

Real-time Anomaly Detection in Customer Intelligence Plattform

Die vorliegende Masterarbeit befasst sich mit dem Aufbau eines Systems zur Verarbeitung von Echtzeit-Daten. Im Fokus steht das Bereitstellen von Computing-Ressourcen, die Skalierbarkeit als auch die Verfügbarkeit des Systems. Anwender/-innen können mit dessen Hilfe eigenständig komplexe Daten-Pipelines aufbauen, ohne sich mit der darunterliegenden Server-Infrastruktur befassen zu müssen. Dazu wird ein Konzept basierend auf bestehenden Open-Source-Komponenten ausgearbeitet und implementiert.

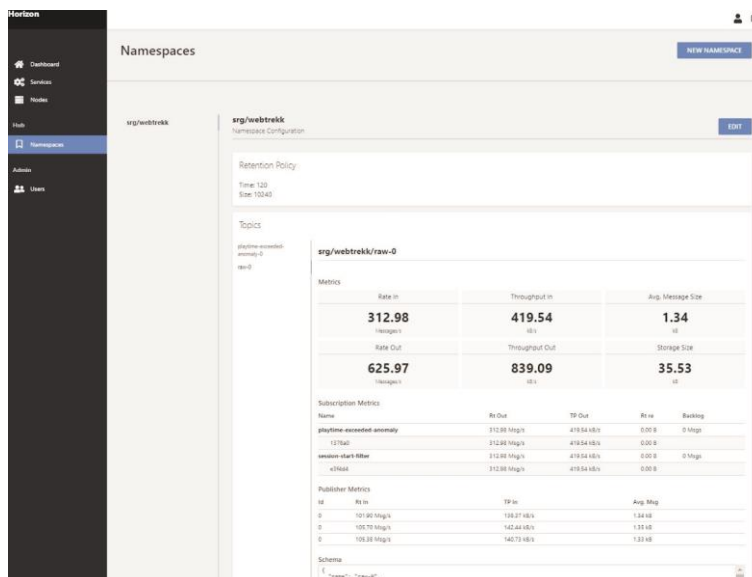
In Zusammenarbeit mit der Schweizerischen Radio- und Fernsehgesellschaft (SRG) werden am Beispiel einer Customer Intelligence Plattform zwei Pipelines als Proof-of-Concept umgesetzt. Marketingdaten aus den Onlineangeboten der SRG werden laufend analysiert und auf Anomalien überprüft. Anomalien sind Datensätze, die in beliebiger Weise von den erwarteten Werten abweichen und damit die Datengrundlage verfälschen. Die Ursachen von Anomalien gilt es daher schnellstmöglich zu beheben.

Neben der Detektion von Anomalien kann das System für andere Anwendungen interessant sein. Insbesondere für Anwendungen, bei denen die zeitnahe Verarbeitung von gesammelten Daten wichtig ist. Das System befindet sich aktuell bei der SRG im produktiven Einsatz. Wie und ob sich das System im Alltag bewährt, zeigt die Zukunft.



Diplomand/in
Roman Fürst

Dozent
Jürgen Spielberger



Die Benutzeroberfläche zur Verwaltung von Data-Pipelines zeigt den aktuellen Durchfluss der einzelnen Datenpakete.