



Malware-Erkennung 2015

Thomas Bottesch
Avira Operations GmbH
02.06.2015

Inhaltsverzeichnis

- Künstliche Intelligenz in der Malware-Erkennung
 - Wie baue ich eine künstliche Intelligenz?
 - Funktion der KI bei Avira
 - Ausblick



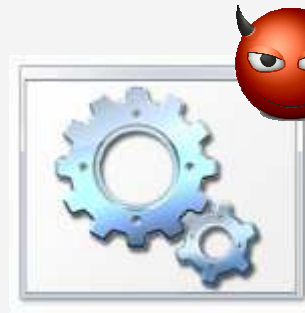
Wie baue ich eine KI, die Malware erkennt?



Word.exe
(Size: 0.8MB)



Acrobat.exe
(Size: 0.3MB)



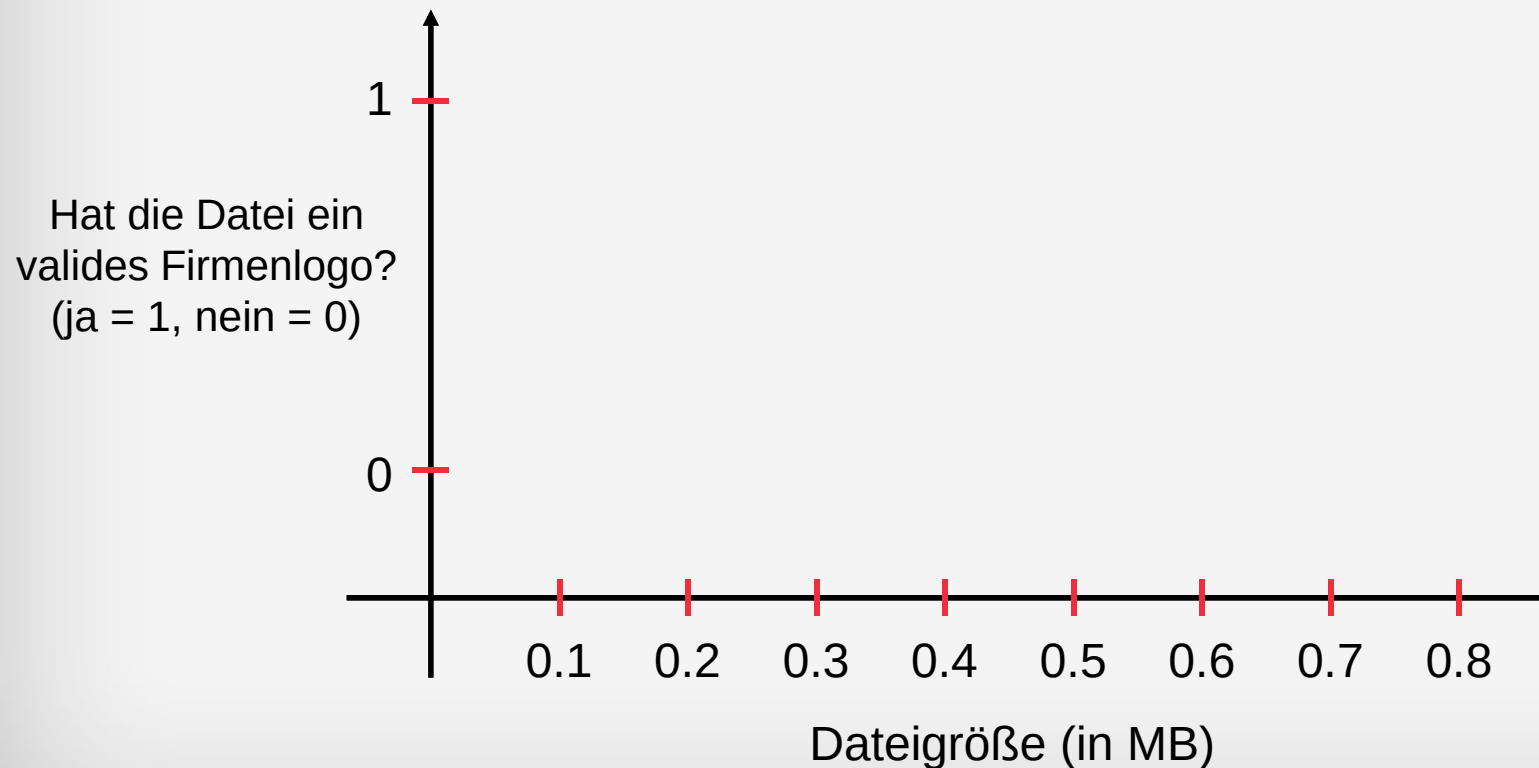
Virus.exe
(Size: 0.2MB)



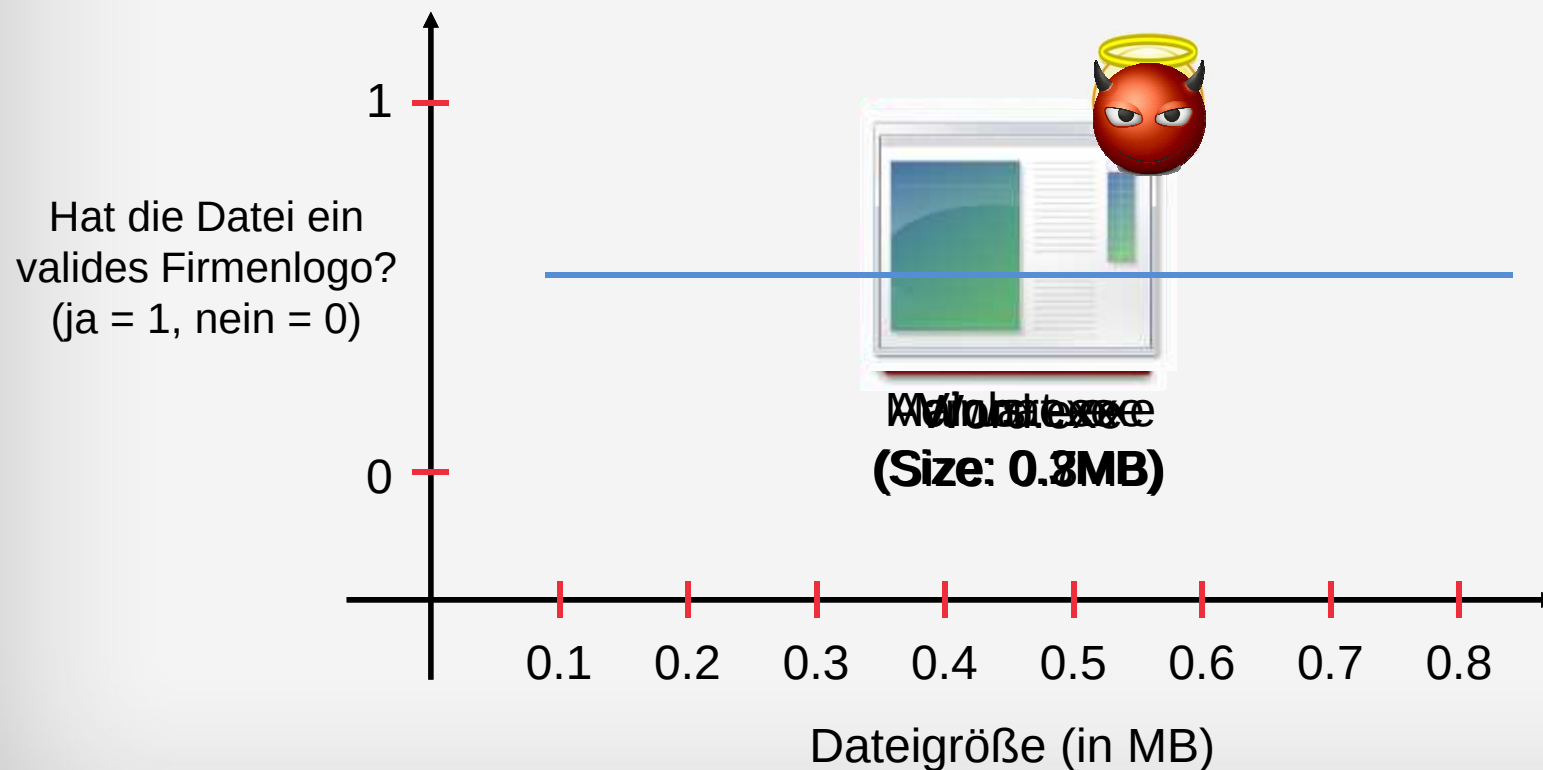
Malware.exe
(Size: 0.7MB)



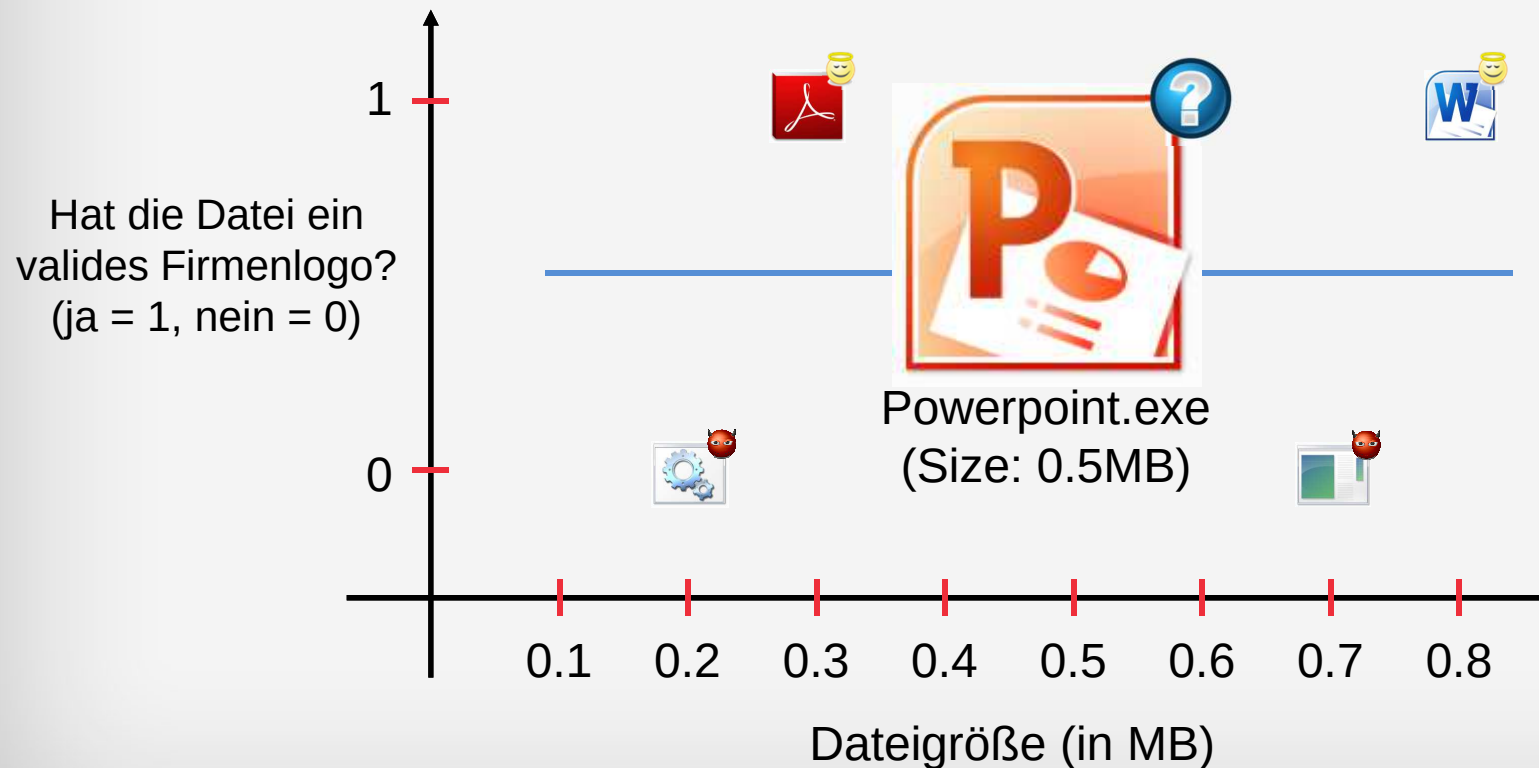
Wie baue ich eine KI, die Malware erkennt?

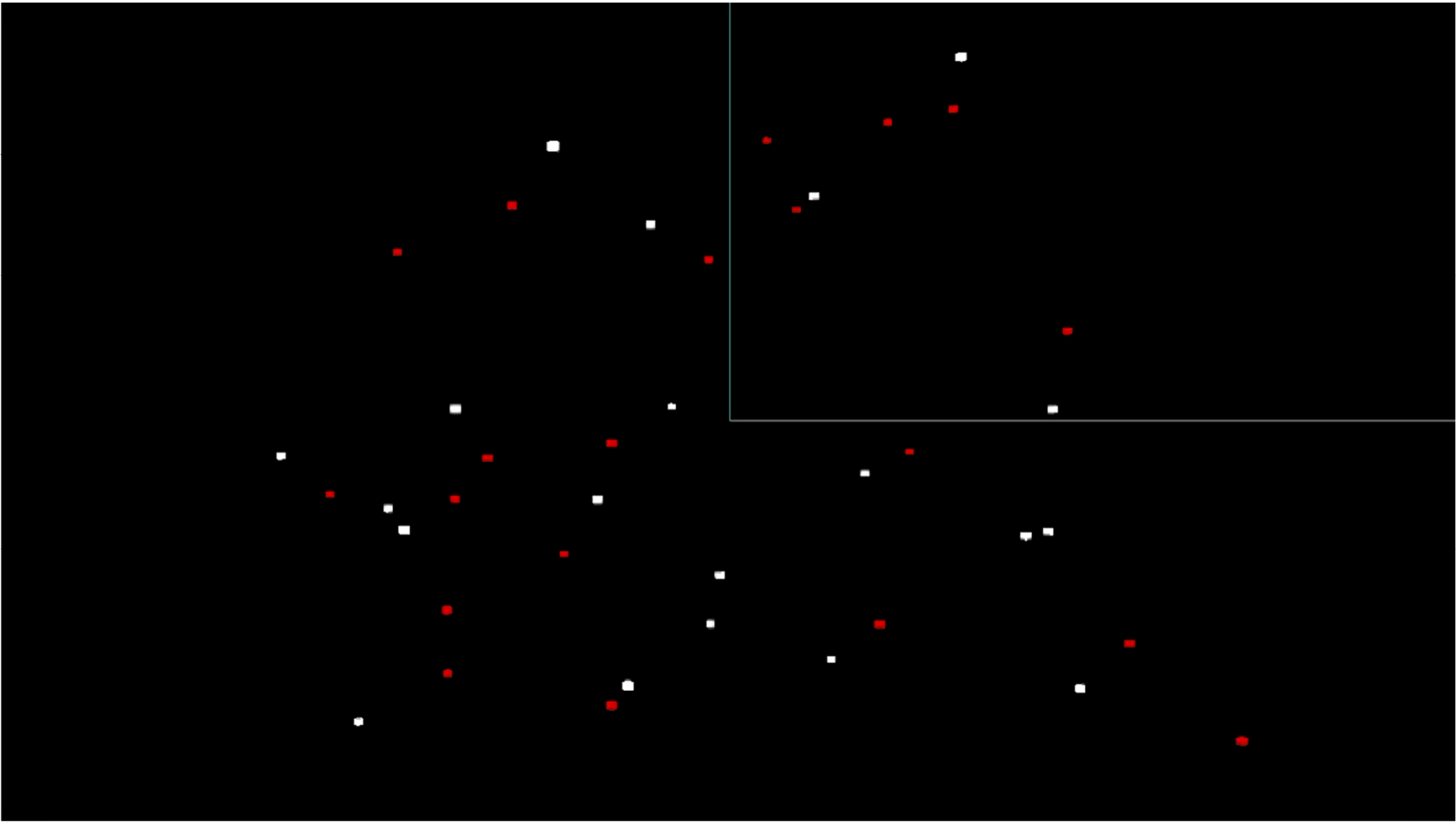


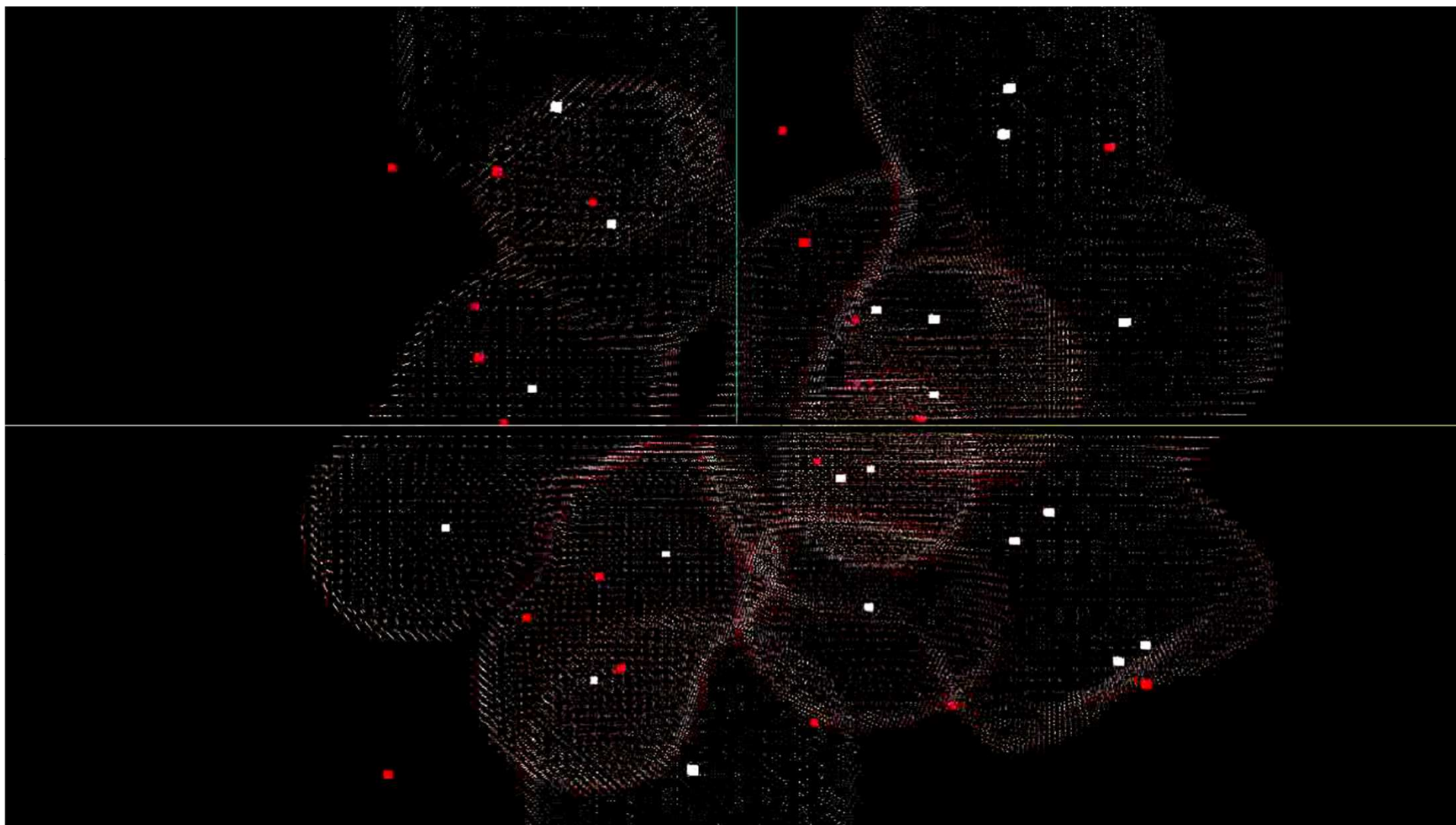
Wie baue ich eine KI, die Malware erkennt?



Wie baue ich eine KI, die Malware erkennt?

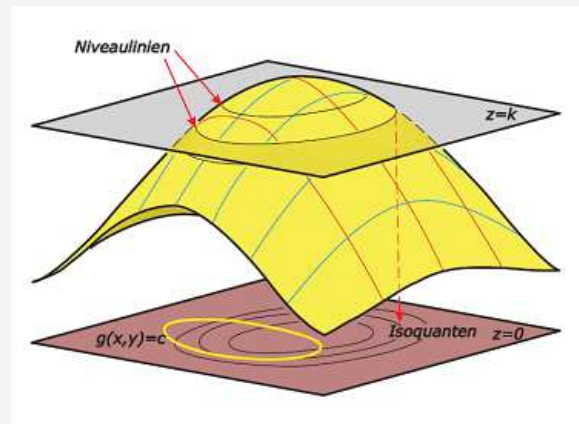






Komplexität die Hyperebene zu finden

$$\min_{\mathbf{w}, \xi, b} \max_{\alpha, \beta} \left\{ \frac{1}{2} \|\mathbf{w}\|^2 - C \sum_{i=1}^n \xi_i - \sum_{i=1}^n \alpha_i [y_i (\mathbf{w} \cdot \mathbf{x}_i - b) - 1 + \xi_i] - \sum_{i=1}^n \beta_i \xi_i \right\}$$



$$f(\mathbf{x}) = \text{sgn}(\langle \mathbf{w}, \mathbf{x} \rangle + b) = \text{sgn} \left(\sum_{i=1}^m \alpha_i y_i \langle \mathbf{x}_i, \mathbf{x} \rangle + b \right)$$



Komplexität die Hyperebene zu finden

- > 7000 Dimensionen (Im Beispiel waren es zwei)
- > 100 Millionen Dateien (Im Beispiel waren es vier)

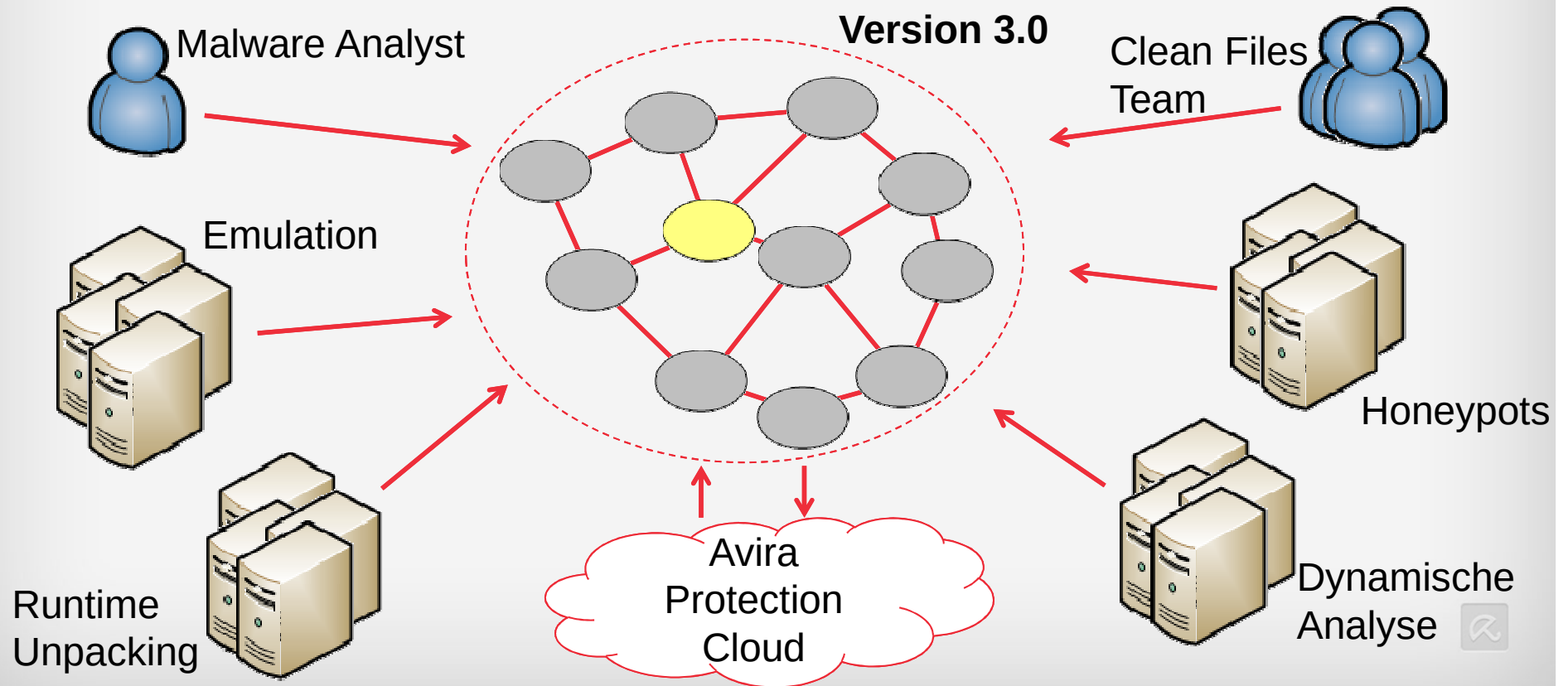
Berechnungszeit der KI (Version 1.0)	
Mit Open-Source Tools	Mit hochoptimiertem CUDA Code
1 PC, mehrere Jahre	16 Server, 1 Tag



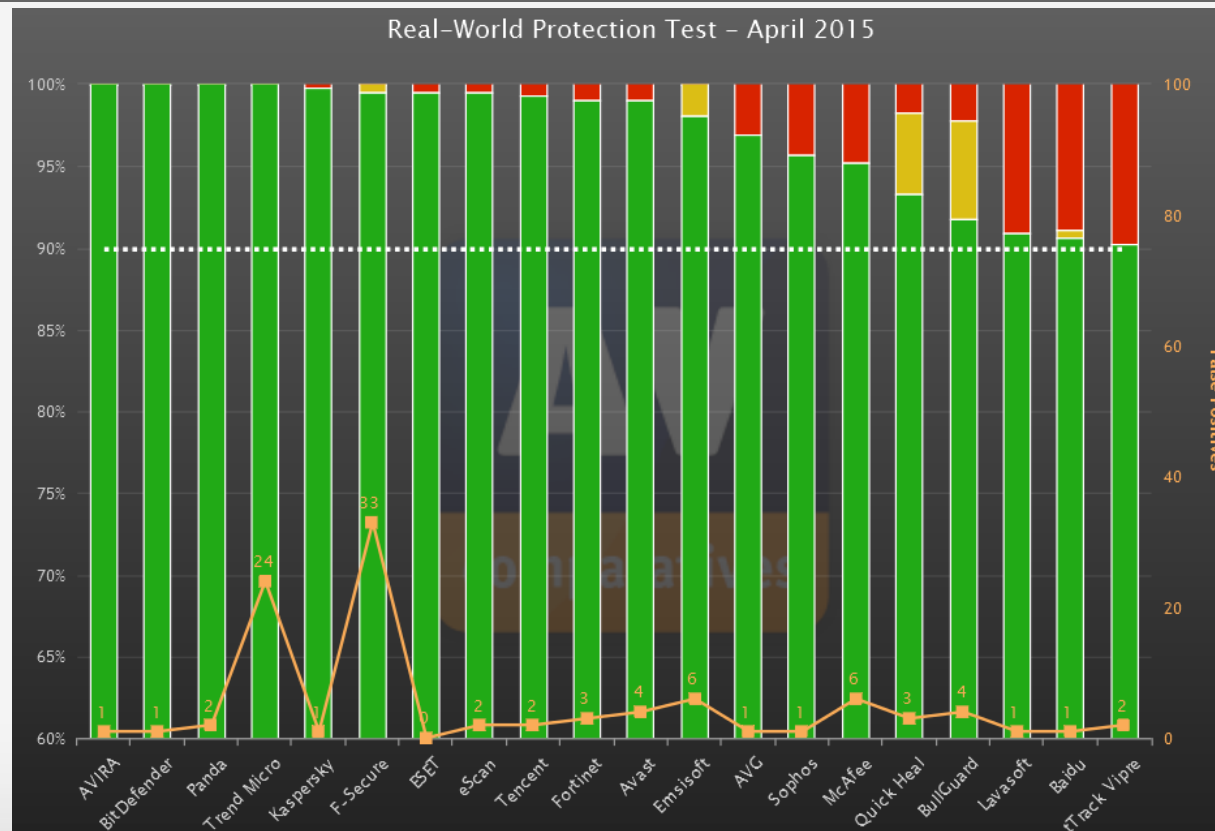
Avira Protection Cloud



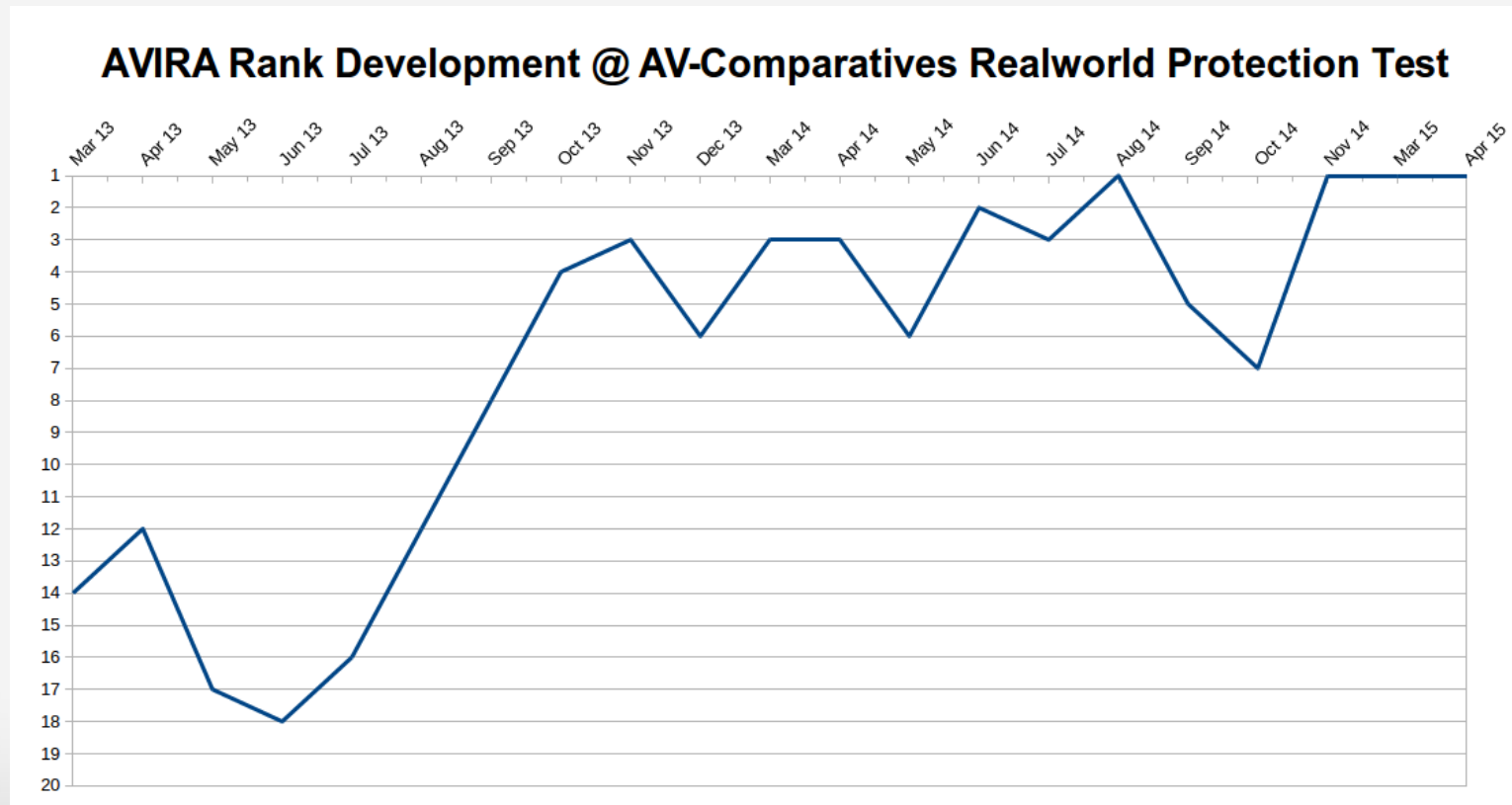
Die Rolle der KI innerhalb von Avira



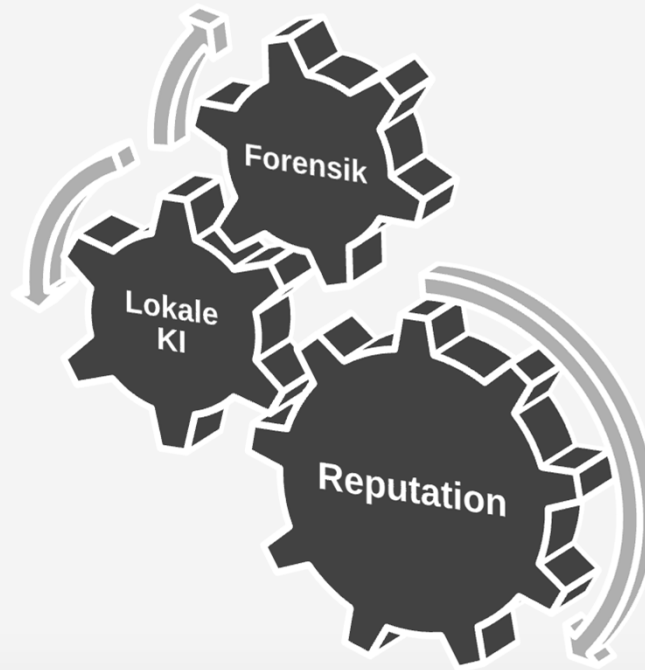
Evaluation der Cloud und der KI



Evaluation der Cloud und der KI



Ausblick





Statistiken: Blockierte Bedrohungen

Blockierte Bedrohungen
(Millionen / Tag)
Windows

