



## OPTIMIERUNG DER DATENVERARBEITUNG MIT APACHE NIFI

### Effiziente LEI-Implementierung durch moderne Technologien

Die internationale Organisation Financial Stability Board, die das globale Finanzsystem überwacht und Empfehlungen ausspricht, hat eine Organisation ins Leben gerufen, um die Implementierung und Nutzung des Legal Entity Identifier (LEI) zu fördern. Ein LEI ist eine global eindeutige Kennung für Rechtsträger im Finanzmarkt, die 2012 eingeführt wurde.

Diese neue Organisation wird von internationalen Behördenvertretern unterstützt. Ihr Ziel ist es, die Transparenz auf den globalen Finanzmärkten zu verbessern.

#### AUF EINEN BLICK:

- Branche: gemeinnützige Organisation
- Ziel: Leistungsoptimierung der Datenverarbeitung
- Technologie: Apache NiFi

#### ÜBER DEN KUNDEN:

Diese Organisation wird von internationalen Behördenvertretern unterstützt. Ihr Ziel ist es, die Transparenz auf den globalen Finanzmärkten zu verbessern.

#### HERAUSFORDERUNG:

Die Organisation bietet eine einzigartige Online-Quelle für offene, standardisierte und hochwertige Referenzdaten über Rechtsträger. Diese Daten ermöglichen es Menschen und Unternehmen, fundierte, wirtschaftlich sinnvolle und verlässliche Entscheidungen bei der Auswahl ihrer Geschäftspartner zu treffen.

Um die Leistung ihrer Datenverarbeitung zu optimieren und die Daten schneller für die Suchfunktion auf ihrer Webseite bereitzustellen, wurde die SHI GmbH beauftragt. Die bestehende Datenverarbeitungs-komponente sollte durch das Open-Source-Tool Apache NiFi ersetzt werden. Im Rahmen eines Proof of Concept (PoC) wurde der Prozess der Datenbereitstellung mit Apache NiFi modelliert.

#### LÖSUNG UND UMSETZUNG:

Der erste Schritt bestand darin, den aktuellen Indexierungsprozess durch einen Workflow in Apache NiFi zu ersetzen. Alle für die Datenbereitstellung wesentlichen Schritte wurden berücksichtigt:

**Datenquellen:** Apache NiFi kann verschiedene Datenquellen anbinden. Je nach Komplexität und Heterogenität der Quellen wurden im PoC einige repräsentative Beispiele ausgewählt und in den Prozess integriert. Der Workflow ist flexibel und kann später um weitere Datenquellen erweitert werden.



**Verarbeitungsschritte:** Die Funktionsweise des derzeit genutzten Indexers wurde analysiert, um relevante Verarbeitungsschritte nachvollziehen zu können. Passende Prozessoren für den Workflow in NiFi wurden ausgewählt, konfiguriert und kombiniert, wobei besonderer Wert auf performante Datenverarbeitung gelegt wurde.

**Ausgabeformat:** Die verarbeiteten Daten wurden für die auf Apache Solr basierende Suchfunktion auf der Webseite bereitgestellt. Die Daten wurden in kompatible Formate gebracht und an Apache Solr weitergeleitet.

### ERGEBNIS:

Vor der Optimierung dauerte die Indexierung in Solr sehr lange, da eine Java-Software zur Datenindexierung eingesetzt wurde. Der Proof of Concept mit minimal konfiguriertem NiFi reduzierte die Indexierungszeit um beeindruckende 54,6 Prozent – mehr als die Hälfte der ursprünglichen Zeit. SHI hat gezeigt, wo Bottlenecks bestehen und wie durch verschiedene Solr-Konfigurationen und Setups zusätzliche Performancegewinne erzielt werden können.

Dank der Zusammenarbeit mit der SHI GmbH und dem Einsatz von Apache NiFi konnte die Organisation die Performance im POC erheblich verbessern. Mit unserem Expertenwissen konnten wir aufzeigen, wie die Datenbereitstellung gestärkt und die Verarbeitung effizienter gestaltet werden kann.

### **Raphael Hoffmann, bei der SHI verantwortlich für Search- und Analytics-Lösungen:**

„Durch die Implementierung von Apache NiFi konnten wir die Effizienz der Datenverarbeitung erheblich steigern und die Indexierungszeit um über 50% reduzieren. Dies zeigt, wie gezielte technologische Lösungen zu messbaren Verbesserungen führen können.“

