamte Team entlastet, weil weniger Zeit in wiederkehrende Schulungs- und Supportanfragen investiert werden muss. Gleichzeitig wächst mit der Interaktion mit einem textbasierten Assistenten nach und nach das Verständnis für die IT-Umgebung, was langfristig zu einer höheren Selbstständigkeit und einer besseren Problemlösungskompetenz führt.

Externe Quellen (Auswahl):

- OpenAI. (2020). *Language Models are Few-Shot Learners*
- Microsoft Research. (2021). *Enhancing User Experience with Conversational AI*
- Google. (2022). *Empowering IT Teams with Large Language Models*

Vorteile und mögliche Bedenken

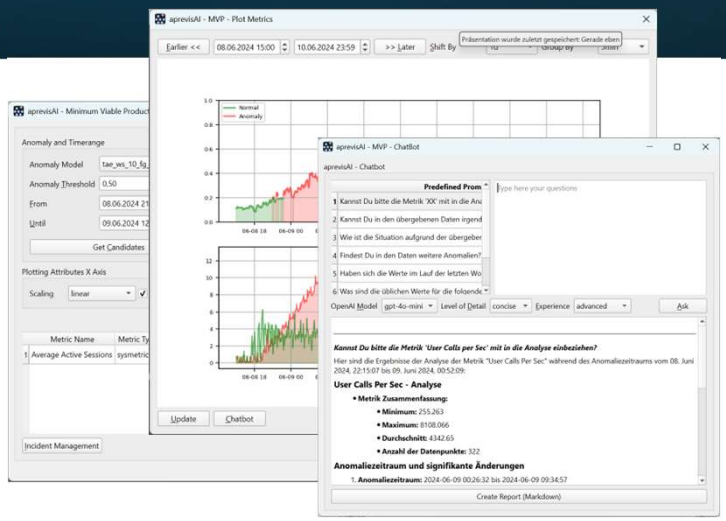
- Vorteile: Zeitersparnis, tiefergehende Analysen, Entdeckung unbekannter Risiken
- Bedenken:
  - Fehlalarm-Rate
  - Datenschutz bei sensiblen Daten
  - Vertrauenswürdigkeit der Modelle
- Bedeutung von Transparenz und Erklärbarkeit in KI-Systemen

Der Einsatz KI-basierter Monitoring-Lösungen ermöglicht eine schnellere und tiefgreifendere Analyse von Systemzuständen, was sowohl die Reaktionszeit bei Problemen als auch das Risiko unerkannter Ausfälle reduziert. Gleichzeitig fördert der Fokus auf Datenkorrelationen und Mustererkennung die Identifizierung bisher unbekannter Risiken oder Engpässe. Dennoch gibt es auch potenzielle Bedenken: Zu viele Fehlalarme können die Akzeptanz der Lösung senken, gerade wenn nicht klar ist, warum eine Anomalie gemeldet wurde. Zudem stellt sich bei sensiblen Daten die Frage nach Datenschutz und Zugriffsrechten, da KI-Modelle in manchen Fällen Einblicke in kritische Log- oder Metadaten benötigen. Ein weiterer Aspekt ist die Vertrauenswürdigkeit der Ergebnisse: Wenn nicht ausreichend dokumentiert ist, wie die KI zu einem bestimmten Schluss kommt, kann das im Team zu Skepsis führen. Aus diesem Grund sind transparente Prozesse und nachvollziehbare KI-Modelle (Explainable AI) ein wesentlicher Faktor, um das volle Potenzial dieser Lösungen auszuschöpfen.

Externe Quellen (Auswahl):

- Gartner. (2021). *Magic Quadrant for AIOps Platforms*
- McKinsey & Company. (2022). *The AI Advantage: Leveraging Data and Analytics*
- IEEE. (2020). *Transactions on Artificial Intelligence: Balancing Efficiency and Explainability in AI Systems*

Unsere Idee - aprevisAI



- KI-gestützte Erkennung von Anomalien
- Ursachenfindung mit Hilfe von GenAI/RAG im Dialog
- Einbindung in existierende Monitoring-Lösungen
- Gefördert von HESSENIDEEN

Ausblick: Zukunft des IT-Monitorings

- Automatisierte Fehlerbehebung (Self-Healing-Systeme)
- Enge Verzahnung von KI-gestützter Anomalieerkennung und Cloud-Management
- Forecasting
- Ausbau domänenspezifischen Wissens für komplexere Analyse (RAG-Ansätze)
- Starke Integration von Security- und Compliance-Aspekten

Die Weiterentwicklung des Monitorings bewegt sich klar in Richtung eines proaktiven, stark automatisierten Ansatzes, bei dem KI-Modelle nicht nur Anomalien erkennen, sondern auch selbstständig Gegenmaßnahmen ergreifen. In Self-Healing-Umgebungen können Systeme etwa Container neu starten, Ressourcen dynamisch skalieren oder Konfigurationen anpassen, ohne dass manuell eingegriffen werden muss. Parallel verschmelzen Monitoring- und Security-Disziplinen immer stärker: KI-Gestützte Verfahren korrelieren Performance-Metriken mit sicherheitsrelevanten Ereignissen, um Angriffe oder Fehlkonfigurationen frühzeitig aufzudecken.
Ein weiterer Trend ist die tiefe Integration von Large Language Models in Observability-Plattformen. LLMs stellen kontextbezogene Erklärungen, Handlungsempfehlungen und Best-Practices unmittelbar bereit – sowohl für On-Premises- als auch für Multi-Cloud- und Edge-Umgebungen. Damit steigt die Transparenz, gleichzeitig sinkt der Schulungsaufwand für neue Werkzeuge. Ergänzt wird dies durch verstärkte Bemühungen um Explainable AI, um Entscheidungswege der Modelle nachvollziehbar zu machen und regulatorische Anforderungen an Datenschutz und Compliance zu erfüllen. Perspektivisch wird Monitoring damit zu einer intelligenten, durchgängigen Steuerungsinstanz, die Leistung, Sicherheit und Kosten in Echtzeit optimiert – ein wesentlicher Baustein für resiliente, zukunftsfähige IT-Landschaften.

Externe Quellen (Auswahl):

- Gartner. (2022). *Market Guide for AIOps Platforms*.
- Forrester. (2022). *The Future of Observability and AIOps*.
- Microsoft. (2021). *Advancing Intelligent Cloud: AI-driven Monitoring and Security*.

Zusammenfassung & Takeaways

- KI-basierte Anomalieerkennung: ein wichtiger Schritt zu proaktivem Monitoring
- Deep-Learning- und LLM-Ansätze ergänzen sich ideal für Ursachenanalyse
- Erfahrene wie unerfahrene Nutzer profitieren vom textbasierten Dialog mit KI
- Nachhaltige Lösung, um in komplexen IT-Umgebungen schnell und effektiv zu reagieren

Das Monitoring von IT-Umgebungen entwickelt sich zunehmend von einer rein reaktiven Fehlererkennung hin zu einem proaktiven und teils sogar selbstkorrigierenden System. Möglich wird dies durch fortlaufende Verbesserungen in den Bereichen Künstliche Intelligenz und Machine Learning, die es erlauben, ungewöhnliche Zustände noch präziser und frühzeitiger zu erkennen. Künftige Lösungen werden verstärkt auf Automatisierung setzen, etwa indem sie bei bestimmten Abweichungen nicht nur Alarmmeldungen erzeugen, sondern automatisch geeignete Gegenmaßnahmen einleiten oder Ressourcen in Echtzeit an veränderte Lastprofile anpassen.
Ein weiterer Trend ist die tiefere Verzahnung von Monitoring und Security, da die Erkennung von Anomalien häufig auch sicherheitsrelevante Ereignisse enthüllt. Gleichzeitig dürften Large Language Models immer leichter in bestehende Plattformen integriert werden, um einfache Erklärungen oder fundierte Ursachenanalysen sowohl Anfängern als auch erfahrenen Administratoren zugänglich zu machen. Gerade in hybriden oder Multi-Cloud-Umgebungen wird eine enge Abstimmung zwischen KI-Modellen und Infrastrukturkomponenten wichtig, damit nicht nur Datenbanken, sondern auch Applikationsserver, Netzwerksegmente und Container-Plattformen verlässlich überwacht werden können. Diese umfassende Perspektive wird maßgeblich dazu beitragen, Ausfallzeiten zu minimieren und kritische Probleme frühzeitig zu lösen.

Externe Quellen (Auswahl):

- Gartner. (2022). *Market Guide for AIOps Platforms*
- Forrester. (2022). *The Future of Observability and AIOps*
- Microsoft. (2021). *Advancing Intelligent Cloud: AI-driven Monitoring and Security*

Demo + Q&A

aprevis
EXPLAINABLE AI

Q&A & Diskussion

- Fragen zur technischen Umsetzung
- Austausch über individuelle Anforderungen
- Möglichkeiten zur Weiterentwicklung und Integration in bestehende Umgebungen

Mögliche Fragen für die Q&A-Folie

1. Wie hoch ist der zusätzliche Ressourcenbedarf für KI-gestützte Anomalieerkennung im Vergleich zu herkömmlichen Monitoring-Lösungen?
2. Welche Kriterien muss man bei der Auswahl eines KI- oder AIOps-Tools berücksichtigen (z. B. Integrationsfähigkeit, Lizenzkosten, Support)?
3. Wie lassen sich Datenschutz und Sicherheit gewährleisten, wenn KI-Systeme Zugriff auf sensible Log- oder Performance-Daten erhalten?
4. Welche Erfolgsfaktoren haben sich bei der Einführung einer KI-basierten Lösung in der Praxis bewährt (z. B. Pilotprojekte, enge Einbindung des Teams)?
5. Wie hoch ist das Risiko von Fehlalarmen und wie können solche False Positives reduziert werden?
6. Kann die KI auch selbstständig Gegenmaßnahmen einleiten oder benötigt man immer noch manuelles Eingreifen?
7. Wie lässt sich die Lösung in bestehende Monitoring-Tools wie Prometheus, Grafana oder andere Systemlandschaften integrieren?
8. Was ist zu beachten, wenn man KI-Verfahren für sehr unterschiedliche Systeme einsetzen möchte (Datenbanken, Container, Cloud-Services)?
9. Wie verfährt man mit Oracle-spezifischen Eigenheiten (z. B. AWR), um eine umfassende Ursachenanalyse zu gewährleisten?
10. Welche Rolle spielt das Know-how der Administratoren bei KI-gestützten Verfahren – können auch weniger erfahrene Mitarbeiter von Anfang an profitieren?