



integrata
cegos

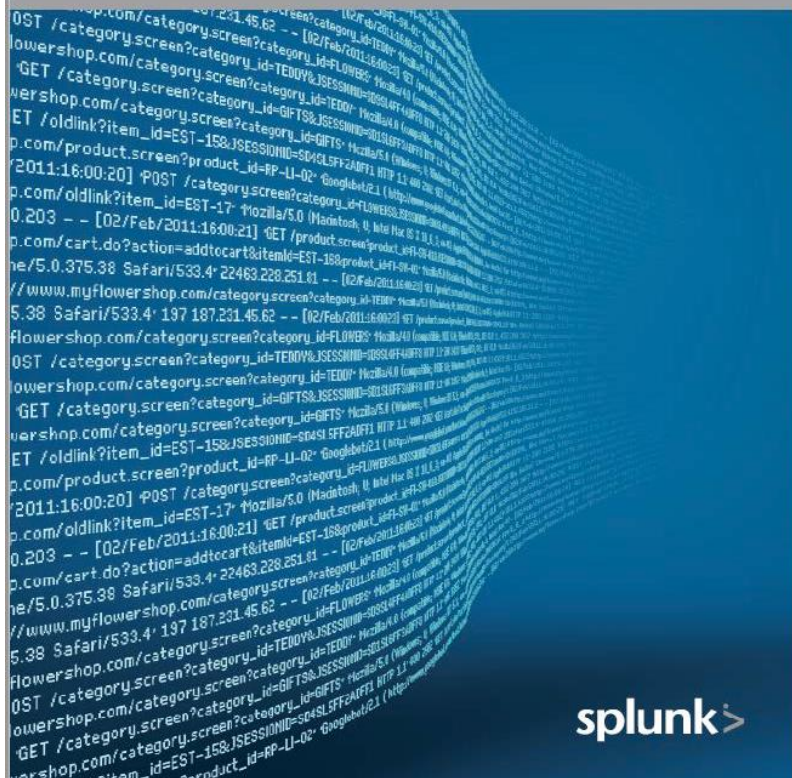
Splunk

Anwender

Exploring Splunk

SEARCH PROCESSING LANGUAGE (SPL)
PRIMER AND COOKBOOK

By David Carasso, Splunk's Chief Mind



splunk>

IT

SECURITY

DEVOPS

PLATFORM

WHY SPLUNK?

EXPLORE ▾

The Data-to-Everything™ Platform

Drive outcomes across Security, IT and
DevOps with the data platform built for
the cloud.

Watch the Video ▶

- Die in diesem Seminar verwendete Werkzeuge und Frameworks sind Open Source
 - LPGL Lizenzmodell
- Dies ist ein Seminar mit praktischen Übungen
 - Damit werden die Inhalte durch Übungen vertieft und verinnerlicht
- Dokumentation und Ressourcen stehen auch im Internet zur Verfügung
- Konventionen
 - Befehle werden in `Courier-Schriftart` dargestellt
 - Dateinamen werden in *kursiver Courier-Schriftart* dargestellt
 - Links werden in unterstrichener Courier-Schriftart dargestellt

© Javacream

Javacream

Dr. Rainer Sawitzki

Alois-Gilg-Weg 6

81373 München

Alle Rechte, einschließlich derjenigen des auszugsweisen Abdrucks, der fotomechanischen und elektronischen Wiedergabe vorbehalten.

Einführung	6
Erstes Arbeiten mit Splunk	27
Suchen	39
Weitere Features	50

1

EINFÜHRUNG

1.1

WAS MACHT SPLUNK EIGENTLICH?

- Splunk kennt Informationen aus verschiedenen Quellen
- Diese Informationen werden von Splunk aufbereitet und mit zusätzlichen Attributen angereichert
- Darauf aufbauend können weitergehende Analysen definiert werden
- Die Ergebnisse werden in repräsentativen Diagrammen und Reports dargestellt
- Auffällige Informationen werden erkannt und hervorgehoben

Aha...



- Prinzipiell ist dies nichts anderes als eine Form der Datenanalyse und -aufbereitung
- So etwas kann bereits mit einem einfachen Tabellenkalkulationsprogramm durchgeführt werden!

zensus_schleswig_holstein.csv - Editor

Datei	Bearbeiten	Format	Ansicht	Hilfe	
Flensburg, Stadt;	82	;	41	;	42
Kiel, Landeshauptstadt;	236	;	114	;	122
Lübeck, Hansestadt;	210	;	100	;	110
Neumünster, Stadt;	77	;	38	;	40
Dithmarschen;	134	;	66	;	68
Herzogtum Lauenburg;	187	;	91	;	96
Nordfriesland;	164	;	80	;	84
Ostholstein;	198	;	95	;	103
Pinneberg;	296	;	144	;	152
Plön;	128	;	62	;	66
Rendsburg-Eckernförde;	269	;	132	;	137
Schleswig-Flensburg;	196	;	97	;	99
Segeberg;	261	;	128	;	133
Steinburg;	131	;	64	;	67
Stormarn;	231	;	112	;	119

Die Daten werden in Excel geladen

A1 ✕ ✓ fx Flensburg, Stadt					
	A	B	C	D	
1	Flensburg, Stadt	82	41	42	
2	Kiel, Landeshauptstadt	236	114	122	
3	Lübeck, Hansestadt	210	100	110	
4	Neumünster, Stadt	77	38	40	
5	Dithmarschen	134	66	68	
6	Herzogtum Lauenburg	187	91	96	
7	Nordfriesland	164	80	84	
8	Ostholstein	198	95	103	
9	Pinneberg	296	144	152	
10	Plön	128	62	66	
11	Rendsburg-Eckernförde	269	132	137	
12	Schleswig-Flensburg	196	97	99	
13	Segeberg	261	128	133	
14	Steinburg	131	64	67	
15	Stormarn	231	112	119	
16					

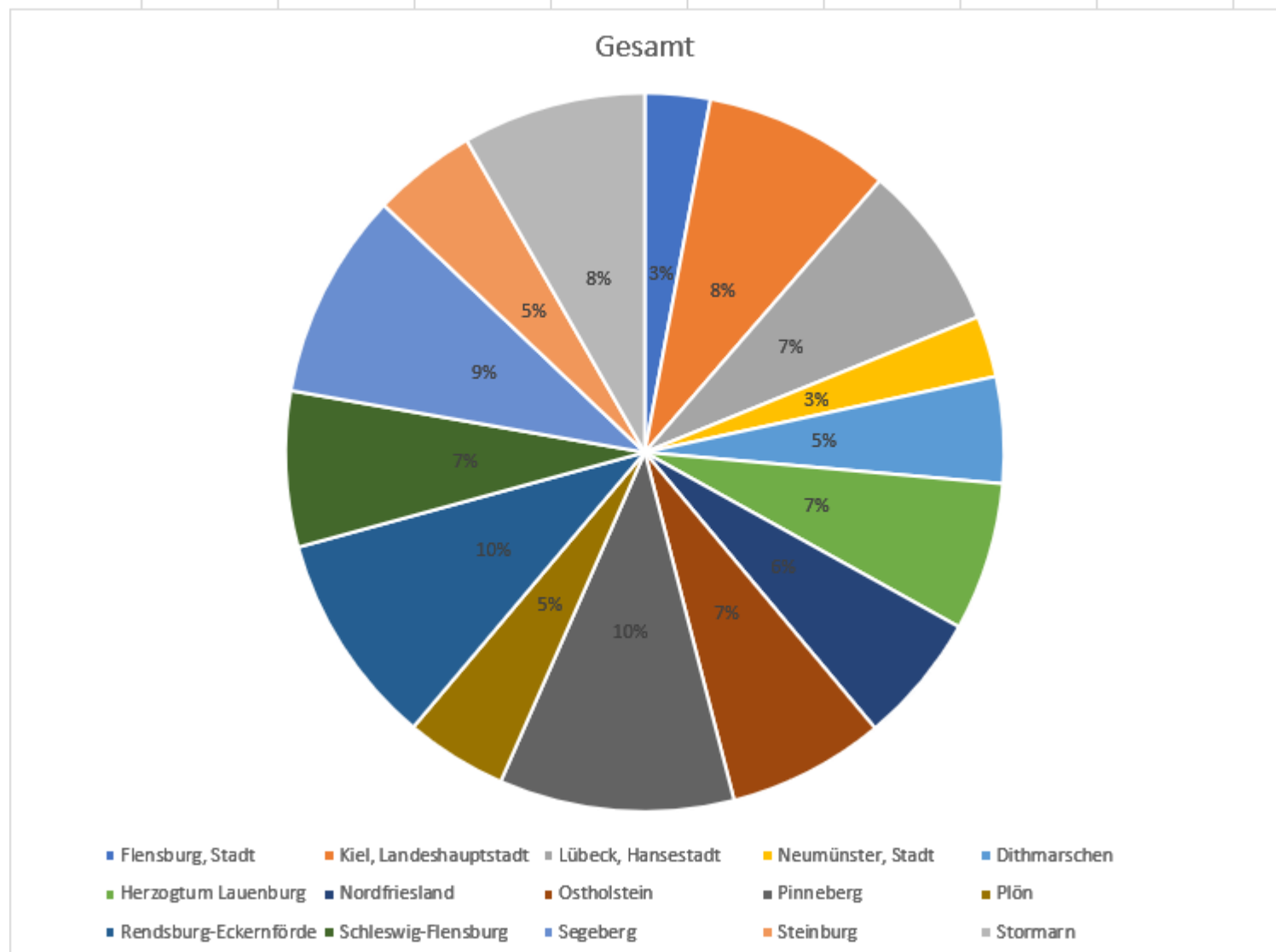
und mit Überschriften versehen

	A	B	C	D
1	Landkreis	Gesamt	Männlich	Weiblich
2	Flensburg, Stadt	82	41	42
3	Kiel, Landeshauptstadt	236	114	122
4	Lübeck, Hansestadt	210	100	110
5	Neumünster, Stadt	77	38	40
6	Dithmarschen	134	66	68
7	Herzogtum Lauenburg	187	91	96
8	Nordfriesland	164	80	84
9	Ostholstein	198	95	103
10	Pinneberg	296	144	152
11	Plön	128	62	66
12	Rendsburg-Eckernförde	269	132	137
13	Schleswig-Flensburg	196	97	99
14	Segeberg	261	128	133
15	Steinburg	131	64	67
16	Stormarn	231	112	119

Zur Analyse werden Filter aktiviert

A2 ✕ ✓ f_x Flensburg, Stadt

	A	B	C	D
1	Landkreis	Gesamt	Männlich	Weiblich
2	Flensburg, Stadt	<div>Nach Größe sortieren (aufsteigend) Nach Größe sortieren (absteigend) Nach Farbe sortieren Tabellenansicht Filter löschen aus "Weiblich" Nach Farbe filtern Zahlenfilter Suchen <input checked="" type="checkbox"/> (Alles auswählen) <input checked="" type="checkbox"/> 40 <input checked="" type="checkbox"/> 42 <input checked="" type="checkbox"/> 66 <input checked="" type="checkbox"/> 67 <input checked="" type="checkbox"/> 68 <input checked="" type="checkbox"/> 84 <input checked="" type="checkbox"/> 96 <input checked="" type="checkbox"/> 99 OK Abbrechen</div>		
3	Kiel, Landeshauptstadt			
4	Lübeck, Hansestadt			
5	Neumünster, Stadt			
6	Dithmarschen			
7	Herzogtum Lauenburg			
8	Nordfriesland			
9	Ostholstein			
10	Pinneberg			
11	Plön			
12	Rendsburg-Eckernförde			
13	Schleswig-Flensburg			
14	Segeberg			
15	Steinburg			
16	Stormarn			
17				
18				
19				
20				
21				
22				



Bedingte Formatierung zur Hervorhebung von Auffälligkeiten

A17				
	A	B	C	D
1	Landkreis	Gesamt	Männlich	Weiblich
2	Flensburg, Stadt	82	41	42
3	Kiel, Landeshauptstadt	236	114	122
4	Lübeck, Hansestadt	210	100	110
5	Neumünster, Stadt	77	38	40
6	Dithmarschen	134	66	68
7	Herzogtum Lauenburg	187	91	96
8	Nordfriesland	164	80	84
9	Ostholstein	198	95	103
10	Pinneberg	296	144	152
11	Plön	128	62	66
12	Rendsburg-Eckernförde	269	132	137
13	Schleswig-Flensburg	196	97	99
14	Segeberg	261	128	133
15	Steinburg	131	64	67
16	Stormarn	231	112	119

- Splunk verarbeitet Daten aus **verschiedensten Quellen**
 - Dateien im Dateisystem, aber auch
 - Datenströme über Netzwerk, Datenbanken...
- Daten können **live** eingespielt werden
- Die Daten müssen als Textdateien gelesen werden, die ansonsten in **beliebigem Format** vorliegen
- Die Menge von Daten, die Splunk verarbeiten kann, ist **beliebig groß**
- Die Analyse der Daten ist einesteils **intuitiv**, andererseits werden auch Verfahren angeboten, die dem Bereich „Künstliche Intelligenz“ bzw. genauer: „**Machine Learning**“ zuzuordnen sind
- Splunk kann **aktive Benachrichtigungen** erzeugen, die ein proaktives Handeln ermöglichen

1.2

ERSTES ARBEITEN MIT SPLUNK

- „Wir betreiben einen Web Shop“
- Problemstellungen
 - Technisch
 - Schnelle Lokalisierung von Fehlern im System
 - Erkennen von auffälligen Betriebssituationen, die ein administratives Eingreifen erfordern
 - Business
 - Was sind Top-Produkte?
 - Wie navigiert ein Kunde innerhalb des Shops?
 - Gibt es Auffälligkeiten im Nutzerverhalten?
 - Welche Suchanfragen führten zu keinen Ergebnissen?
 - Wie viele potenzielle Kaufvorgänge werden an welchen Stellen abgebrochen?

- Schritt 1
 - In welchen Quellen könnten relevante Informationen stecken?
- Schritt 2
 - Welche Teile der Informationen sind relevant, um die Probleme lösen zu können?
- Schritt 3
 - Wie können die Informationen dargestellt werden, um Lösungen schnell und intuitiv zu erfassen?

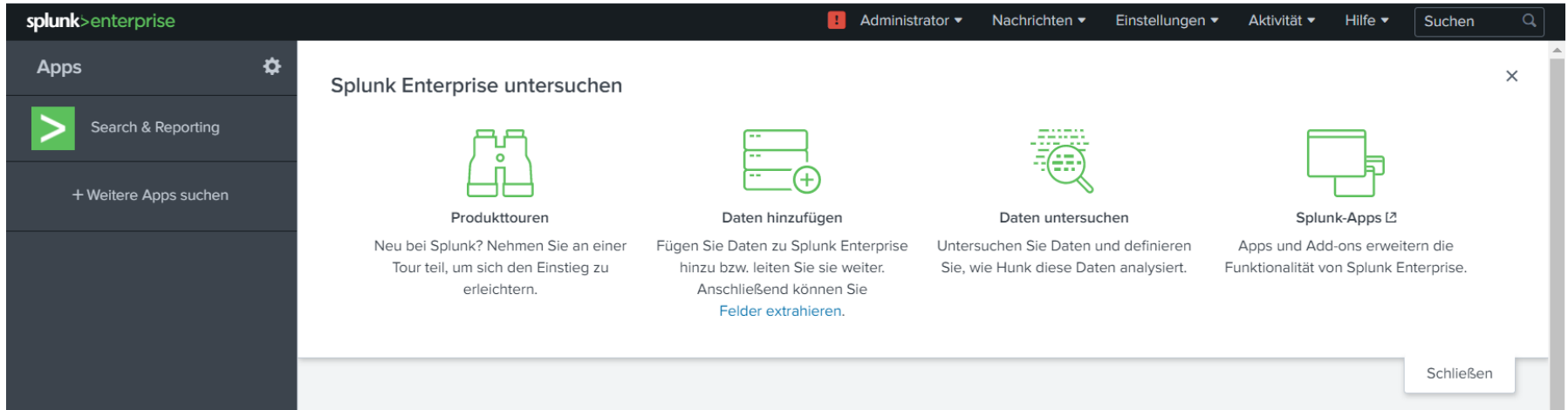
- Schritt 1
 - In welchen Quellen könnten relevante Informationen stecken?
 - „Der Web Shop besteht aus drei Web Servern und einem Mail-Server, die sämtliche Zugriffe in ihre jeweilige Log-Datei schreiben. Weiterhin protokolliert die Shop-Anwendung sämtliche Verkaufsaktivitäten.“
- Schritt 2
 - Welche Teile der Informationen sind relevant, um die Probleme lösen zu können?
 - „Fehlermeldungen enthalten den Text *Error* oder *Warning*, Zugriffe eines Kunden sind über die *IPAddress* verfügbar. Alle aufgerufenen Seiten des Shops werden zusammen mit der Uhrzeit und der Kundeninformation gespeichert.“
- Schritt 3
 - Wie können die Informationen dargestellt werden, um Lösungen schnell und intuitiv zu erfassen?
 - „Ein täglich erstellter Report zeigt alle Seitenaufrufe sortiert nach Zugriffszahlen und ermöglicht eine Beurteilung der Kundenaktivitäten. Tauchen in den Server-Logs Fehlermeldungen auf soll sofort ein Administrator benachrichtigt werden.“

- Lassen Sie sich nicht von der Vielzahl der eben angesprochenen technischen Begriffe demotivieren!
 - Sie müssen **kein Systemadministrator** sein, um sinnvoll mit Splunk arbeiten zu können
- Kümmern Sie sich auch nicht um technische Details wie beispielsweise „wie muss ich Splunk konfigurieren, damit die Log-Dateien in Echtzeit übertragen werden?“
 - Für diese Aufgabe gibt es **Spezialisten**, die Splunk in die System-Landschaft integrieren
 - Sie werden in den seltensten Fällen als Splunk-Anwender solche Aufgaben lösen müssen

1.3

PRAXIS

Eine Test-Splunk-Instanz



The screenshot shows the Splunk Enterprise web interface. The top navigation bar includes the 'splunk>enterprise' logo, a search bar, and links for 'Administrator', 'Nachrichten', 'Einstellungen', 'Aktivität', and 'Hilfe'. The left sidebar shows the 'Apps' menu with 'Search & Reporting' selected and a link to '+ Weitere Apps suchen'. The main content area is titled 'Splunk Enterprise untersuchen' and contains four cards:

- Produkttouren**: Neu bei Splunk? Nehmen Sie an einer Tour teil, um sich den Einstieg zu erleichtern.
- Daten hinzufügen**: Fügen Sie Daten zu Splunk Enterprise hinzu bzw. leiten Sie sie weiter. Anschließend können Sie [Felder extrahieren](#).
- Daten untersuchen**: Untersuchen Sie Daten und definieren Sie, wie Hunk diese Daten analysiert.
- Splunk-Apps**: Apps und Add-ons erweitern die Funktionalität von Splunk Enterprise.

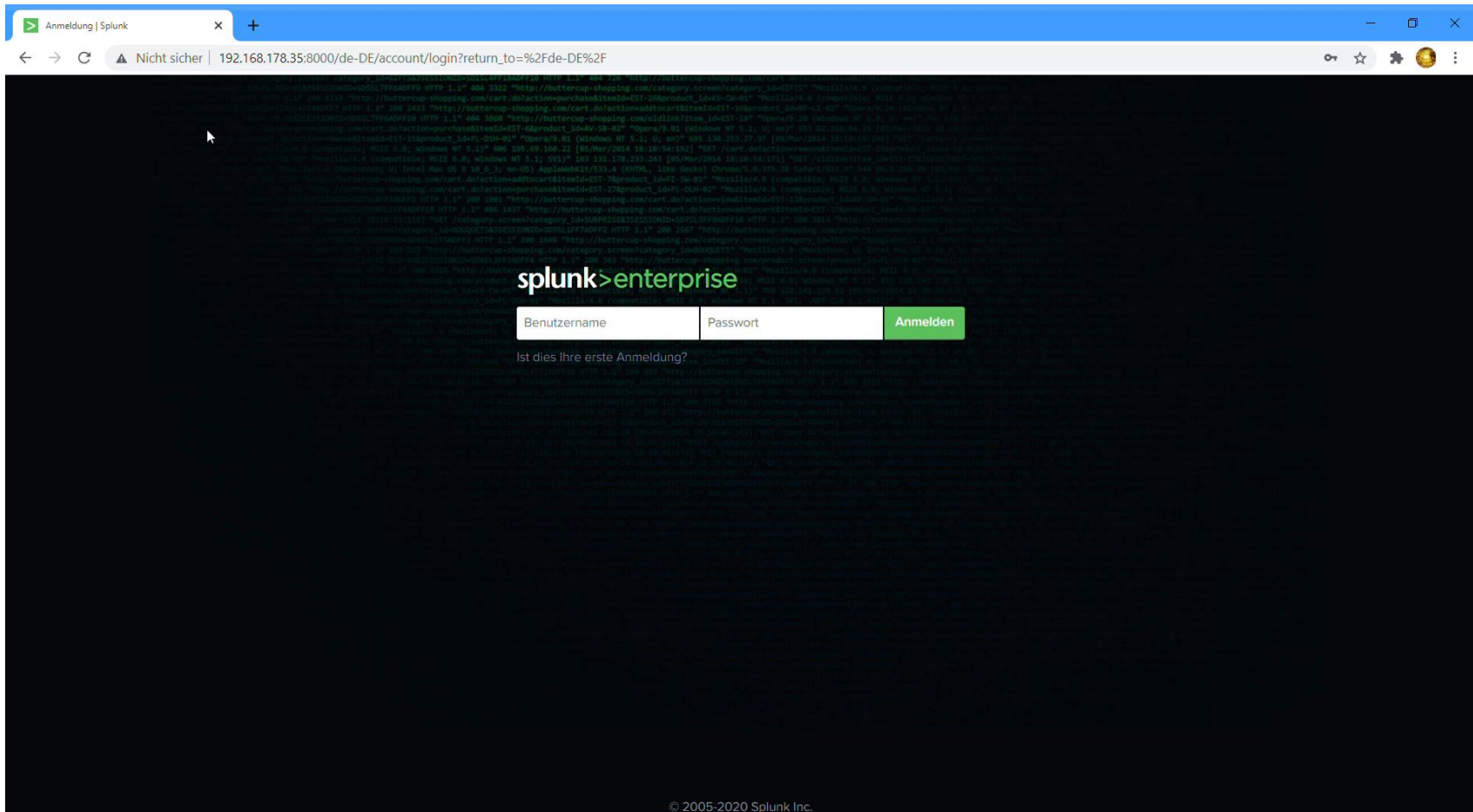
A 'Schließen' (Close) button is located in the bottom right corner of the main content area.

Die Test-Daten als ZIP-Archiv

C:\Users\Rainer Sawitzki\Downloads\tutori			tzki\Downloads\tutorialdata.zip\www3\		
Name	Größe	Gepac	Name	Größe	
mailsv	1 115 194		access.log	4 043 241	
vendor_sales	2 056 592		secure.log	1 103 496	
www1	5 432 793				
www2	5 109 542				
www3	5 146 737				

Bestandteil des offiziellen Splunk-Tutorials,
Download unter
<http://docs.splunk.com/images/Tutorial/tutorialdata.zip>

Demonstration: Hinzufügen der Daten



2

ERSTES ARBEITEN MIT SPLUNK

2.1

BEGRIFFE

- Jegliche Information, die in Splunk eingebracht wird, ist ein „Ereignis“, ein „Event“
- Im entspricht einer Zeile einer Log-Datei genau einem Event
 - Bestimmte Informationen wie detaillierte Fehlerbeschreibungen können mehrzeilige Blöcke erfordern
 - Splunk erkennt solche Blöcke dann als jeweils einen einzigen Event
- Events werden von Datenquellen („Data Sources“) geliefert
 - Dateien
 - Netzwerk
 - Ergebnisse eines Programmlaufs
- Wichtig:
 - Events werden in der Praxis automatisch kontinuierlich nach Splunk übertragen
 - Das hier eben praktizierte händische Hochladen einer Datei ist eher die Ausnahme

- Jeder Event hat mindestens 4 Felder
 - source
 - Die Datenquelle
 - sourcetype
 - Was für ein Typ von Information ist dies?
 - Splunk kennt einen Satz von „pretrained“ Source-Typen, die automatisch erkannt werden
 - Eigene Source-Typen können von einem Splunk-Administrator eingerichtet werden
 - host
 - Von welcher „Maschine“ wurde der Event generiert
 - Meistens ein Server im Netzwerk
 - Bei uns ist das der beim Hochladen vergeben Host-Name bzw. der Name eines darin enthaltenen Verzeichnisses
 - _time
 - Der Zeitstempel des Events

- Felder eines Events werden „indiziert“
 - Das ist eine Begriff aus der Datenbank-Welt
 - Ein Index ermöglicht eine schnelle Suche nach diesem Feld
- Beim Eintragen von Daten werden diese automatisch nach weiteren Feldern untersucht
 - Welche Felder gefunden werden „entscheidet“ Splunk im Standard alleine
 - Selbstverständlich können auch eigene Regeln definiert und Splunk beigebracht werden
 - Trennzeichen
 - Komplexe „Reguläre Ausdrücke“
 - Dies ist allerdings nicht für alle Anwender relevant

2.2

EINE EINFACHE SUCHE

Die Eingabe des Suchausdrucks

[Suche](#)[Analysen](#)[Datensets](#)[Berichte](#)[Benachrichtigungen](#)[Dashboards](#)[Search & Reporting](#)

Suche

Letzte 24 Stunden ▾



Kein Abruf von Beispielereignissen ▾

Modus "Ausführlich" ▾

Vorgehensweise beim Suchen

Wenn Sie mit den Suchfunktionen nicht vertraut sind oder weitere Informationen wünschen, sollten Sie eine der folgenden Ressourcen zu Rate ziehen.

[Dokumentation](#)[Trainings](#)

Gewünschte Suche

329.592

vor 2 Monaten

vor einem Monat

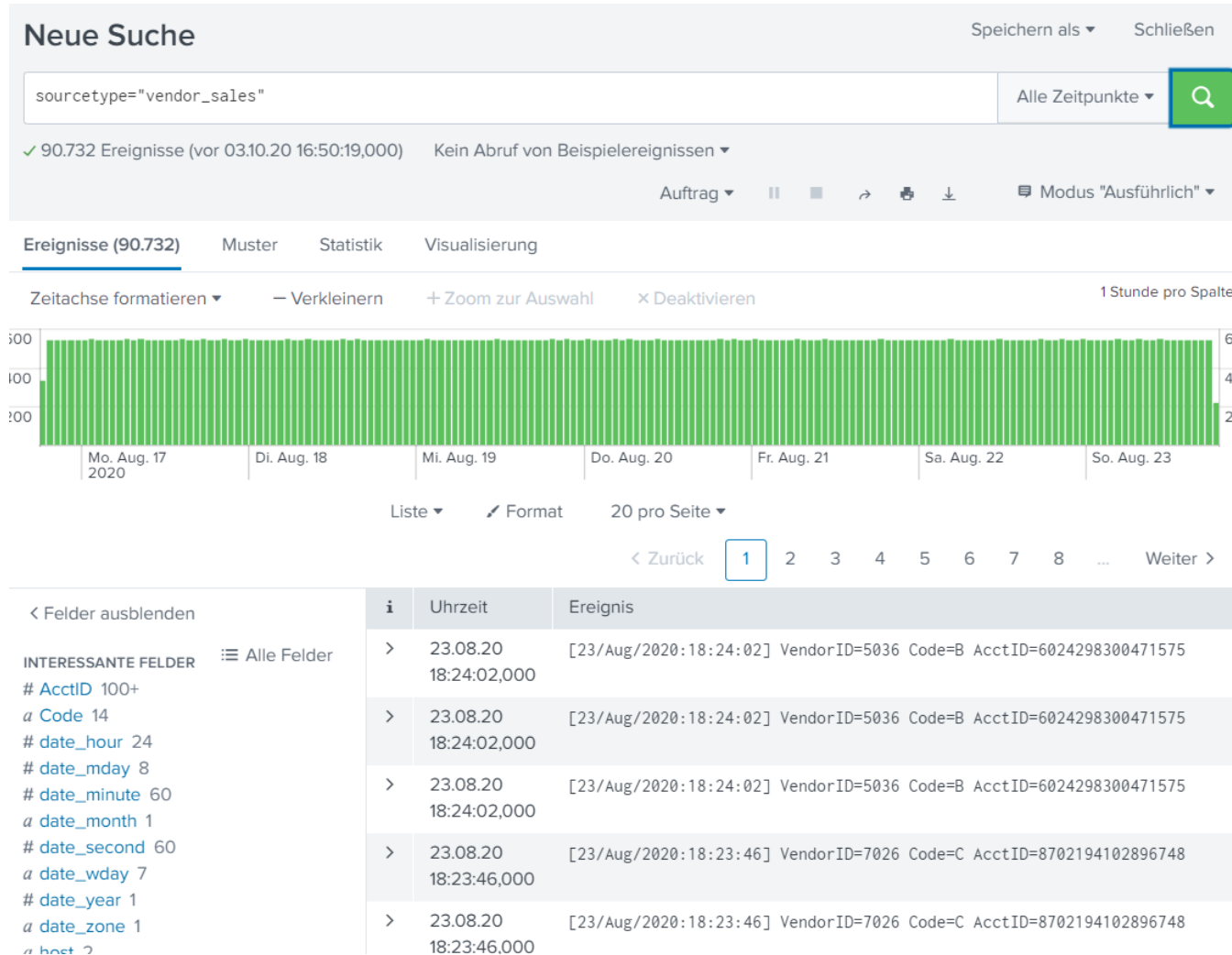
Ereignisse

FRÜHESTES EREIGNIS LETZTES EREIGNIS

INDIZIERT

[Datenzusammenfassung](#)[> Suchverlauf](#)

Beispiel und Darstellung der Ergebnisse



Weitere Auswahl mit „interessanten Feldern“

date_wday

7 Werte, 100 % der Ereignisse

Ausgewählt

Berichte

Oberste Werte

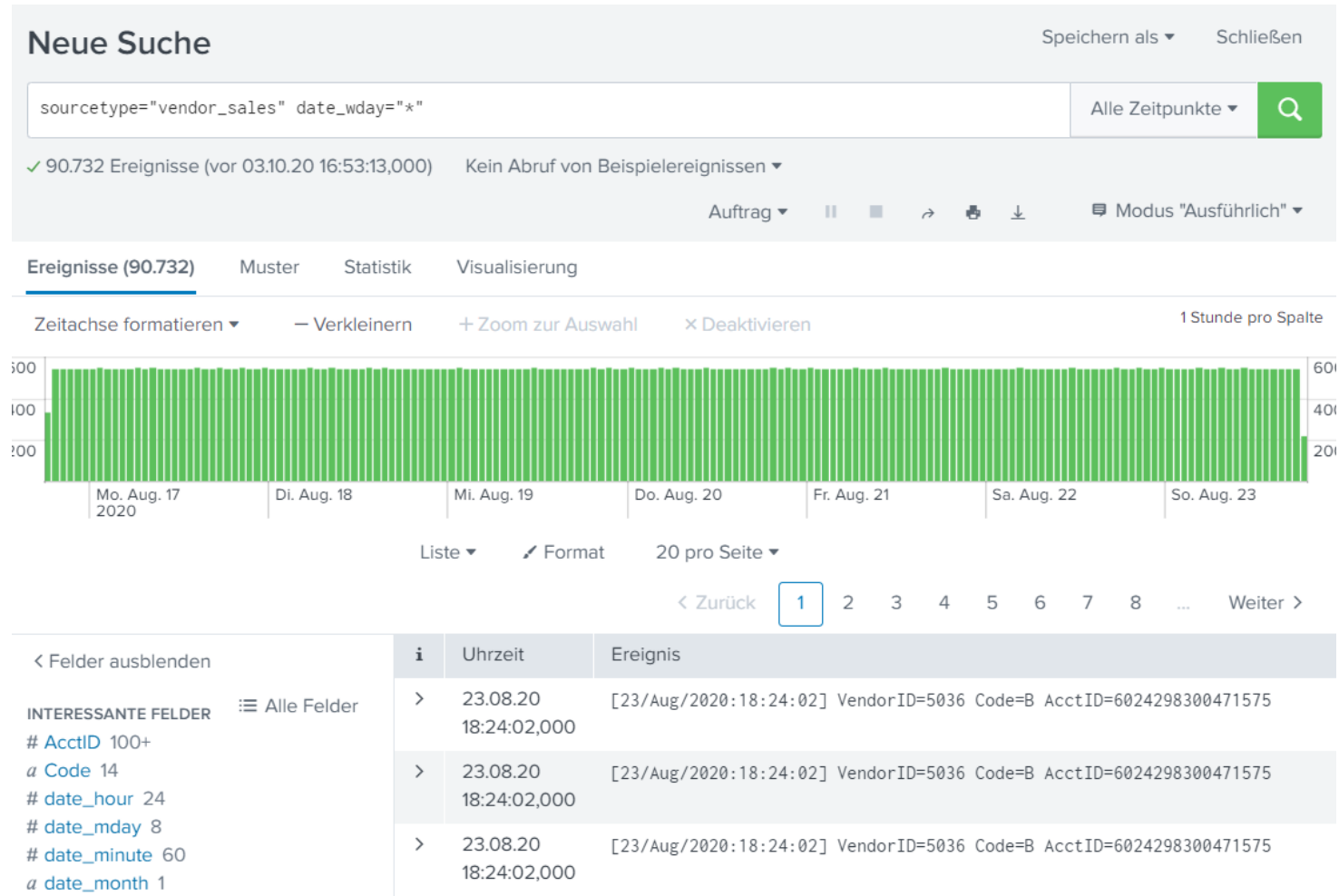
Spitzenwerte bis zum Zeitpunkt

Seltene Werte

Ereignisse mit diesem Feld

Werte	Anzahl	%	
sunday	12.969	14,294 %	
monday	12.963	14,287 %	
saturday	12.963	14,287 %	
friday	12.960	14,284 %	
thursday	12.960	14,284 %	
wednesday	12.960	14,284 %	
tuesday	12.957	14,28 %	

Neue Suche mit dem „interessanten Feld“



Neue Suche

Speiche

sourcetype="vendor_sales" date_wday="|"

✓ 90.732 Ereignisse (vor 03.10.20 16:53:13,

date_wday="friday"	Übereinstimmender Begriff
date_wday="monday"	Übereinstimmender Begriff
date_wday="saturday"	Übereinstimmender Begriff
date_wday="sunday"	Übereinstimmender Begriff
date_wday="thursday"	Übereinstimmender Begriff
date_wday="tuesday"	Übereinstimmender Begriff
date_wday="wednesday"	Übereinstimmender Begriff
sourcetype="vendor_...les" date_wday=""	Übereinstimmende Suche

Ereignisse (90.732)

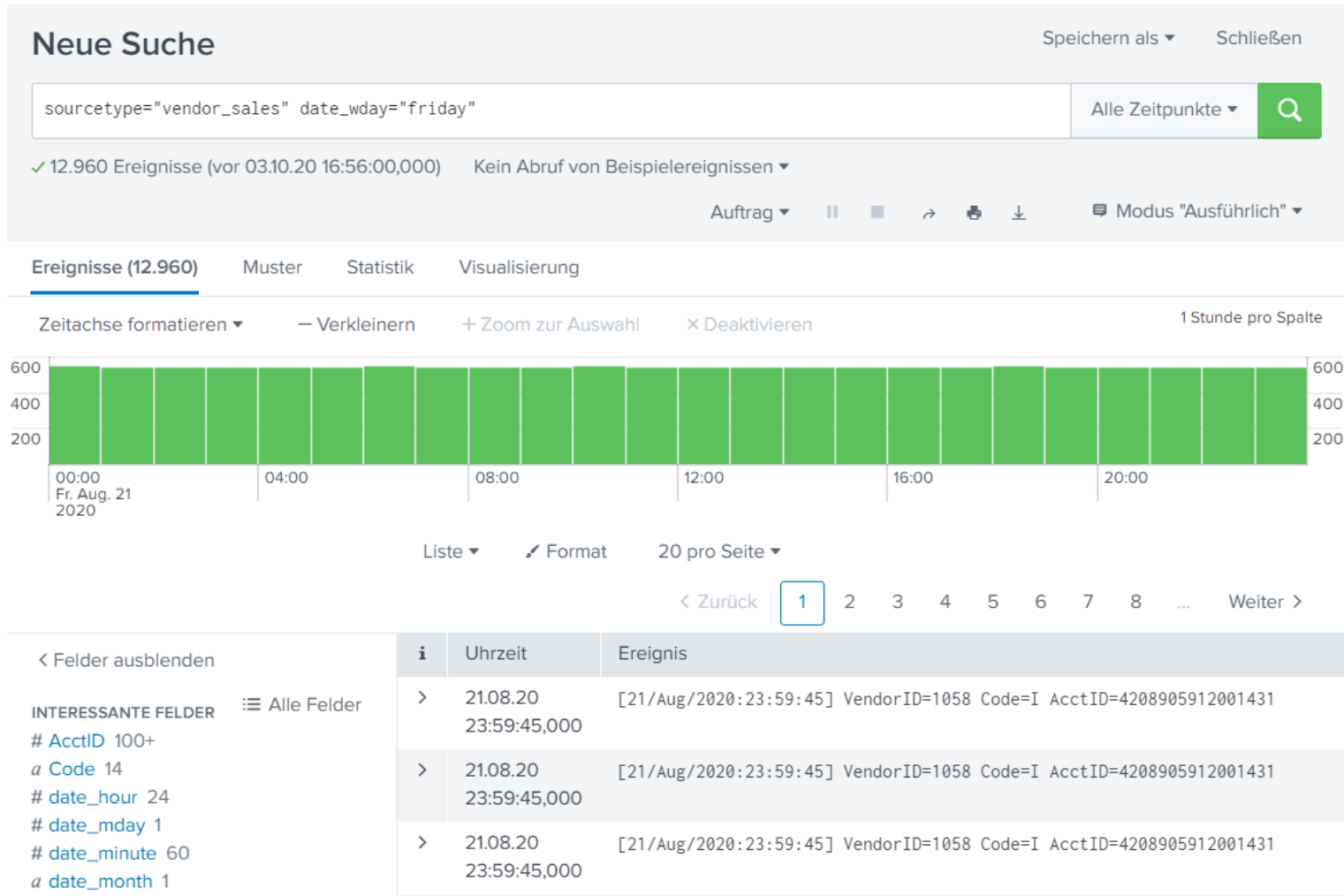
Muster

Statis

Zeitachse formatieren ▼

— Verkleinere

Das neu gefiltertes Ergebnis



3

SUCHEN

3.1

SUCHEN MIT DER SEARCH PROCESSING LANGUAGE

- Wir wissen, dass Splunk Events verarbeitet, die intern in Felder umgewandelt werden
- Ein Suchausdruck kombiniert nun diese Werte und liefert nur die Events, die zu den Kriterien passen
- Basis-Syntax
 - `feld1=wert1 feld2=wert2`

- Im Detail sind die Suchausdrücke alles andere als trivial und müssen syntaktisch korrekt formuliert werden
 - So müssen beispielsweise Werte mit Leerzeichen in Anführungszeichen gesetzt werden
 - Diese müssen dann aber auch korrekt abgeschlossen werden
- Werden mehrere Feld-Werte-Paare angegeben werden diese logisch mit UND verknüpft
 - Eine Oder-Verknüpfung erfolgt mit dem Schlüsselwort OR
- Im Endeffekt ist die Formulierung einer Suche die Aufgabe eines Programmierers!

- Assistent während der Eingabe
- Online-Referenz
 - <https://docs.splunk.com/Documentation/Splunk/8.0.6/SearchReference/Commandsbycategory>

- Für Splunk-Anwender ohne Programmier-Vorkenntnisse ist die Erstellung komplexer und effizienter Suchen recht mühsam
- Pragmatischer Ansatz
 - Helfen lassen: „Wie kann ich aus den Daten folgende Informationen extrahieren?“
 - Ein Anwender kann sich komplexe Abfragen als Bericht anzeigen lassen
 - Dieser wird von einem Splunk-Experten erstellt und abgespeichert
 - Ebenso können eigene sogenannte „Event Typen“ erstellt werden, die beliebig komplexe Such-Ausdrücke beinhalten können
 - Der Anwender benutzt dann nur noch den Ausdruck
`event_type=custom_event_type`

3.2

ANALYSIEREN MIT DER SEARCH PROCESSING LANGUAGE

- Das Thema „Programmierung“ wird hier noch weiter getrieben
- Hier ist die Darstellung der abstrakten Arbeitsweise von Splunk sinnvoll
 - Die Suche extrahiert aus der gesamten Splunk bekannten Event-Menge die passenden Events
 - Diese Ergebnismenge kann nun als neue Basismenge betrachtet werden, auf die weitere Kommandos angewendet werden können
 - Das neue Kommando wird mit einem String „|“, der sogenannten „Pipe“ eingeleitet



- `search`
 - Die Suche selbst ist bereits ein Kommando
 - `search sourcetype=„*“` ist äquivalent zu `sourcetype=„*“`
- `top <feld>`
 - Gruppiert die Treffermenge automatisch nach dem angegebenen Feld und zeigt die höchsten 5 Treffer an
 - Das Ergebnis von `top` ist damit nicht mehr ein Event sondern ein „Result“ mit den Feldern „count“ und „percent“

- Die Anzahl von Splunk-Kommandos ist beträchtlich
- Eine Übersicht wieder in der Online-Referenz
- Hinweis
 - Auch Berichte und damit Diagramme sind Kommandos!
 - `date_hour= "10" | timechart count by clientip limit="10"`

4

WEITERE FEATURES

4.1

DATENANALYSE

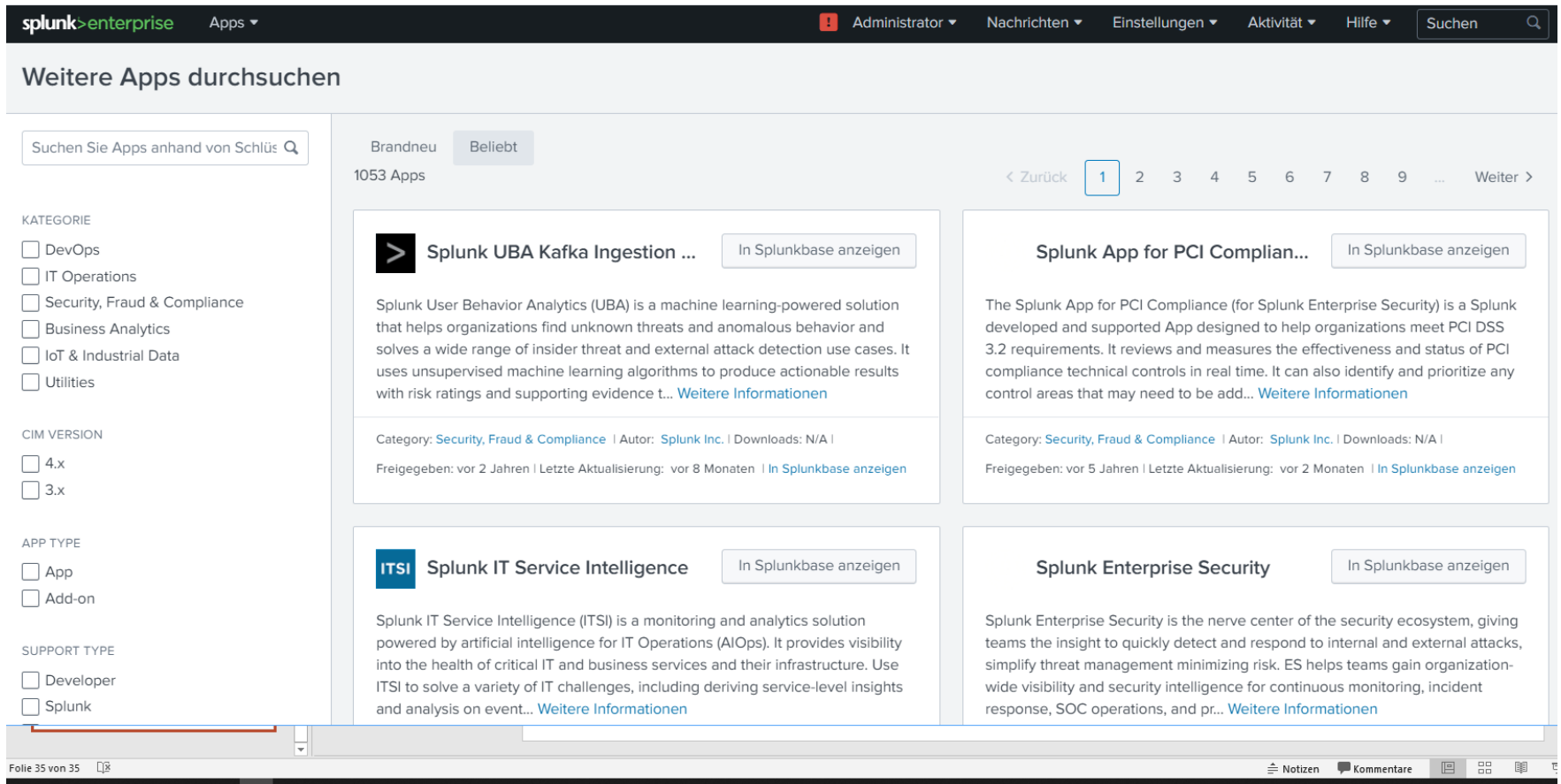
- Datasets
 - Hier werden Assoziationen = Beziehungen zwischen verschiedenen Datenquellen ermöglicht
 - Ein Datensatz definiert ein „Data Model“
 - dieses besteht aus einem Benutzer-definierten Satz von Feldern
- Das Common Information Model (CIM) gruppiert mehrere Data Models
 - und liefert somit eine normalisierte Datenmenge

- Mit Berichten können Ergebnisse visualisiert werden
- Die Berichte können als PDF exportiert werden
- Weiterhin können Berichte regelmäßig automatisiert ausgeführt werden
- Jede Bericht-Ausführung kann im Anschluss weitere Aktionen auslösen
 - Versand des Berichts als PDF
 - Benachrichtigungen an einen Administrator

4.2

ADD ONS

- Bisher haben wir ausschließlich 2 Apps gesehen:
 - Die Home-App mit der Startseite des Servers
 - Die Suchen-App
- Splunk bietet in seinem Marketplace zusätzliche Applikationen an, die, teilweise Lizenz-pflichtig, installiert werden können
- Solche Apps erweitern den Funktionsumfang von Splunk „beliebig“



The screenshot displays the Splunk App Marketplace interface. At the top, a navigation bar includes the Splunk logo, 'enterprise' label, 'Apps' dropdown, and user menu with 'Administrator', 'Nachrichten', 'Einstellungen', 'Aktivität', and 'Hilfe'. A search bar is on the right. Below the navigation bar, the heading 'Weitere Apps durchsuchen' is followed by a search input field. On the left, filters for 'KATEGORIE' (DevOps, IT Operations, Security, Fraud & Compliance, Business Analytics, IoT & Industrial Data, Utilities) and 'CIM VERSION' (4.x, 3.x) are shown. The main area displays a grid of app cards. The first card is 'Splunk UBA Kafka Ingestion ...' with a description of its machine learning capabilities and a 'In Splunkbase anzeigen' button. The second card is 'Splunk App for PCI Complian...', describing its role in meeting PCI DSS requirements. The third card is 'ITSI Splunk IT Service Intelligence', detailing its monitoring and analytics features. The fourth card is 'Splunk Enterprise Security', explaining its function as the nerve center of the security ecosystem. Each card includes category, author, download status, and release date information. A pagination bar at the top of the grid shows page 1 of 9. The bottom of the screenshot shows a presentation footer with 'Folie 35 von 35' and navigation icons.

splunk>enterprise Apps Administrator Nachrichten Einstellungen Aktivität Hilfe Suchen

Weitere Apps durchsuchen

Suchen Sie Apps anhand von Schlüsselwörtern

Brandneu Beliebt

1053 Apps

< Zurück 1 2 3 4 5 6 7 8 9 ... Weiter >

KATEGORIE

- ☐ DevOps
- ☐ IT Operations
- ☐ Security, Fraud & Compliance
- ☐ Business Analytics
- ☐ IoT & Industrial Data
- ☐ Utilities

CIM VERSION

- ☐ 4.x
- ☐ 3.x

APP TYPE

- ☐ App
- ☐ Add-on

SUPPORT TYPE

- ☐ Developer
- ☐ Splunk

Splunk UBA Kafka Ingestion ...

In Splunkbase anzeigen

Splunk User Behavior Analytics (UBA) is a machine learning-powered solution that helps organizations find unknown threats and anomalous behavior and solves a wide range of insider threat and external attack detection use cases. It uses unsupervised machine learning algorithms to produce actionable results with risk ratings and supporting evidence t... [Weitere Informationen](#)

Category: [Security, Fraud & Compliance](#) | Autor: [Splunk Inc.](#) | Downloads: N/A |
Freigegeben: vor 2 Jahren | Letzte Aktualisierung: vor 8 Monaten | [In Splunkbase anzeigen](#)

Splunk App for PCI Complian...

In Splunkbase anzeigen

The Splunk App for PCI Compliance (for Splunk Enterprise Security) is a Splunk developed and supported App designed to help organizations meet PCI DSS 3.2 requirements. It reviews and measures the effectiveness and status of PCI compliance technical controls in real time. It can also identify and prioritize any control areas that may need to be add... [Weitere Informationen](#)

Category: [Security, Fraud & Compliance](#) | Autor: [Splunk Inc.](#) | Downloads: N/A |
Freigegeben: vor 5 Jahren | Letzte Aktualisierung: vor 2 Monaten | [In Splunkbase anzeigen](#)

ITSI Splunk IT Service Intelligence

In Splunkbase anzeigen

Splunk IT Service Intelligence (ITSI) is a monitoring and analytics solution powered by artificial intelligence for IT Operations (AIOps). It provides visibility into the health of critical IT and business services and their infrastructure. Use ITSI to solve a variety of IT challenges, including deriving service-level insights and analysis on event... [Weitere Informationen](#)

Splunk Enterprise Security

In Splunkbase anzeigen

Splunk Enterprise Security is the nerve center of the security ecosystem, giving teams the insight to quickly detect and respond to internal and external attacks, simplify threat management minimizing risk. ES helps teams gain organization-wide visibility and security intelligence for continuous monitoring, incident response, SOC operations, and pr... [Weitere Informationen](#)

Folie 35 von 35 Notizen Kommentare