

Revolution innerhalb der Revolution

Der Aufstieg der nicht überwachten Lernverfahren

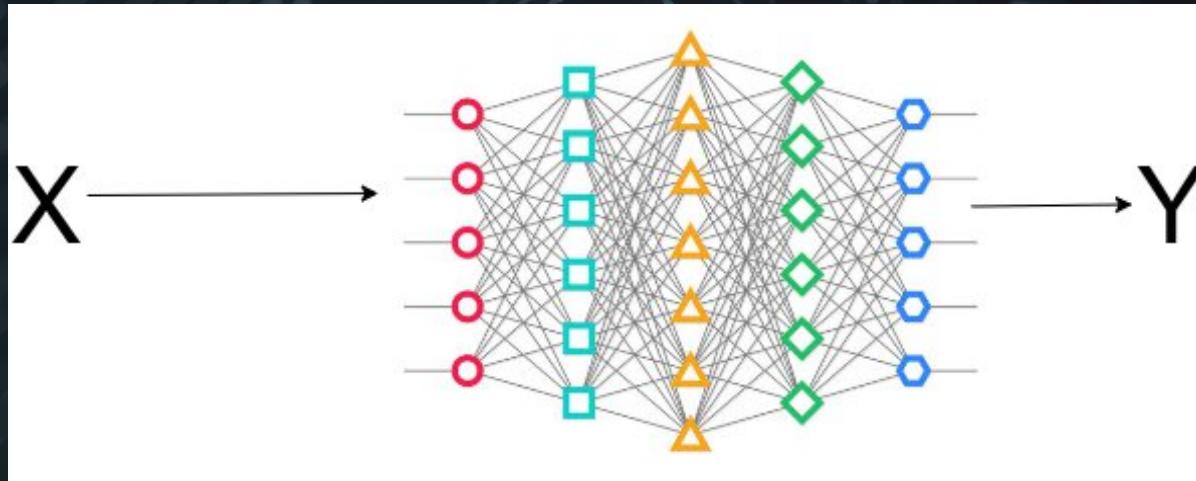
Michael Kopp



institute of advanced
research in artificial
intelligence

Welches "Skalierungsproblem"?

- Überwachte Lernverfahren dominieren
- Sie benoetigen "Labels", also das Ziel das erreicht werden soll.



- Die meisten "deep learning" Methoden basierten bis vor kurzem auf dieser Art des Lernens (und Variationen davon ...).

Na und?

Beispiel: Ampel erkennen bei "Selbstfahrenden Autos"

- Leicht einen Prototypen mit Labels in einer Stadt zu bauen.
- Aber mit unterschiedlichen Lichtverhaeltnissen ...
- Unterschiedlichen Ampeln ...
- Unterschiedlichen Umgebungen fuer eine Ampel ...
- ... in 180 Laendern oder so ...
- ... und man kann nicht mit 99% accuracy zufrieden sein ...
- ... kein Ende der Annotation!



Fazit

- Echt-welt Probleme haben statistisch gesehen eine "Fettschwanzverteilung", d.h. in der Regel sieht man nur Ausnahmen!
- Man wird nicht aus dem Annotieren herauskommen

So skaliert das nicht.

A new sheriff in town ...

Koennen Maschinen nicht selbst ohne "Labels" (also Zielvorgaben) lernen ?

2020/2021 - Dieser Traum scheint wahr zu werden!

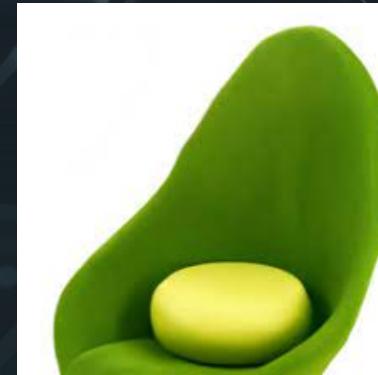
Buzzwords:

SimCLR, OpenAI/CLIP, contrastive learning, semi-supervised learning, CURL, ...

A flavour of examples ...

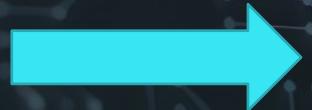
Auf Internetz-daten skaliert trainiert - unter Ausnutzung der Tatsache, dass mehrere Datenquellen fuer das selbe Objekt vorhanden sind (Text und Bild) ... fuehrt zu

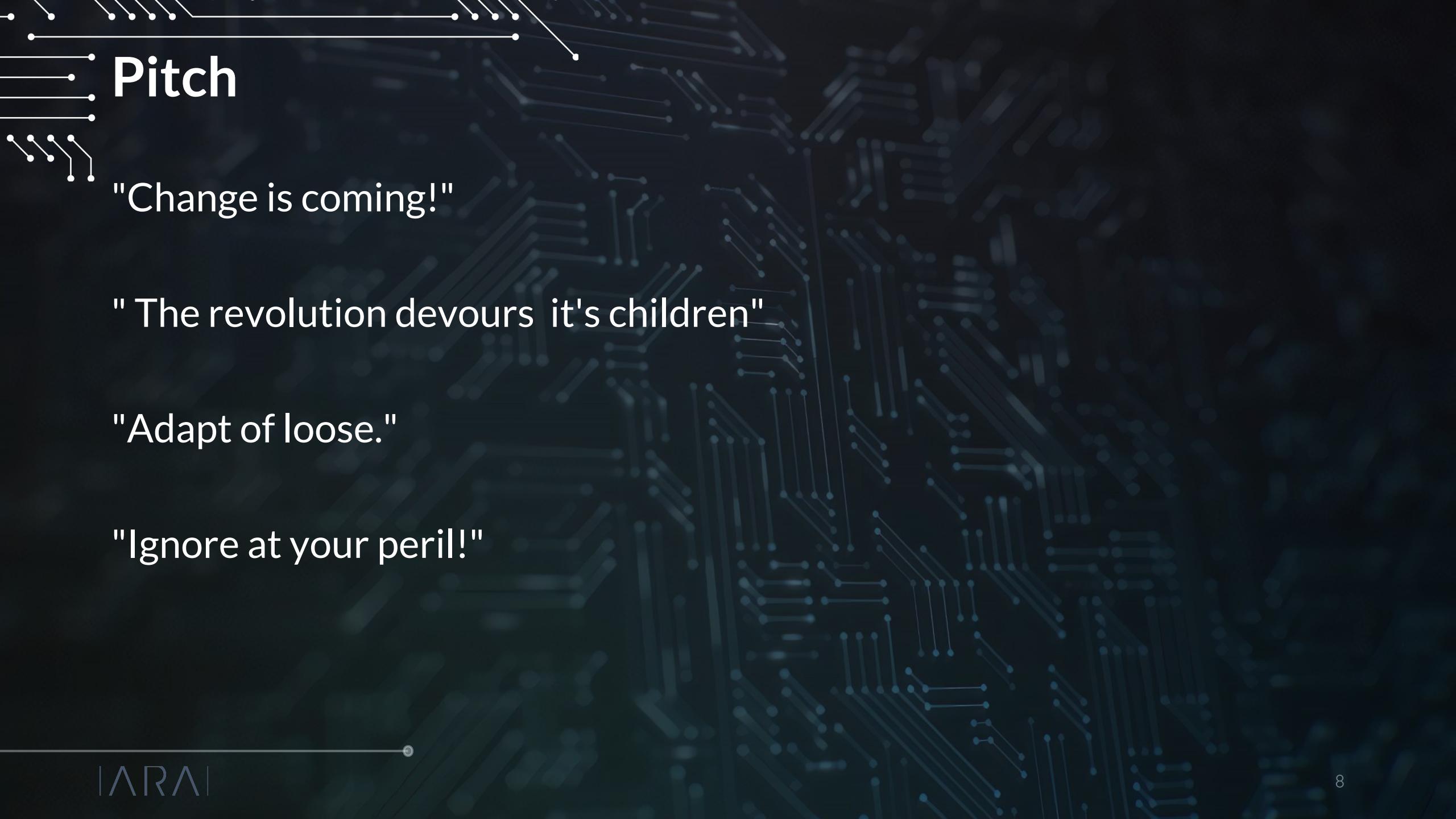
- "chair like an avocado"
- Ohne die Labels zu kennen bessere Klassifikation auf bekannten Datensaetzen (zero-shot learning)!



A flavour of examples ...

"painting of a city built on top of a giant turtle walking slowly towards the viewer with a clear blue sky and a lush green landscape"



A dark blue background featuring a dense, abstract circuit board pattern of light blue lines and dots.

Pitch

"Change is coming!"

"The revolution devours its children"

"Adapt or loose."

"Ignore at your peril!"