

# Log Analysis Machine Learner (LAMaLearner)



# Agenda

- Log Management
- Security Information and Event Management (SIEM)
- Herausforderungen
- Wie LAMaLearner hilft
- MAD2 & LAMaLearner
- Anwendungsfälle
- Demo

# Log Management

- Computer und Netzwerkkomponenten schreiben Log Files
- Diese können ausgewertet werden
  - Forensisch nach einem Fehler oder Vorfall
  - In Nahezu-Echtzeit während ein System arbeitet um Fehler zu bemerken
- Log Files können zentral gesammelt werden
  - Server mit Potentiell mehr Speicherplatz als auf individuellen Computern
  - Analyse auch möglich wenn Quell-Computer offline ist
- Kann für Compliance notwendig sein.



The image displays three screenshots of log files, each in a separate window titled 'Editor'.

**auth.log - Editor**

```
Datei Bearbeiten Format Ansicht Hilfe
May 19 06:29:32 hosting sshd[3540]: Address 188.114.89.141 maps to 141.89.114.188.ip4.netren.pl, but this does not
May 19 06:29:32 hosting sshd[3540]: pam_unix(sshd:auth): authentication failure; logname= uid=0 euid=0 tty=ssh ruse
May 19 06:29:34 hosting sshd[3540]: Failed password for root from 188.114.89.141 port 55205 ssh2
May 19 06:29:37 hosting sshd[3540]: Failed password for root from 188.114.89.141 port 55205 ssh2
May 19 06:29:37 hosting sshd[3540]: Received disconnect from 188.114.89.141: 11: [preauth]
```

**error.log - Editor**

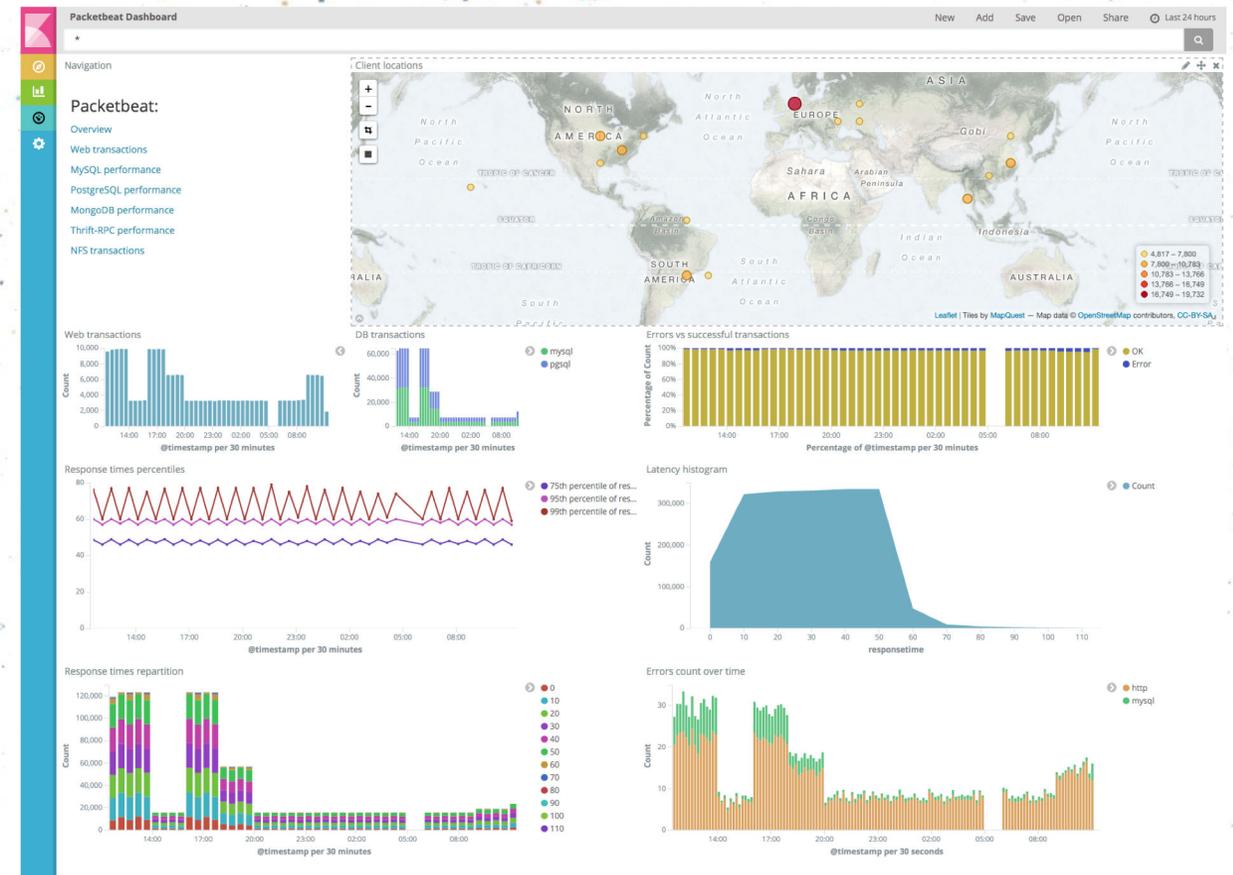
```
Datei Bearbeiten Format Ansicht Hilfe
[Sun May 19 06:27:30 2019] [error] python_init: Python version mismatch, expected '2.7.2+', found '2.7.3'.
[Sun May 19 06:27:30 2019] [error] python_init: Python executable found '/usr/bin/python'.
[Sun May 19 06:27:30 2019] [error] python_init: Python path being used '/usr/lib/python2.7/;/usr/lib/python2.7/plat
```

**syslog - Editor**

```
Datei Bearbeiten Format Ansicht Hilfe
May 22 06:25:17 hosting rsyslogd: [origin software="rsyslogd" swVersion="5.8.11" x-pid="3636" x-info="http://www.rs
May 22 06:25:18 hosting postfix/smtpd[26093]: connect from localhost[127.0.0.1]
May 22 06:25:18 hosting postfix/smtpd[26093]: 61227B7AD5: client=localhost[127.0.0.1]
```

# Security Information and Event Management (SIEM)

- Security zentrische Erweiterung zum normalen Log Management
- Sammelt Logs aus heterogenen Komponenten
- Ermöglicht Visualisierung von Event Häufigkeiten
  - Anzahl fehlgeschlagener Log-Ins je Zeitintervall
- Ermöglicht Definition von Alarmierungsregeln.  
Z.B.:
  - 3 fehlgeschlagene Log-Ins auf einer Maschine in einer Minute
  - 15 Firewall Rejections in einer Minute
  - 1 Malware erkannt ...



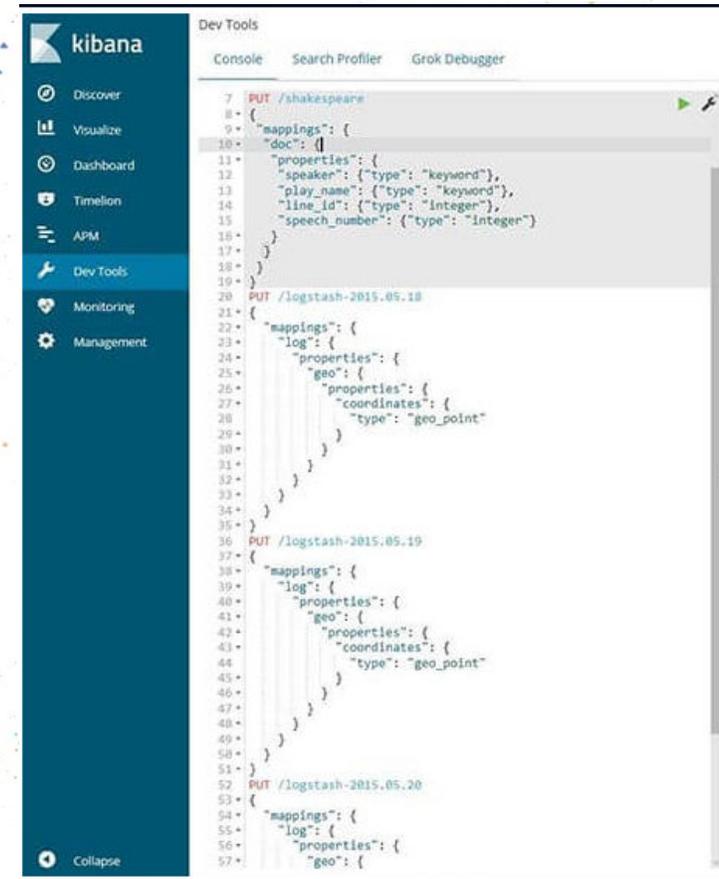
# SIEM und Log Management Herausforderungen

- Volume & Velocity
  - Log Daten einer Organisation können schnell hunderte GB erreichen
  - Sammeln und Auswerten kann Systeme überfordern
- Variety
  - Es gibt keine verbindliche Normen für Logs
  - Jedes System schreibt Logs nach anderen Mustern z.B.:
    - Name, IP, Status, Meldungstext
    - Meldungstext, Zeit, Name, Status
  - Aktuelle SIEM Systeme funktionieren nur mit manuell definierten Mustern
    - Meldungstexte können völlig individuell sein
    - Namen und IPs sind Organisationsspezifisch



# Beispiel dafür Variety Herausforderungen

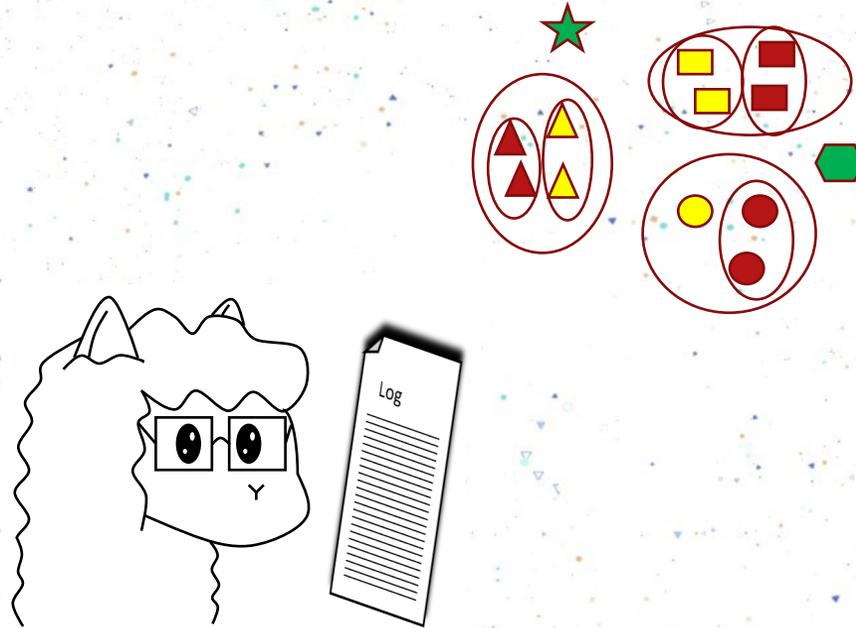
- Damit man in ElasticStack Events aggregiert werden können, muss ein entsprechendes Muster vorgegeben werden.
- Dafür muss das Muster bekannt sein.
- Das ist schwer, wenn man nicht vorher weiss, wonach man eigentlich sucht.



```
7 PUT /shakespeare
8 {
9   "mappings": {
10    "doc": {
11      "properties": {
12        "speaker": {"type": "keyword"},
13        "play_name": {"type": "keyword"},
14        "line_id": {"type": "integer"},
15        "speech_number": {"type": "integer"}
16      }
17    }
18  }
19 }
20 PUT /logstash-2015.05.18
21 {
22   "mappings": {
23     "log": {
24       "properties": {
25         "geo": {
26           "properties": {
27             "coordinates": {
28               "type": "geo_point"
29             }
30           }
31         }
32       }
33     }
34   }
35 }
36 PUT /logstash-2015.05.19
37 {
38   "mappings": {
39     "log": {
40       "properties": {
41         "geo": {
42           "properties": {
43             "coordinates": {
44               "type": "geo_point"
45             }
46           }
47         }
48       }
49     }
50   }
51 }
52 PUT /logstash-2015.05.20
53 {
54   "mappings": {
55     "log": {
56       "properties": {
57         "geo": {
```

# LAMaLearner lernt selbst und hilft

- LAMaLearner analysiert Log Einträge mit unbekanntem Format
- LAMaLearner erzeugt Cluster inhaltlich ähnlicher Log Einträge
- Cluster können hierarchisch sein
- LAMaLearner erzeugt gemeinsame Muster für erkannte Cluster



# LAMaLearner erzeugt hierarchische Cluster aus Logs

- 0 - 2019.05.16 - 09:58:00 - Router myrouter2 interface 10.10.10.12 down  
[Delete](#)
- 1 - 2019.05.16 - 09:58:01 - Router myrouter1 interface 192.168.13.1 down  
[Delete](#)
- 2 - 2019.05.16 - 09:58:02 - Router myrouter5 interface 192.168.22.33 down  
[Delete](#)
- 3 - 2019.05.16 - 09:58:03 - Router myrouter6 interface 192.168.22.31 down  
[Delete](#)
- 4 - 2019.05.16 - 09:58:04 - Router myrouter6 interface 192.168.22.31 down  
[Delete](#)
- 5 - 2019.05.16 - 09:58:05 - Connection from 192.168.1.1  
[Delete](#)
- 6 - 2019.05.16 - 09:58:06 - RSA key-gen for Tobi completed  
[Delete](#)
- 7 - 2019.05.16 - 09:58:07 - Password authentication for Willi accepted  
[Delete](#)
- 8 - 2019.05.16 - 09:58:08 - Interface 5 shut down  
[Delete](#)
- 9 - 2019.05.16 - 09:58:09 - RSA key-gen for Manfred completed  
[Delete](#)
- 10 - 2019.05.16 - 09:58:10 - Connection from 192.168.3.5  
[Delete](#)
- 11 - 2019.05.16 - 09:58:02 - Router myrouter5 interface 192.168.22.33 down  
[Delete](#)

## Available clusters:

- Id: 0: Router interface down Total Risk Score: 0 [Events](#) Labels:
  - Id: 1: Router myrouter5 interface 192.168.22.33 down Total Risk Score: 0 [Events](#) Labels:
  - Id: 2: Router myrouter6 interface 192.168.22.31 down Total Risk Score: 0 [Events](#) Labels:
- Id: 3: Connection from Total Risk Score: 0 [Events](#) Labels:
- Id: 4: RSA key-gen for completed Total Risk Score: 0 [Events](#) Labels:

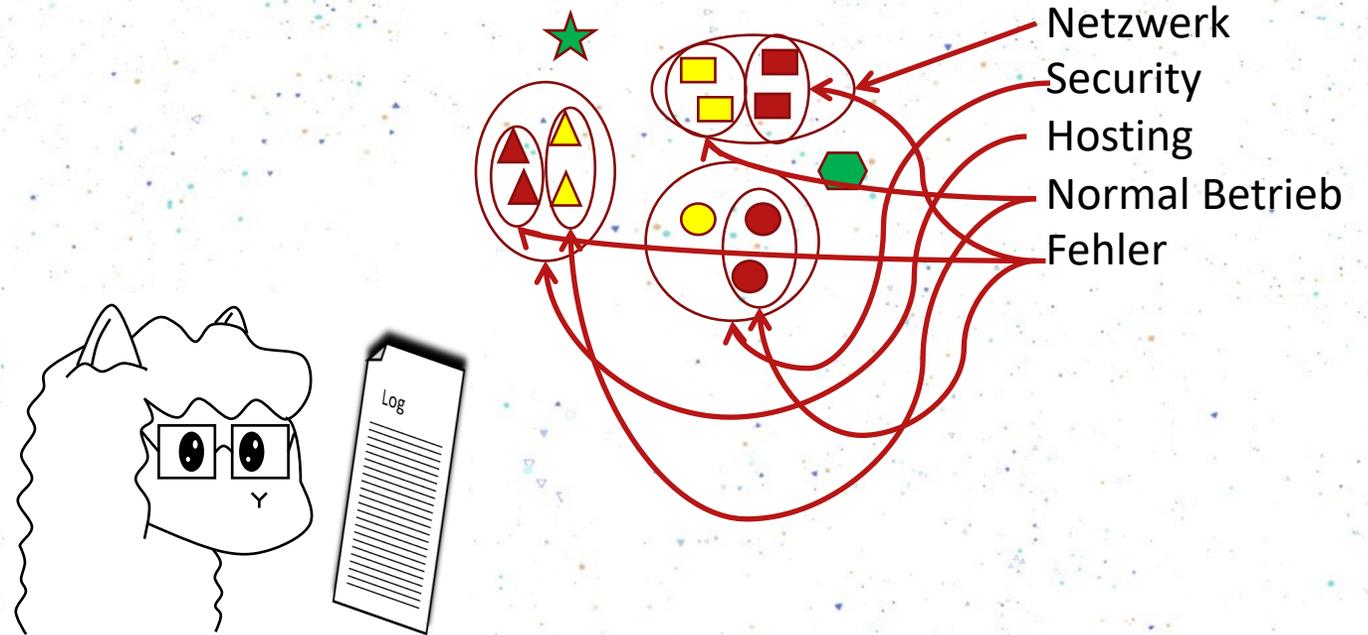
## Events that are not in clusters:

### [Show in Events](#)

- Id: 7
- Id: 8

# LAMaLearner lernt selbst und hilft

- LAMaLearner ermöglicht Annotation von Clustern mit beliebigen Labeln
- Mehrere Label können pro Cluster vergeben werden



# Cluster werden durch frei wählbare Labels annotiert

## Available clusters:

- Id: 0: Router interface down Total Risk Score: 5 [Events](#) Labels:  
failure ✓   
network ✓   
success ▾
- Id: 1: Router myrouter5 interface 192.168.22.33 down Total Risk Score: 5 [Events](#) Labels:  
failure ✓   
network ✓   
success ▾
- Id: 2: Router myrouter6 interface 192.168.22.31 down Total Risk Score: 5 [Events](#) Labels:  
failure ✓   
network ✓   
success ▾
- Id: 3: Connection from Total Risk Score: 0 [Events](#) Labels:  
success ▾
- Id: 4: RSA key-gen for completed Total Risk Score: -3.5 [Events](#) Labels:  
success ✓   
security ✓   
failure ▾

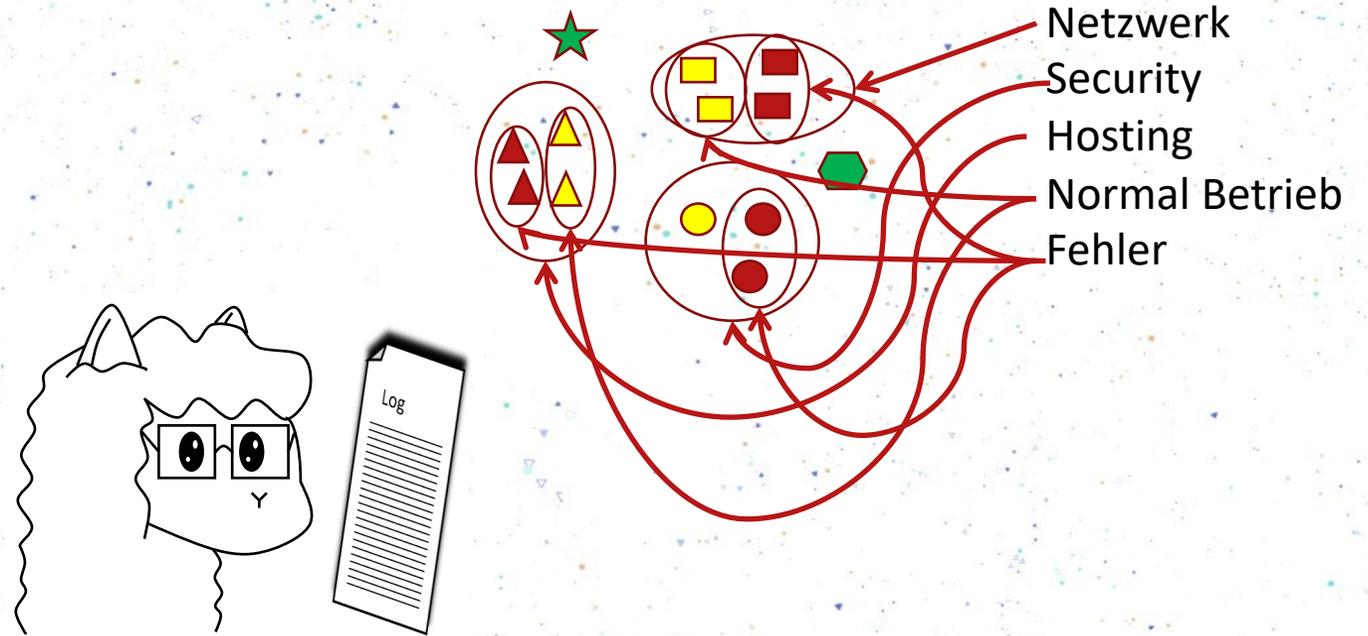
# LAMaLearner lernt selbst und hilft

- Sind genügend Cluster annotiert kann LAMaLearner selbst lernen Cluster zu unterscheiden.

Z.B.:

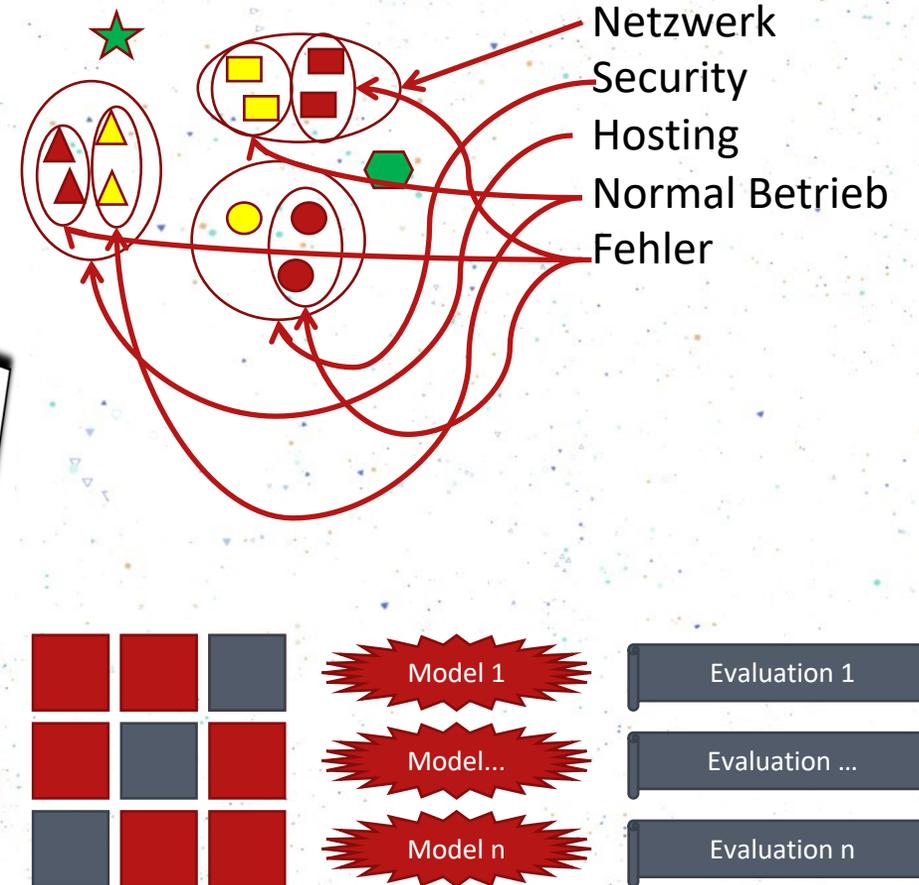
- Viereck → Netzwerk
- Dreieck → Hosting
- Kreis → Security
- Rot → Fehler
- !Rot → Normal Betrieb

- Dieses Modell wird komplett eigenständig gelernt.



# So lernt LAMaLearner und testet sich

- LAMaLearner testet sich selber mit n-fold cross-validation
  - n ist eine natürliche Zahl. Als Beispiel hier n=3.
  - Alle bekannten Zuordnungen werden in n Teilmengen unterteilt
  - Dann in n Iterationen:
    - n-1 Teilmengen zum lernen
    - 1 Teilmenge zum testen
- Bestes Modell wird ausgewählt und gespeichert



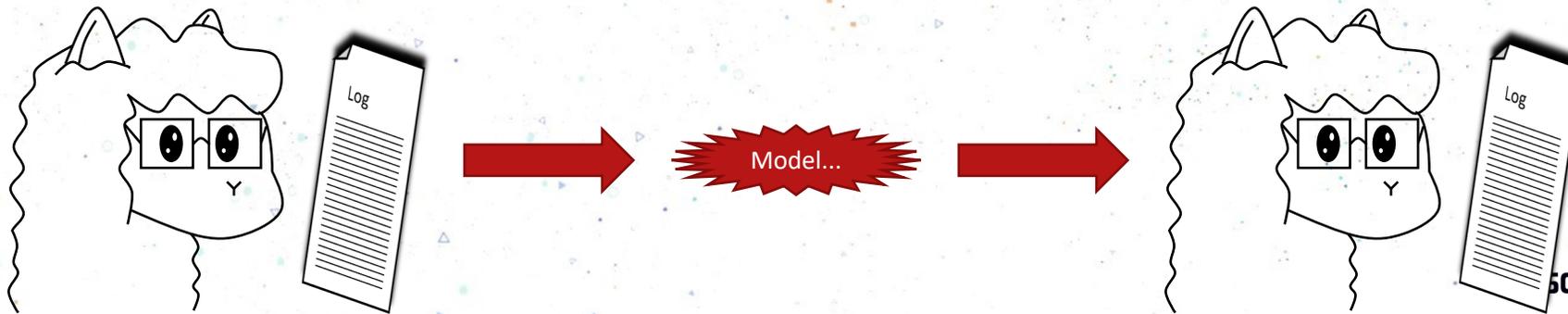
# LAMaLearner generiert Modelle

- Modelle speichern Labels und Regeln diese zu identifizieren
- Modelle können gespeichert und zwischen LAMaLearner Instanzen übertragen werden.
- Evaluationsdaten werden als Metadaten ins Modell integriert

- Creator: Tobias Eljasik-Swoboda
- Log Type: Windows Security Event Log
- Usage: Tells failures from successes
- Creation Date: 2019-09-02 13:50:48.149
- Version: 0.10.1-SNAPSHOT
- Training description: Trained and evaluated on 744 clusters using 3-fold cross-validation. Selection Policy: MicroaverageF1

Fold:		0						
Microaverage Precision:	1	Microaverage Recall:	1	Microaverage F1:	1			
Macroaverage Precision:	1	Macroaverage Recall:	1	Macroaverage F1:	1			
Category Id:	Category Label:	True Positives	False Positives	False Negatives	Precision	Recall	F1	
0	success	123	0	0	1	1	1	
1	failure	4	0	0	1	1	1	
Fold:		1						
Microaverage Precision:	1	Microaverage Recall:	1	Microaverage F1:	1			
Macroaverage Precision:	1	Macroaverage Recall:	1	Macroaverage F1:	1			
Category Id:	Category Label:	True Positives	False Positives	False Negatives	Precision	Recall	F1	
0	success	124	0	0	1	1	1	
1	failure	2	0	0	1	1	1	
Fold:		2						
Microaverage Precision:	1	Microaverage Recall:	1	Microaverage F1:	1			
Macroaverage Precision:	1	Macroaverage Recall:	1	Macroaverage F1:	1			
Category Id:	Category Label:	True Positives	False Positives	False Negatives	Precision	Recall	F1	
0	success	125	0	0	1	1	1	
1	failure	1	0	0	1	1	1	

© copyright ONTEC AG & Schoeller Network Control GmbH 2019



# Wie wird Effektivität gemessen?

- True Positive (TP): Getesteter Cluster hatte tatsächlich Label das es haben sollte.
- False Positive (FP): Getesteter Cluster wurde dem falschen Label zugeordnet
- False Negative (FN): Getester Cluster hätte dieses Label haben sollen, hatte aber ein anderes
- Precision: Verhältnis TP zu FP: Wieviel vom gefundenen richtig ist.
- Recall: Verhältnis TP zu FP: Wieviel vom zu Findenden gefunden wurde
- F1: Geometrisches Mittel aus Precision und Recall
- Microaverage: Gesamtwerte aus einzelnen TP, FP und FN berechnen
- Macroaverage: Durchschnittswerte

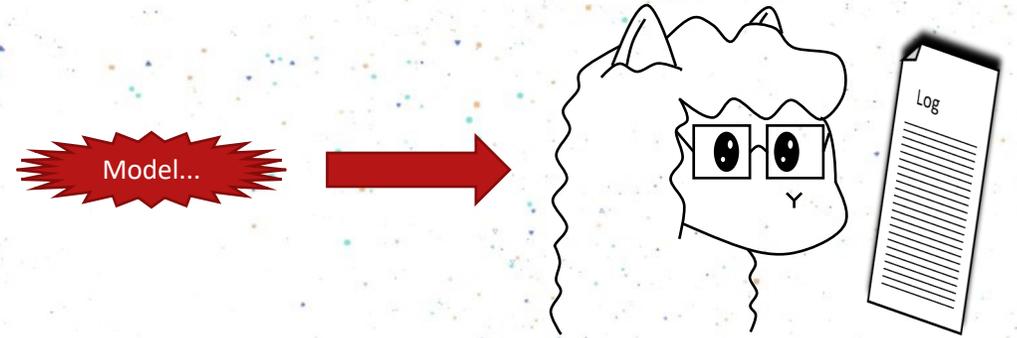
- Creator: Tobias Eljasik-Swoboda
- Log Type: Windows Security Event Log
- Usage: Tells failures from successes
- Creation Date: 2019-09-02 13:50:48.149
- Version: 0.10.1-SNAPSHOT
- Training description: Trained and evaluated on 744 clusters using 3-fold cross-validation. Selection Policy: MicroaverageF1

Fold:		0						
Microaverage Precision:		1	Microaverage Recall:		1	Microaverage F1:		1
Macroaverage Precision:		1	Macroaverage Recall:		1	Macroaverage F1:		1
Category Id:	Category Label:	True Positives	False Positives	False Negatives	Precision	Recall	F1	
0	success	123	0	0	1	1	1	
1	failure	4	0	0	1	1	1	
Fold:		1						
Microaverage Precision:		1	Microaverage Recall:		1	Microaverage F1:		1
Macroaverage Precision:		1	Macroaverage Recall:		1	Macroaverage F1:		1
Category Id:	Category Label:	True Positives	False Positives	False Negatives	Precision	Recall	F1	
0	success	124	0	0	1	1	1	
1	failure	2	0	0	1	1	1	
Fold:		2						
Microaverage Precision:		1	Microaverage Recall:		1	Microaverage F1:		1
Macroaverage Precision:		1	Macroaverage Recall:		1	Macroaverage F1:		1
Category Id:	Category Label:	True Positives	False Positives	False Negatives	Precision	Recall	F1	
0	success	125	0	0	1	1	1	
1	failure	1	0	0	1	1	1	

© copyright ONTEC AG & Schoeller Network Control GmbH 2019

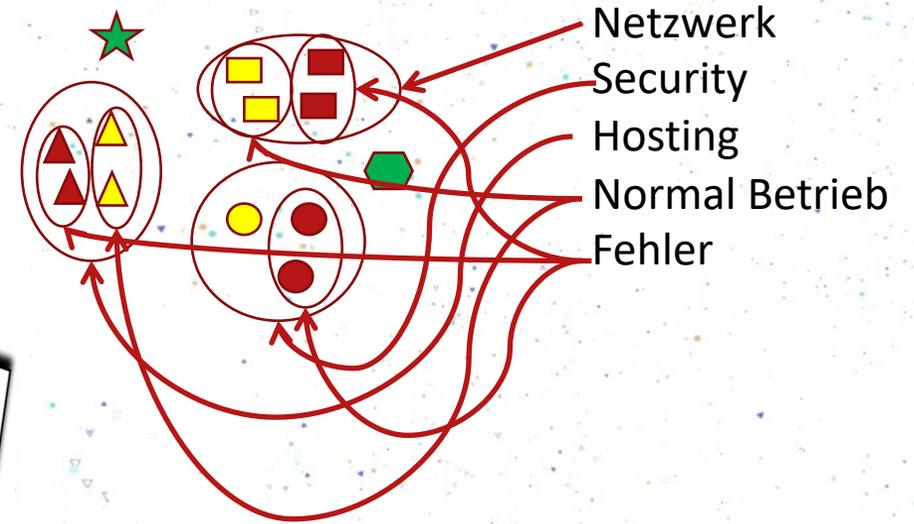
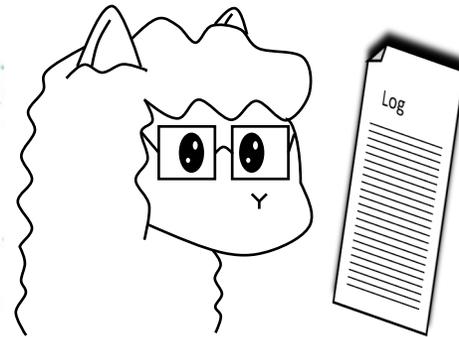
# LAMaLearner macht, was das gelernte Modell sagt

- Das heisst, LAMaLearner Instanzen müssen nicht bei 0 Anfangen.
- Neue Instanzen könnten mit vorhandenen Modellen / Wissen gestartet werden.
- So kann LAMaLearner Events mit gelernten Labels annotieren.
- Dies kann z.B. einen Pattern Matcher in Log Beats ersetzen ohne selber Pattern definieren zu müssen.



# Alternativ: Regel basierte Labels

- Neben dem Machine Learning Model kann LAMaLearner auch Label nach gefundenen Begriffen zuordnen.
- Diese können variabel mit Regeln kombiniert werden:
  - Nur Modell
  - Modell AND Regeln
  - Modell OR Regeln
  - Nur Regeln



My rules:  
Down → Fehler  
Success → Normal Betrieb  
Router → Netzwerk

# LAMaLearner lernt selbst und hilft

- LAMaLearner mit Regeln und/oder gelerntem Modell kann
  - Einzelne Events mit völlig unbekanntem Format Labeln zuordnen.
  - Neue Cluster finden und diese auch Labeln zuordnen.
- LAMaLearner kann alarmieren
  - Labels haben zugeordneten Risk Score.
  - Wenn Gesamt Risk Score eines Events oder Clusters > Threshold, wird externes Kommando aktiviert.



Model...

My rules:  
Down → Fehler  
Success → Normal Betrieb  
Router → Netzwerk



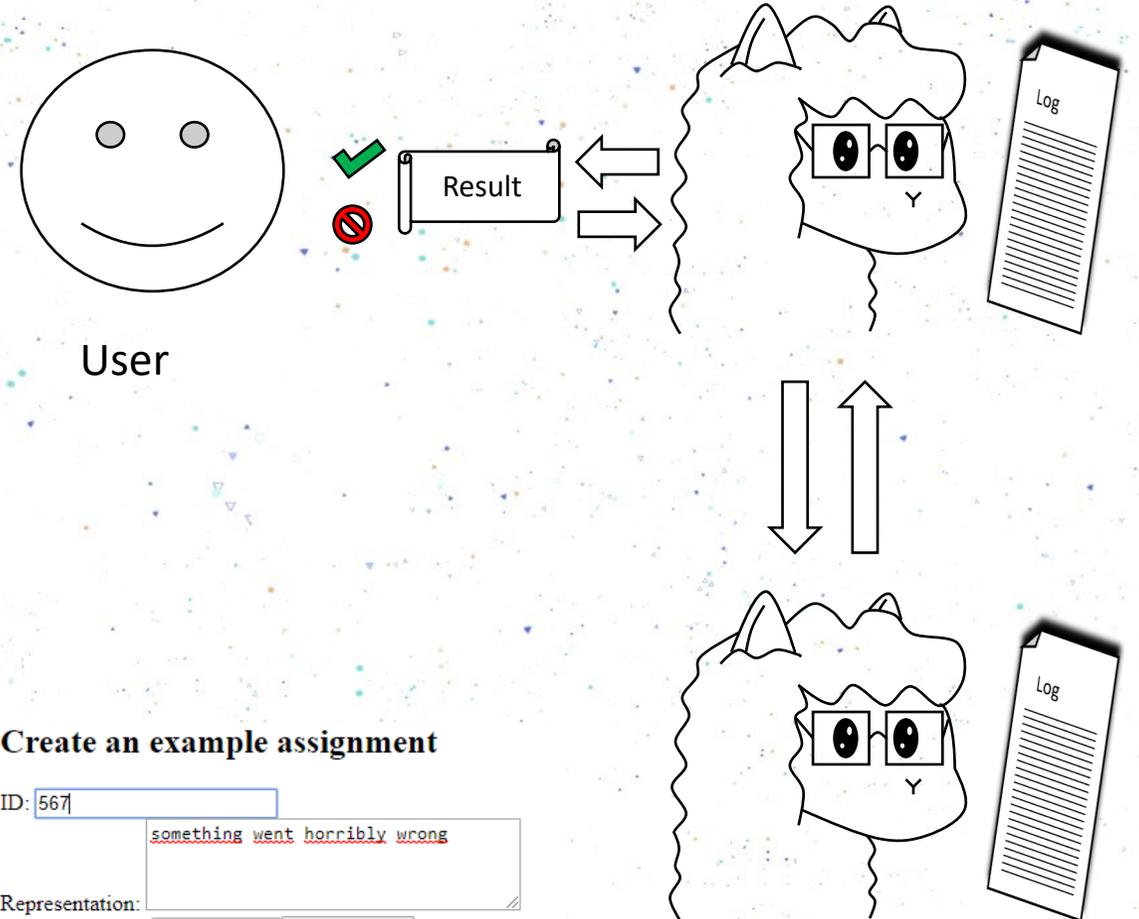
Labels: All Events success failure network security

### Available labelings:

- Id: 218: 4625 Microsoft-Windows-Security-Auditing Logon Audit\_Failure Security dc02.schoeller.local MAGGIES SCHOELLER NtLmSsp\_NTLM MAGGIE 0 0x0 %2313 0xc0000064 Labels:  
failure    
success   add
- Id: 560: 4625 Microsoft-Windows-Security-Auditing Logon Audit\_Failure Security dc02.schoeller.local MAGGIES SCHOELLER NtLmSsp\_NTLM MAGGIE 0 0x0 %2313 0xc0000064 Labels:  
failure    
success   add
- Id: 853: 4625 Microsoft-Windows-Security-Auditing Logon Audit\_Failure Security dc02.schoeller.local MAGGIES SCHOELLER NtLmSsp\_NTLM MAGGIE 0 0x0 %2313 0xc0000064 Labels:  
failure    
success   add
- Id: 854: 4625 Microsoft-Windows-Security-Auditing Logon Audit\_Failure Security dc02.schoeller.local MAGGIES SCHOELLER NtLmSsp\_NTLM MAGGIE 0 0x0 %2313 0xc0000064 Labels:  
failure    
success   add

# LAMaLearner lernt immer weiter

- User kann LAMaLearner Ergebnisse immer bestätigen oder korrigieren.
- Jeder Hinweis wird als Faktenwissen in LAMaLearner Instanz gespeichert
  - Event 1 hat Label A
  - Cluster 3 hat Label B
  - ...
- Fakten können zwischen LAMaLearner Instanzen kopiert und zwischengespeichert werden.
- Fakten können editiert und manuell angelegt werden



## Create an example assignment

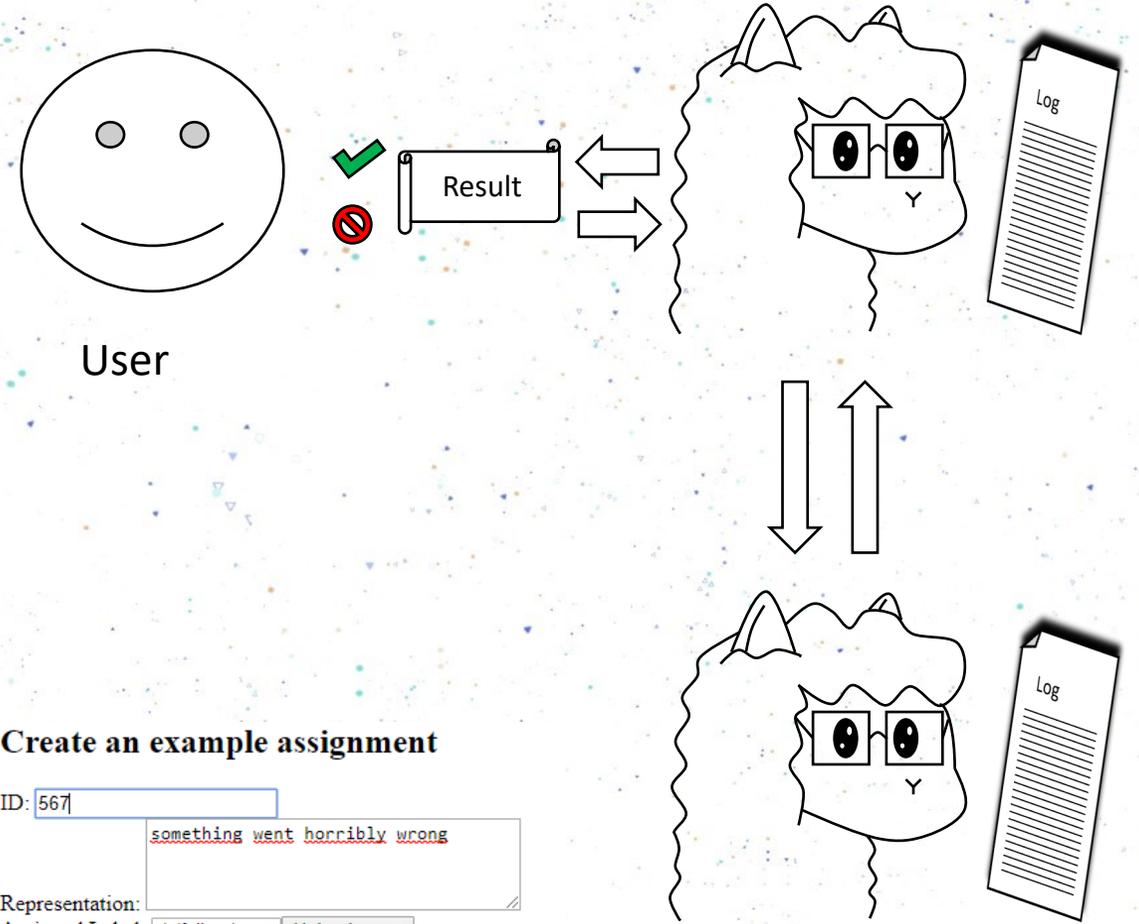
ID:

Representation:

Assigned Label:

# LAMaLearner lernt immer weiter

- Aus mehr Fakten können bessere Modelle generiert werden.
- Bessere Modelle bringen bessere Ergebnisse.
- Modelle können zwischen LAMaLearner Instanzen kopiert werden.



## Create an example assignment

ID:

Representation:

Assigned Label:

# LAMaLearner vereinfacht Variety Herausforderungen

- LAMaLearner lernt Muster aus Log Events:  
Keine manuelle Definition notwendig.
- LAMaLearner macht Logs Übersichtlicher.
  - Millionen Events
  - 10 – 100 Cluster
  - 2-5 Labels
- LAMaLearner kann auch neue Formate Labeln
  - Z.B. Modell aus bekannten Logs erzeugen und dann auf neues Format anwenden.
  - Hier: Gelernt aus Windows Security Event Log – Gelabelt: Firewall Log

Labels:

## Available labelings:

- Id: 0: QoS Firewall goliath1 Accept 85.15.47.176 hades\_DC\_ext (37.252.250.219) smtp (TCP/25) smtp Traffic Accepted from 85.15.47.176 to 37.252.250.219 25fcfad3-0100-00c0-5ce5-181200000014 @A@@B@1558476000@C@10047419 aresmgr (192.168.105.13) inbound bond1.50 false false 114 34558 25 TCP (6) External Internal smtp hades.dmz3.ontec.at (192.168.123.15) 0 0 206 1 2019-05-22T09:36:18Z Connection goliath-policy aresmgr {1A6AE460-5890-8D44-AC5D-03984E46FDC1} Today 11:33:35 AMAccess 0 incoming to Hades 194 5044f09d-5722-4603-8828-332844078160 goliath-policy Security Labels:

success

failure

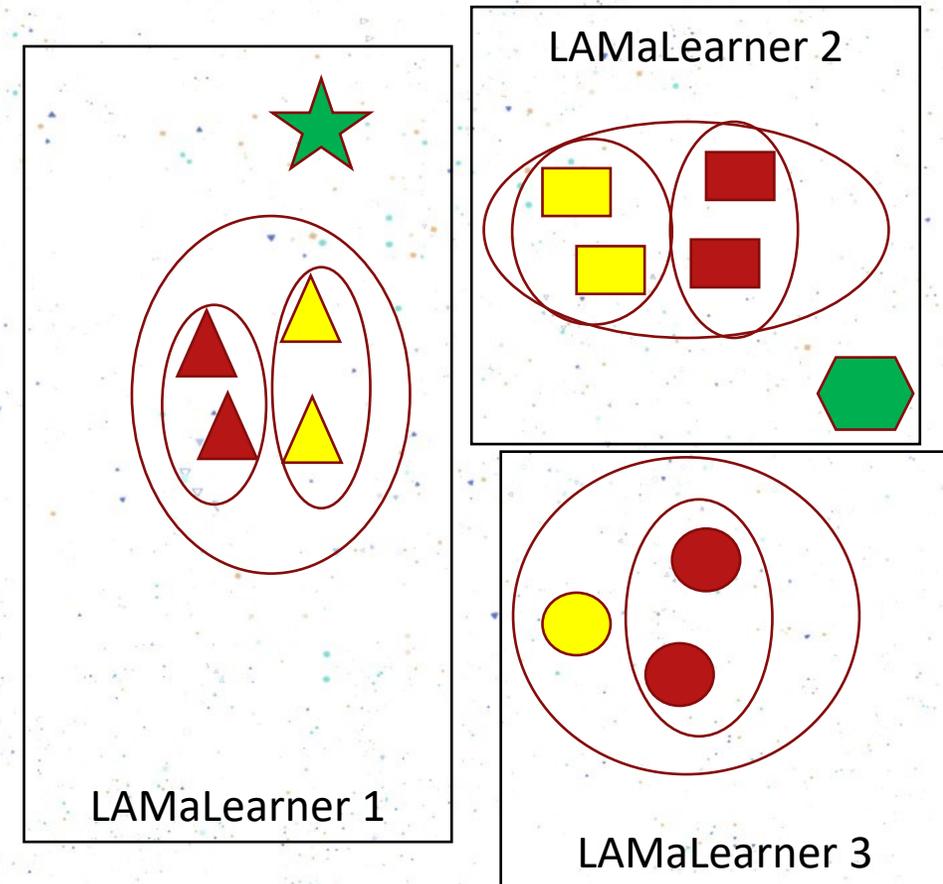
- Id: 1: QoS Firewall goliath1 Accept gca-app-dev-a01.dev0.ontec.at (172.27.30.41) docker-mgmt-01.olymp.ontec.at (192.168.100.150) https (TCP/443) https Traffic Accepted from 172.27.30.41 to 192.168.100.150 25fcfad3-0000-00c0-5ce5-181200000020 @A@@B@1558476000@C@10047418 aresmgr (192.168.105.13) inbound bond2.344 false false 112 56612 443 TCP (6) Internal Internal https 2019-05-22T09:36:18Z Connection goliath-policy aresmgr {1A6AE460-5890-8D44-AC5D-03984E46FDC1} Today 11:33:35 AMAccess 0 97 2003b088-4781-449a-a609-e6a5c95ea125 goliath-policy Security Labels:

success

failure

# LAMaLearner ist Ready für Volume & Velocity

- Events können auf mehrere LAMaLearner Instanzen verteilt werden.
- Einzelne Instanzen können separat
  - Cluster finden
  - Cluster labeln
  - Events labeln
- Ergebnisse einzelner LAMaLearner Instanzen können wieder zusammengeführt werden.



# LAMaLearner ist mobil und flexibel

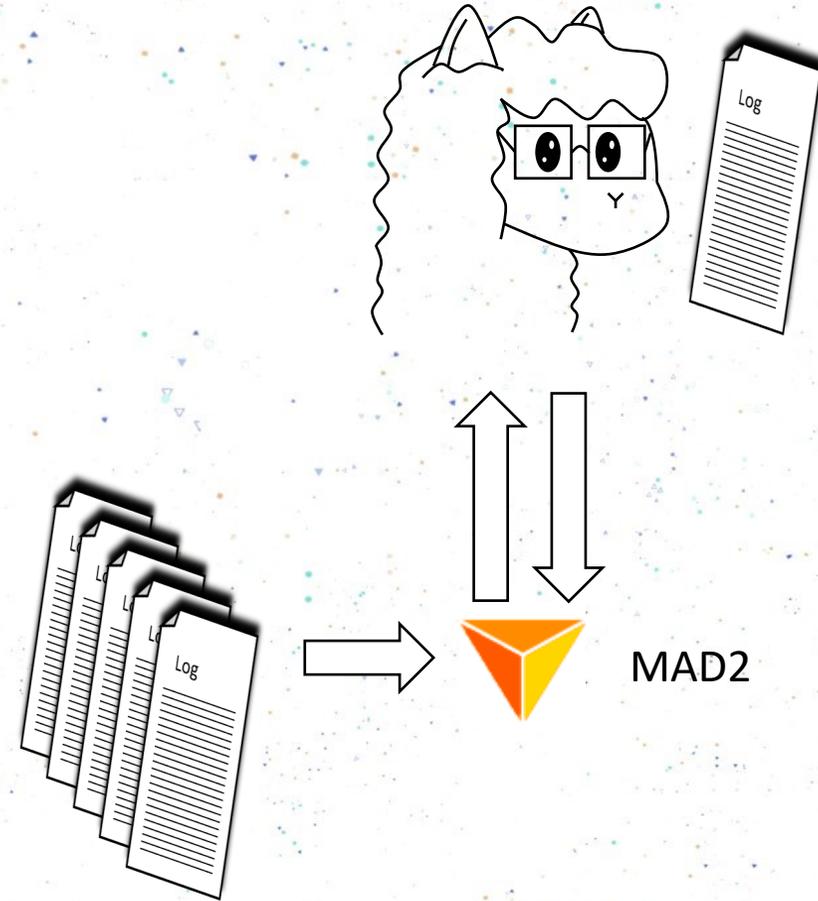
- LAMaLearner braucht keine Verbindung in eine Cloud
- LAMaLearner läuft überall wo es Java gibt
  - Getestet mit Java 9.0.1
  - auf Windows 10, Ubuntu Linux und Mac OS X
- LAMaLearner kommuniziert über eine einfache REST/JSON API
  - Port in Config File konfigurierbar
  - Mehrere Instanzen auf gleichem Rechner möglich
- Bringen Sie LAMaLearner zu Ihren Daten, nicht Ihre Daten zum LAMaLearner
  - Alle Funktionen in 24 MB Software
  - Model Dateien < 1 MB
- Scale Out Anywhere: Last Abhängig können LAMaLearner Work Loads auf beliebig viele Maschinen verteilt werden.
  - Auf Laptop mit Windows 10, Intel i7-7820 (Quad Core, 2.9 GHz) / 16 GB RAM:
    - Clustering: 1000 Events in ca. 3 Sekunden. 336 Events pro Sekunde
    - Event Labeling: 1000 Events in 100 ms. 10000 Events pro Sekunde

# Log files als REST / JSON ?!

- Log Files werden normalerweise nicht als JSON file geschrieben.
- Im ELK Stack werden die Events in Elastic Search gespeichert.
- Elastic Search hat auch eine REST / JSON Schnittstelle.
- Logstash sammelt Log Files für Elastic Search
- LAMaLearner nutzt zu diesem Zweck MAD2

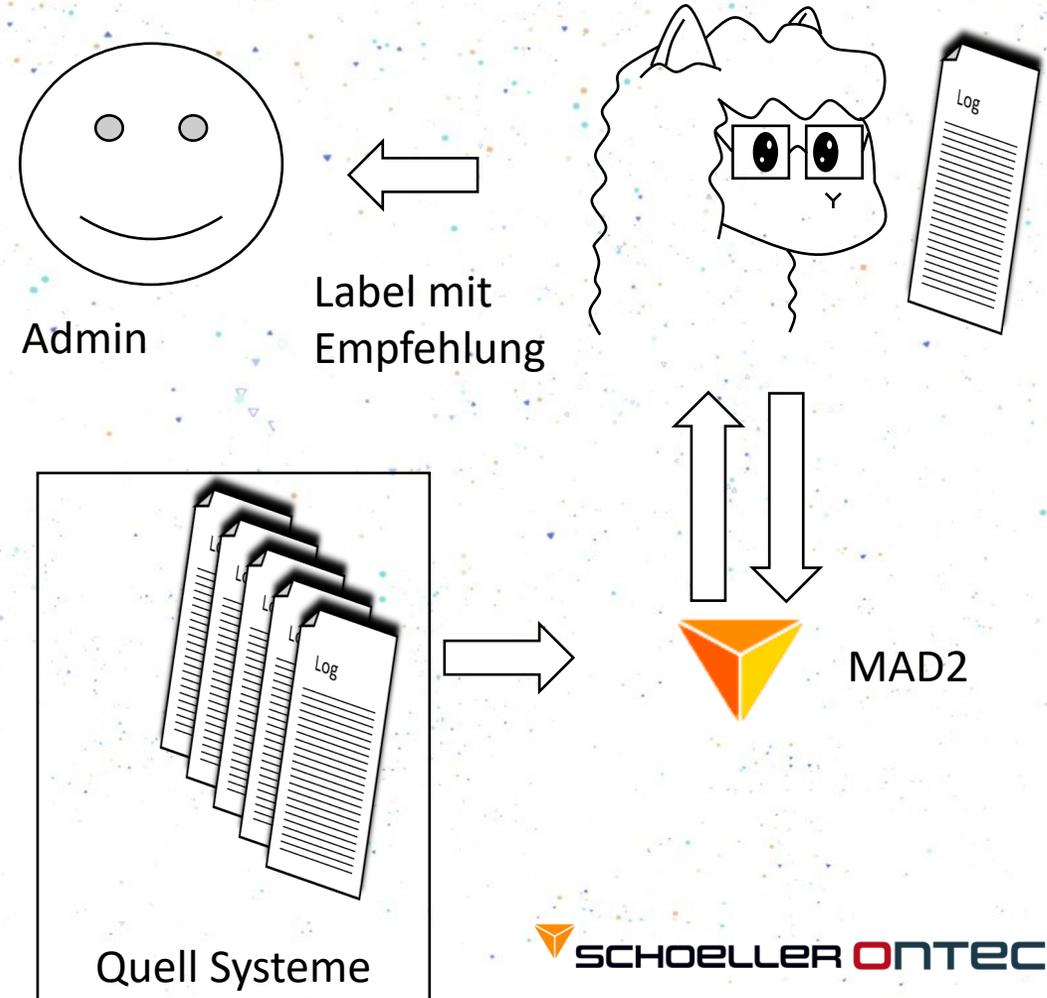
# LAMaLearner und MAD2

- Das Modular Abstract Data processing Tool (MAD2)
  - Kann Log Files aus diversen Quellsystemen sammeln
  - Speichert Log Files zentral
  - Verarbeitet 10.000 Events pro Sekunde / Instanz
  - Kompressionsrate: 3% der Original Log Größe
  - Generiert Reports nach fixen Regeln
- Bestehende Schnittstelle zur Verwendung von LAMaLearner aus MAD2
- Zwei Tools zum intelligenten Log Management von SCHOELLER NETWORK CONTROL



# LAMaLearner und MAD2

- Geschriebene Logs können mittels MAD2 und LAMaLearner ausgewertet werden.
- Bestimmte gefundene Labels können bestimmte Handlungen implizieren.
- Diese können Administratoren vorgeschlagen werden.
- Darauf aufbauende SIEM Features
  - Visualisierung Event Label Häufigkeiten pro Zeitintervall
  - Alarmierungsregeln
  - Abweichungsberechnung (Unterschiede zu „normalen“ Tagen)



# Haben Log Management Lösungen nicht schon AI?

- Existierende AI Features verarbeiten aggregierte Time-Series Daten.
  - Input: Kennlinien über Zeit
  - Output: Wo die Kennlinien nicht normal aussehen
- Das bedingt, dass man die Kennlinien hat
- Dafür muss richtiges Pattern Matching stattfinden.
- LAMaLearner ersetzt die manuelle Pattern Definition.





# Beispiel Anwendungsfälle LAMaLearner

# Alarmieren, wenn Fehler im Windows Event Log auftauchen

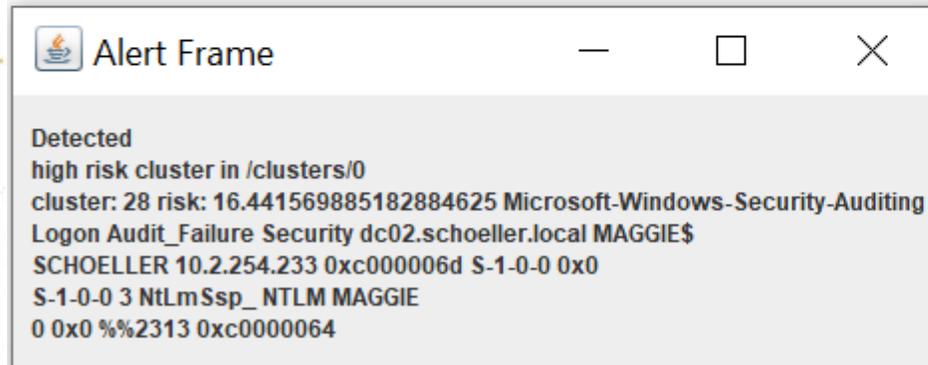
## Available clusters:

- Id: 28: 4625 Microsoft-Windows-Security-Auditing Logon Audit\_Failure Security dc02.schoeller.local MAGGIES  
SCHOELLER 10.2.254.233 0xc000006d S-1-0-0 0x0 S-1-0-0 3 NtLmSsp\_ NTLM MAGGIE 0 0x0 %%2313 0xc0000064  
Total Risk Score: 16.44156988518288 [Events](#) Labels:

failure  

security  

success ▾ add



# Herausfinden, was über die FW läuft

- Unsupervised clustering
- Labels aus Clustern ablesen:
  - 14 Cluster statt 50 Meldungen
  - Zum Beispiel:
    - https
    - Monitoring (snmp / check\_mk)
    - VPN
    - Domain Controller
  - Model für Label Identifikation

Labels:

## Available clusters:

- Id: 0: QoS Firewall goliath1 Accept gca-app-dev-a01.dev0.ontec.at (172.27.30.41) docker-mgmt-01.olymp.ontec.at (192.168.100.150) https (TCP/443) https Traffic Accepted from 172.27.30.41 to 192.168.100.150 aresmgr (192.168.105.13) inbound bond2.344 false false 443 TCP (6) Internal Internal https 2019-05-22T09:36:18Z Connection goliath-policy aresmgr {1A6AE460-5890-8D44-AC5D-03984E46FDC1} Today 11:33:35 AMAccess 0 97 2003b088-4781-449a-a609-e6a5c95ea125 goliath-policy Security Total Risk Score: 0.6567741540852687 [Events](#) Labels:

https  

- Id: 1: VPN goliath1 Encrypt monitor.olymp.ontec.at (192.168.100.50) 172.28.2.124 snmp-read (UDP/161) Encrypted in community vpn\_semantic\_rz aresmgr (192.168.105.13) inbound bond2.100 false false 161 UDP (17) Internal External snmp-read 0 71 1 2019-05-22T09:36:18Z Connection goliath-policy aresmgr {1A6AE460-5890-8D44-AC5D-03984E46FDC1} Today 11:33:35 AMAccess 0 Monitor to all Customers 635 ad97f236-8615-4c01-9648-21501f66b4da goliath-policy Security goliath.olymp.ontec.at (192.168.100.10) IKE ESP: AES-256 + SHA1 + PFS (group 2) semantic\_rz (188.172.243.157) vpn\_semantic\_rz VPN Total Risk Score: 1.139735130647125 [Events](#) Labels:

vpn  

- Id: 2: VPN goliath1 Encrypt monitor.olymp.ontec.at (192.168.100.50) 172.28.2.127 snmp-read (UDP/161) Encrypted in community vpn\_semantic\_rz aresmgr (192.168.105.13) inbound bond2.100 false false 161 UDP (17) Internal External snmp-read 0 71 1 2019-05-22T09:36:18Z Connection goliath-policy aresmgr {1A6AE460-5890-8D44-AC5D-03984E46FDC1} Today 11:33:35 AMAccess 0 Monitor to all Customers 635 ad97f236-8615-4c01-9648-21501f66b4da goliath-policy Security goliath.olymp.ontec.at (192.168.100.10) IKE ESP: AES-256 + SHA1 + PFS (group 2) semantic\_rz (188.172.243.157) vpn\_semantic\_rz VPN Total Risk Score: 0.5327207724940327 [Events](#) Labels:

monitoring  

- Id: 3: QoS Firewall goliath1 Accept docker-ontec-dev-01.docker.loc (172.52.10.20) cu-services02.dmz3.ontec.at (192.168.123.19) domain-udp (UDP/53) domain-udp Traffic Accepted from 172.52.10.20 to 192.168.123.19 aresmgr (192.168.105.13) inbound bond2.220 false false 53 UDP (17) Internal Internal domain-udp 2019-05-22T09:36:18Z Connection goliath-policy aresmgr {1A6AE460-5890-8D44-AC5D-03984E46FDC1} Today 11:33:35 AMAccess 0 96 c922e8eb-96aa-42b3-8fb0-a38a9944d4b7 goliath-policy Security Total Risk Score: 0.8118063355965119 [Events](#) Labels:

domain controller  

# Nach Updates richtig labeln

- Version X loggt in spezifischem Format.
- Daraus kann LAMaLearner ein Modell generieren das Fehler von Normal unterscheidet.
- Version X+1 loggt plötzlich ganz anders.
- LAMaLearner Modell funktioniert immer noch, wo Parser Regeln versagen.

```
▼ events:  
  ▼ 0:  
    id: 0  
    time: ""  
    message: "Component router 1 with IP 192.168.13.1 is down 0"  
  ▼ 1:  
    id: 1  
    time: ""  
    message: "Component router 1 with IP 192.168.13.1 is working correctly 7"  
  ▼ 2:  
    id: 2  
    time: ""  
    message: "Component router 1 with IP 192.168.13.1 is >79% traffic capacity 10"  
▼ events:  
  ▼ 0:  
    id: 0  
    time: ""  
    message: "System 192.168.13.1 (router 1) down 0"  
  ▼ 1:  
    id: 1  
    time: ""  
    message: "System 192.168.13.1 (router 1) working correctly 7"  
  ▼ 2:  
    id: 2  
    time: ""  
    message: "System 192.168.13.1 (router 1) >79% traffic capacity 10"
```

# Neue Logformate verstehen

## Available clusters:

- Id: 0: ::1 - - "GET /server-status?auto HTTP/1.1" 200 441 "-" "Python-urllib/2.7" Total Risk Score: 0 [Events](#) Labels:
- Id: 1: 127.0.0.1 - - "GET / HTTP/1.1" 403 202 "-" "Mozilla/5.0 (ISPCconfig monitor)" Total Risk Score: 0 [Events](#) Labels:
- Id: 2: ::1 - - "GET /server-status?auto HTTP/1.1" 200 439 "-" "Python-urllib/2.7" Total Risk Score: 0 [Events](#) Labels:
- Id: 3: ::1 - - "GET /server-status?auto HTTP/1.1" 200 440 "-" "Python-urllib/2.7" Total Risk Score: 0 [Events](#) Labels:
- Id: 4: ::1 - - "GET /server-status?auto HTTP/1.1" 200 437 "-" "Python-urllib/2.7" Total Risk Score: 0 [Events](#) Labels:
- Id: 5: ::1 - - "GET /server-status?auto HTTP/1.1" 200 438 "-" "Python-urllib/2.7" Total Risk Score: 0 [Events](#) Labels:
- Id: 6: ::1 - - "GET /server-status?auto HTTP/1.1" 200 436 "-" "Python-urllib/2.7" Total Risk Score: 0 [Events](#) Labels:

# Aufgaben Mischen

- Log Typ
- Nominal & Failure

- Id: 54: 4625 Microsoft-Windows-Security-Auditing Logon Audit **Failure** Security dc02.schoeller.local MG MG-NB 10.2.200.54 0xc000006d S-1-0-0 0x0 S-1-0-0 3 NtLmSsp\_ NTLM MG-NB 0 0x0 %%2313 0xc000006a Total Risk Score: 0.532842053868346 [Events](#) Labels:

Windows Security Log  

failure  

FW Log

- Id: 55: 4769 Microsoft-Windows-Security-Auditing Kerberos\_Service\_Ticket\_Operations Audit\_Success\_Security dc02.schoeller.local bb@SCHOELLER.LOCAL SCHOELLER.LOCAL DC02\$ S-1-5-21-831815799-2714527801-2835520152-1108 0x40810000 0x12 ::ffff:10.2.200.49 0x0 Total Risk Score: 0.08708915947350236 [Events](#) Labels:

Windows Security Log  

nominal  

FW Log

- Id: 99: 4648 Microsoft-Windows-Security-Auditing Logon Audit\_Success\_Security dc02.schoeller.local tecadmin SCHOELLER 10.2.100.124 22249 {00000000-0000-0000-0000-000000000000} S-1-5-18 DC02\$ SCHOELLER 0x3e7 0x200 C:\Windows\System32\lsass.exe {00000000-0000-0000-0000-000000000000} localhost localhost Total Risk Score: 0.06499146357122207 [Events](#) Labels:

Windows Security Log  

nominal  

FW Log

- Id: 102: 127.0.0.1 - - "GET / HTTP/1.1" 403 202 "-" "Mozilla/5.0 (ISPCconfig monitor)" Total Risk Score: 0.06698409963412978 [Events](#) Labels:

Access Log  

nominal  

Windows Security Log

# “Eigentlich Egal“ Events identifizieren

- Häufig stehen in Log Files viele irrelevante Meldungen die nicht über einzelne Felder voneinander unterschieden werden können
- LAMaLearner kann lernen relevant von irrelevant zu unterscheiden.

# Denglische Meldungen verstehen

- Für LAMaLearner ist die natürliche Sprache der Log Meldungen irrelevant.
- Solange Cluster manuell mit Labels annotiert werden, kann LAMaLearner lernen Events Labels zuzuordnen.

# Label Typen pro Zeitintervall aggregieren

- MAD2 überträgt Events über definierte Zeiträume
- LAMaLearner annotiert Labels für den jeweiligen Zeitraum
- MAD2 erzeugt Reports über die Label Häufigkeiten pro Zeitraum.  
Z.B.:
  - Anzahl fehlgeschlagener Log Ins
  - Anzahl Firewall Rejections
  - Anzahl erkannte Malware
  - ...
- Durch Machine Learning basiertes Modell können Labels auch in veränderten Log Formaten erkannt werden.
- Basierend auf erkannten Labels pro Event oder Anzahl Label pro Zeitraum können Alarmierungen durchgeführt werden
- Auf Aggregationen können Trendprognosen berechnet und Anomalien festgestellt werden.

# Log Typen auseinanderhalten

- 47 Cluster statt 18110 Events (Aufmerksamkeitsschwelle 5)
- 56 Cluster statt 18110 Events (Aufmerksamkeitsschwelle 2)
- Model für Cluster mit Aufmerksamkeitsschwelle 2:
  - 4 FP/FN
  - 52 TP
  - 93% F1

Labels:  Windows Security Log  Access Log  FW Log

Teach Lama

## Available clusters:

- Id: 0: 4624 Microsoft-Windows-Security-Auditing Logon Audit\_Success\_Security bi01.schoeller.local S-1-5-18 BI01\$ SCHOELLER 0x3e7 S-1-5-18 SYSTEM NT-AUTORITÄT 0x3e7 5 Advapi\_Negotiate {00000000-0000-0000-0000-000000000000} 0 0x200 C:\Windows\System32\services.exe %%1833 Total Risk Score: 0.8111982989657337 [Events](#)  
Labels:  
 Windows Security Log    
 Access Log  add
- Id: 1: 4672 Microsoft-Windows-Security-Auditing Special\_Logon Audit\_Success\_Security bi01.schoeller.local S-1-5-18 SYSTEM NT-AUTORITÄT 0x3e7 SeAssignPrimaryTokenPrivilege\_SeTcbPrivilege\_SeSecurityPrivilege\_SeTakeOwnershipPrivilege\_SeLoadDriverPrivilege\_S Total Risk Score: 0.8503780783447816 [Events](#) Labels:  
 Windows Security Log    
 Access Log  add
- Id: 2: 4672 Microsoft-Windows-Security-Auditing Special\_Logon Audit\_Success\_Security dc02.schoeller.local S-1-5-18 DC02\$ SCHOELLER SeSecurityPrivilege\_SeBackupPrivilege\_SeRestorePrivilege\_SeTakeOwnershipPrivilege\_SeDebugPrivilege\_SeSystemEnviro Total Risk Score: 0.8550176628505113 [Events](#) Labels:  
 Windows Security Log    
 Access Log  add

# LAMaLearner

Tobias Eljasik-Swoboda



**SCHOELLER  
ONTEC**

SCHOELLER NETWORK CONTROL GmbH, ONTEC AG  
Ernst-Melchior-Gasse 24/DG, A-1020 Wien, Austria

Tel.: +43 1 20 55 20-0, Fax: +43 1 20 55 20-20  
[office@ontec.at](mailto:office@ontec.at), [www.ontec.at](http://www.ontec.at)