

Sherlock Holmes im Orbit

Künstliche Intelligenz erkennt
Biotope auf Satellitenbildern,
unterstützt durch
Synthetic Data Sets

Download dieses Slide Decks:

DE: sherlock-ki.climatehackerz.com

EN: sherlock-ai.climatehackerz.com

Jörg Osarek, May 2023

m:contact at climatehackerz.com

T: +49-151-23-0-24-333

CLIMATE HACKERZ

2023 - 2035



KEEP
CALM
AND
TRANS-
FORM

www.ClimateHackerz.com



The video, that watches you.
analysis of facial expressions and your mood
with a simple webcam → Emotion Database with API

Affdex Your Vote View Results

Expressiveness

Age Bands

You Your Curve

All

5

40 © Skilltower Institute, 2016

Level 5: fa
expression

Emotion A
with PC o
www.affectiv

Zur Person Jörg Osarek

- Homburger Bub, geboren 1970
- Industriekaufmann (IHK) 1991
Hewlett-Packard
- **IT-Consulting** seit 1992
- Selbständiger Unternehmensberater seit
2003,
- **Mitgründer its-people** Verbund,
4 Jahre Geschäftsführer its-people
Hochtaunus GmbH
- Verlag Gordon's Arcade seit 2012
- <http://skilltower.com/publications.html>
(60+ Einträge)
- YAM-Track-Record seit den 1980ern*
- **Gründer ClimateHackerz.com 2019**



DIGILITY KEYNOTE

VIRTUAL REALITY ANALYTICS:
HOW VR AND AR CHANGE
BUSINESS INTELLIGENCE

JÖRG OSAREK
CEO, SKILLTOWER INSTITUTE
PARTNER, ITS-PEOPLE



Zur Person Jörg Osarek: Publizierte Bücher

2023: Work in Progress



Reisefuehrer.ClimateHackerz.com

2008



2016



2012



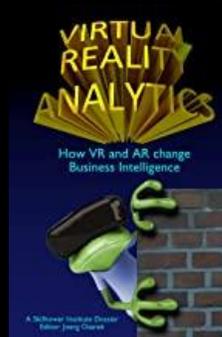
2014



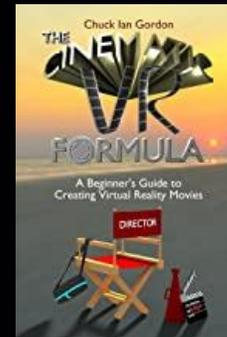
2012



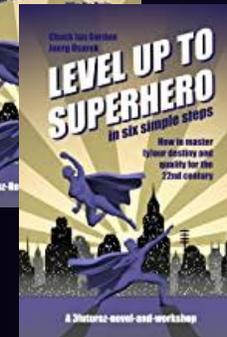
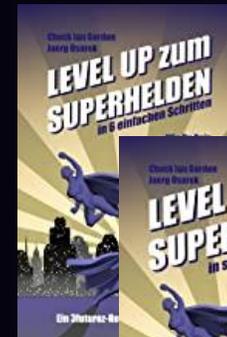
2016



2017



2020



The Climate Transformation (Travel) Guide: Steps



Baseline – where are we now?



Your new Cathedral:
VMAP + CLI + CEC

Kritische
Erfolgsfaktoren

- 1=VMAP-
New Cathedral Plan
- 2=Finance
- 3=Neutralize
Slow Downs
- 4=Assets
- 5=(Exp.) Growth
Strategy

Travel Preparations



Your Journey:
Navigating the transformation



Workouts (Habits)



Steps Beyond



Die Klimatransformation ist nur eine der aktuellen Prüfungen unserer Zivilisation

Transformation der Kultur > Erreichen der Balance von Macht und Verantwortung durch Unabhängigkeit und Eigentum an der neuen Welt permanenter Nachhaltigkeit (Emissionen und Kreislaufwirtschaft unter Kontrolle einer verantwortungsvollen Zivilisation)



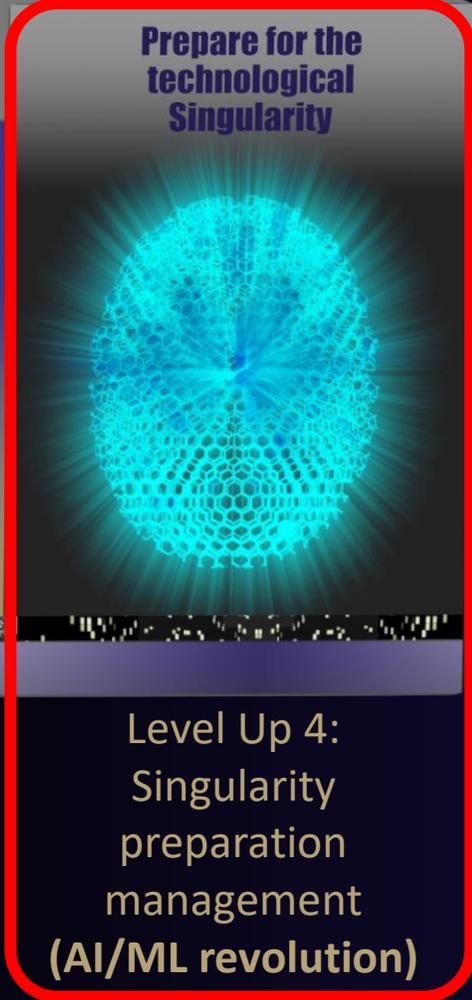
Level Up 1:
Empower
our character
and orgs



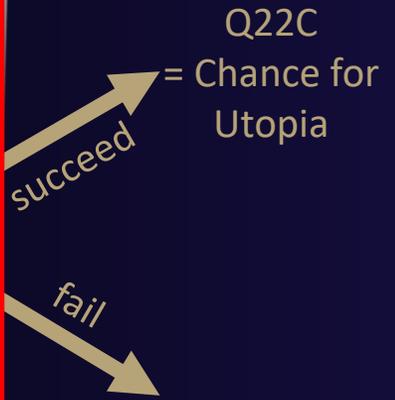
Level Up 2:
Solve climate crisis
with Doughnut
economics (by
Kate Raworth)



Level Up 3:
Bring balance to
our civilization



Level Up 4:
Singularity
preparation
management
(AI/ML revolution)



www.Q22Century.org



Sherlock Holmes im Orbit

Künstliche Intelligenz erkennt Biotope auf Satellitenbildern,
unterstützt durch Synthetic Data Sets

zu einem feierlichen Anlass wie diesem
habe ich mir gedacht,
ich trage das auch etwas feierlich vor.



Der Anstoß



Ob CSRD, Green Deal, entwaldungsfreie Lieferketten
weil wir Nachhaltigkeit gern ohne Abholzung hätten.

Millionenfach muss man bald mit geschulten Blicken
Biotope kartieren und entwaldete Lücken
auf Satellitenbildern und von Daten mit Geobezug
Doch Kartierer sind teuer und es gibt nicht genug.

Daher ist künstliche Intelligenz oder kurz KI
zur Massenauswertung gefragt, wie noch nie.





Die AUFGABE

Die Aufgabe war einfach zu benennen.
Ein KI-System soll Biotope automatisch erkennen.
Wie Sherlock Holmes, der im Orbit kreist
und Biotope metergenau ausweist.

Damit wir wissen, ob und wie das geht
wo Wissenschaft und Technik heute steht
und auch, wie man dieses System skaliert
haben wir einen Proof of Concept durchgeführt.

Was wir gelernt beim POC, will ich mich eilen
mit Freude hier mit Euch zu teilen.





POC - Was haben wir gelernt

Frei nach Hadden's Regel für Investition: Warum nur einen Beweis wenn Du zwei kannst bekommen für den doppelten Preis?

Zwei Anbieter aus Forschung und Wirtschaft wir wählten.

Mit bohrenden Fragen wir beide quälten.

Challenge gemischt mit gegenseitiger Befruchtung war das Resultat.

Wir lernten beide gut kennen und erhielten Klarheit für den weiteren Pfad.

befruchtende zwei Anbieter Challenge (Coopetition)



POC - Was haben wir gelernt



Mit der KI experimentieren, selbst Hand anzulegen war enorm wichtig und für unser Team ein Segen.

So wurde daraus ein Lern- und Verständnisprojekt, bessere Ergebnisse spielerisch gemeinsam entdeckt.



POC - Was haben wir gelernt



Bild rein, Antwort raus, war der Wunsch,
sehr schlicht
doch zeigt die Praxis: so gehts leider nicht.

Kernidee: Einen offenen KI-Werkzeugkasten für mannigfaltige Aufgaben erstellen.





POC - Was haben wir gelernt

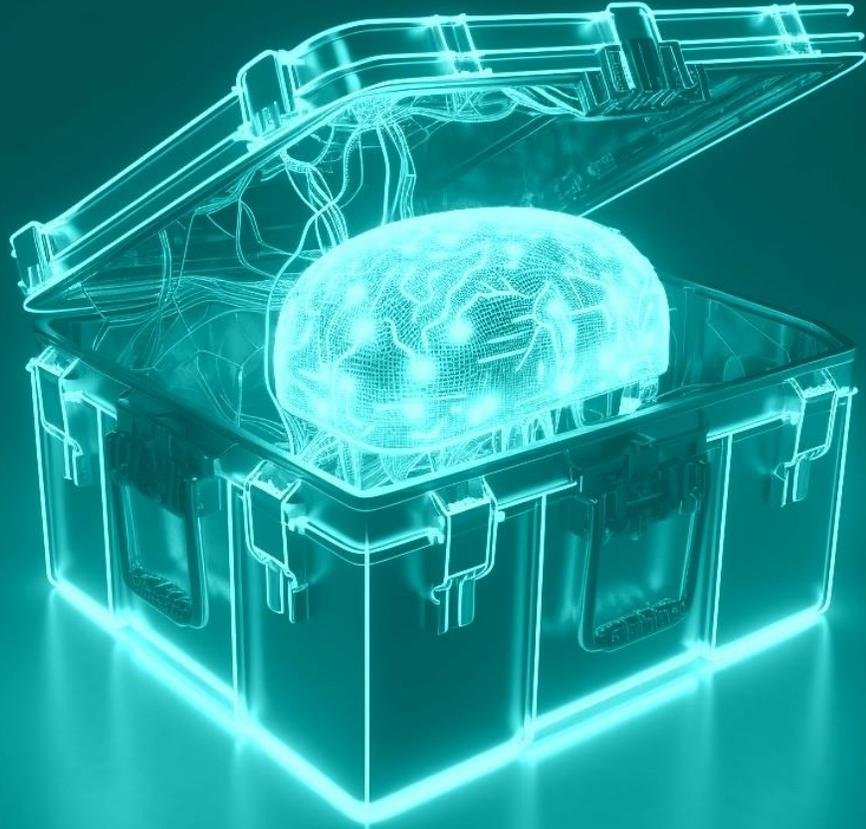
Explainable AI lässt Annahmen prüfen und Ideen lenken.
XAI ermöglicht, der KI zuzusehen beim Denken.*

Kernidee: Einen offenen KI-Werkzeugkasten für mannigfaltige Aufgaben erstellen.

* wichtig gerade jetzt in Zeiten sich nähernder Gefahr durch Krisen und die nahende Singularität: explainable AI – XAI. Zitat von - <https://www.ibm.com/watson/explainable-ai> : (Explainable artificial intelligence (XAI) is a set of processes and methods that allows human users to comprehend and trust the results and output created by machine learning algorithms. Explainable AI is used to describe an AI model, its expected impact and potential)



POC - Was haben wir gelernt



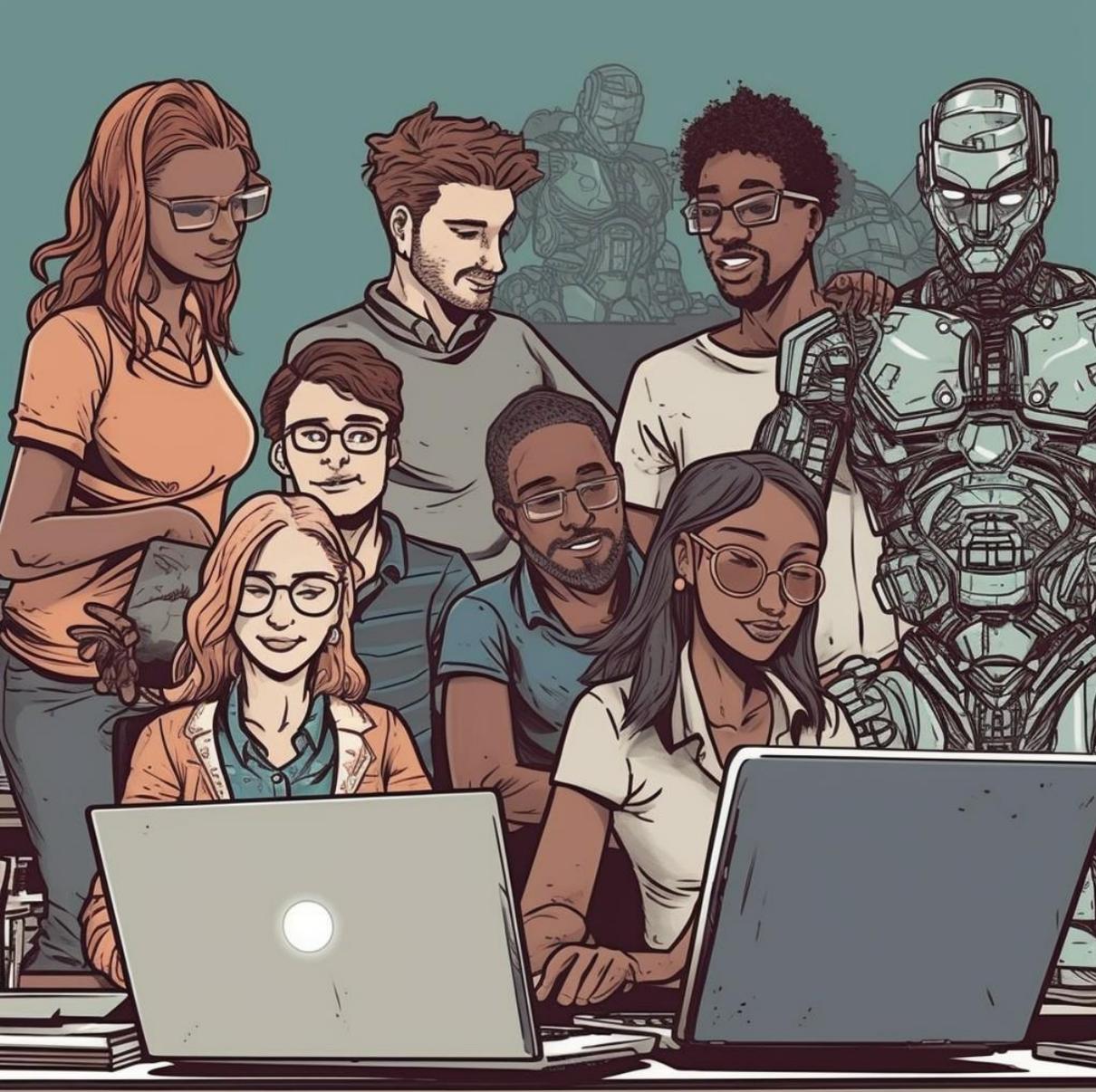
Ein spezialisiertes System: wie bauen wir das nur?
Als dynamische modulare Prozessarchitektur.

Einen offenen Werkzeugkasten für viele Szenarien
für heutige und zukünftige Auswertungs-Arien.

Denn in dieser beschleunigten, turbulenten Zeit
bietet nur Flexibilität Investitionssicherheit.

Kernidee: Einen offenen KI-Werkzeugkasten für mannigfaltige Aufgaben erstellen.





POC - Was haben wir gelernt

Am Allerwichtigsten: ein Team interdisziplinärer Enthusiasten.

Die können ihn bauen diesen magischen Machine-Learning-Kasten.

Kernidee: Einen offenen KI-Werkzeugkasten für mannigfaltige Aufgaben erstellen.



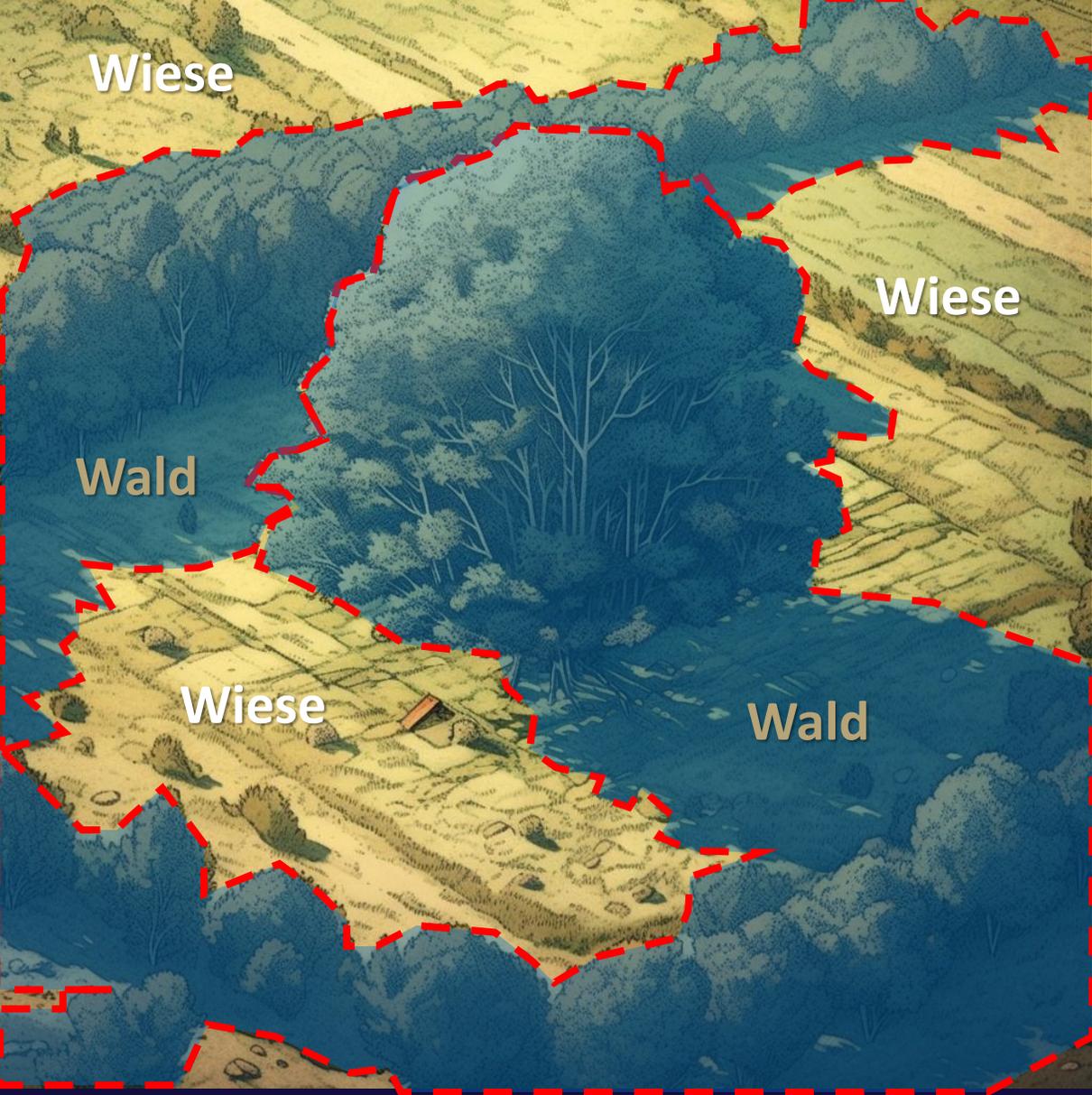


POC - Was haben wir gelernt

Auch wir sind in einige Fallen getappt.
Welche fragt Ihr und wie es dann
trotzdem klappt?

Lessons Learned: Fallstricke und Lösungsansätze





POC - Was haben wir gelernt

Für Watershed-Segmentierung war Baum gleich Schatten.

Wald wurde erkannt, obwohl wir dort Wiese hatten.

Lessons Learned: Fallstricke und Lösungsansätze





POC - Was haben wir gelernt

Obacht auch bei der Geo-Projektion:

wichtig: Koordinaten mit hoher Präzision.

Liegt ein Weg dicht an einem Fließgewässer

ist die Biotoperkennung schlecht, statt besser.

Das System lag schon mal 7 Meter daneben

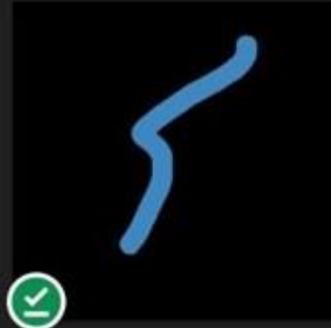
aber auf Feldwegen können keine Fische leben.

Abhilfe schafft Testmanagement-Denken und -
Handeln

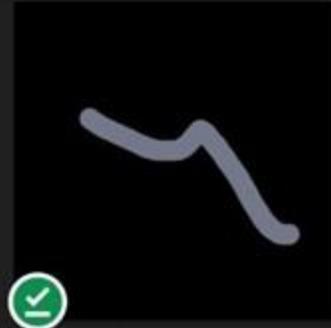
Explainable AI hilft beim entdecken, beim
Lenken und Wandeln.



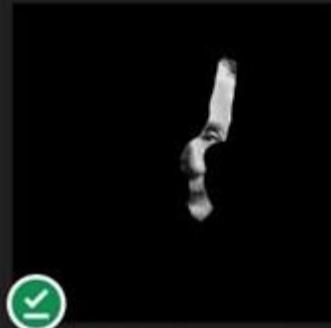
POC - Was haben wir gelernt



fake-Water-A.png



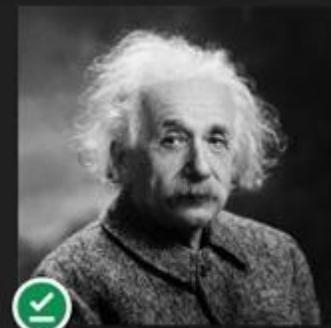
fake-Water-B.png



fake-Water-C2-Albert_Einstein_Head_cleaned
_CC0.png



fake-Water-D.png



fake-Water-C1-Albert_Einstein_Head_cleaned
_CC0.png

Einen Bias* entdeckten wir durch Hinterfragen.
Zu viele Wasserläufe kamen aus den Vorhersagen.
Egal welche Farbe, bei länglicher Form
war die Wasserlaufquote einfach enorm.
Selbst Einsteins Foto als länglicher Schnitt:
als Wasserlauf erkannt. So ein Pech.

Einsteins ganzes Foto wurde klassifiziert
als Geröllhalde - wir hatten nur Biotope programmiert.
Die KI verwechselte Noise mit Signal.
Ein Training mit "Anderem" macht das zukünftig egal.

Lessons Learned: Fallstricke und Lösungsansätze

Problem: Bias erkennt lange Formen als Wasser. Dieser ist durch veränderte Trainingsstrategie zu vermeiden, denn es können z.B. auch Wege oder Gebäude sein.

Bilder von Flüssen mit nicht ausreichender geistiger Schöpfungshöhe von Jörg Osarek

Bild von Einstein - Photograph by Oren Jack Turner, Princeton, N.J. Public domain - https://commons.wikimedia.org/wiki/File:Albert_Einstein_1947.jpg





POC - Was haben wir gelernt

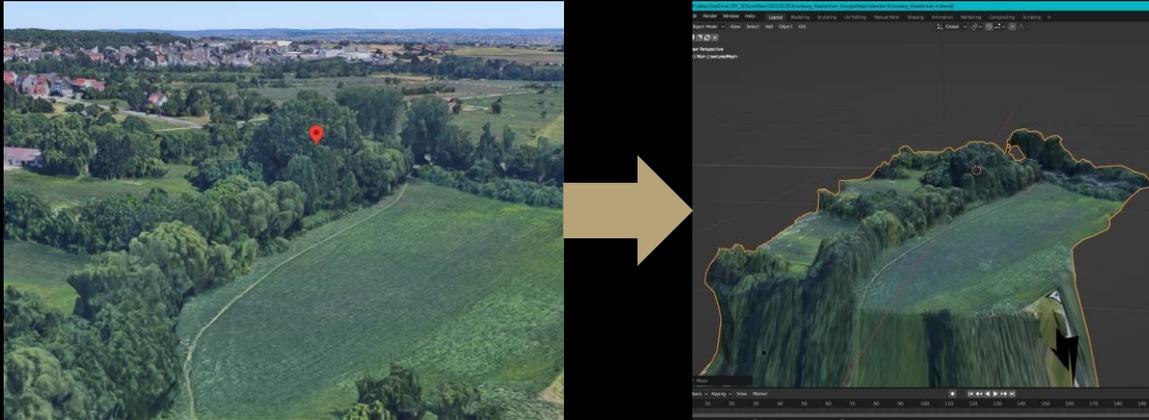
Die Biotoperkennung war teils gut, teils geraten.
Die KI hatte viel zu wenig gelabelte Trainingsdaten.

Labeln durch Menschen stellt Wirtschaftlichkeitsfragen.
Heißt das, wir sind hier schmachvoll geschlagen?

Problem: Die Trainingsdatenmenge war deutlich zu gering.

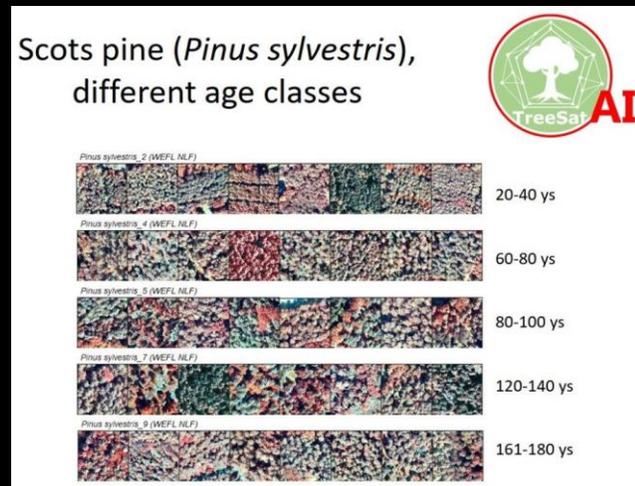


POC - Was haben wir gelernt



Für einen Teil helfen 3D Höhenmodelle*
Weitere Spektralbänder helfen an anderer Stelle.

TreeSat AI Study:
<https://www.tu.berlin/geoinformation/forschung/projekte/laufende-projekte/treesatai>
und
<https://zenodo.org/record/6598391#.ZCLBFXZByUk> – letzte Seite des PDFs:



Die TreeSatAI-Studie konnte den Nachweis liefern.
Mit Infrarot erkennt man sogar das Alter von Kiefern.

Problem: Die Trainingsdatenmenge war deutlich zu gering.

* 3D Höhenmodelle links oben Google Maps Screenshot > rechts 3D Modell in Blender aus Photogrammetrie – Wäldchen bei Kronberg Ts.

- Sat Radar Canopy height model = weitere Informationskanäle erschließen und die KI damit trainieren.

- Auch über weitere Spektren möglich. + Multiple Spektralbänder nir, infrarot, Studie TreeSat AI

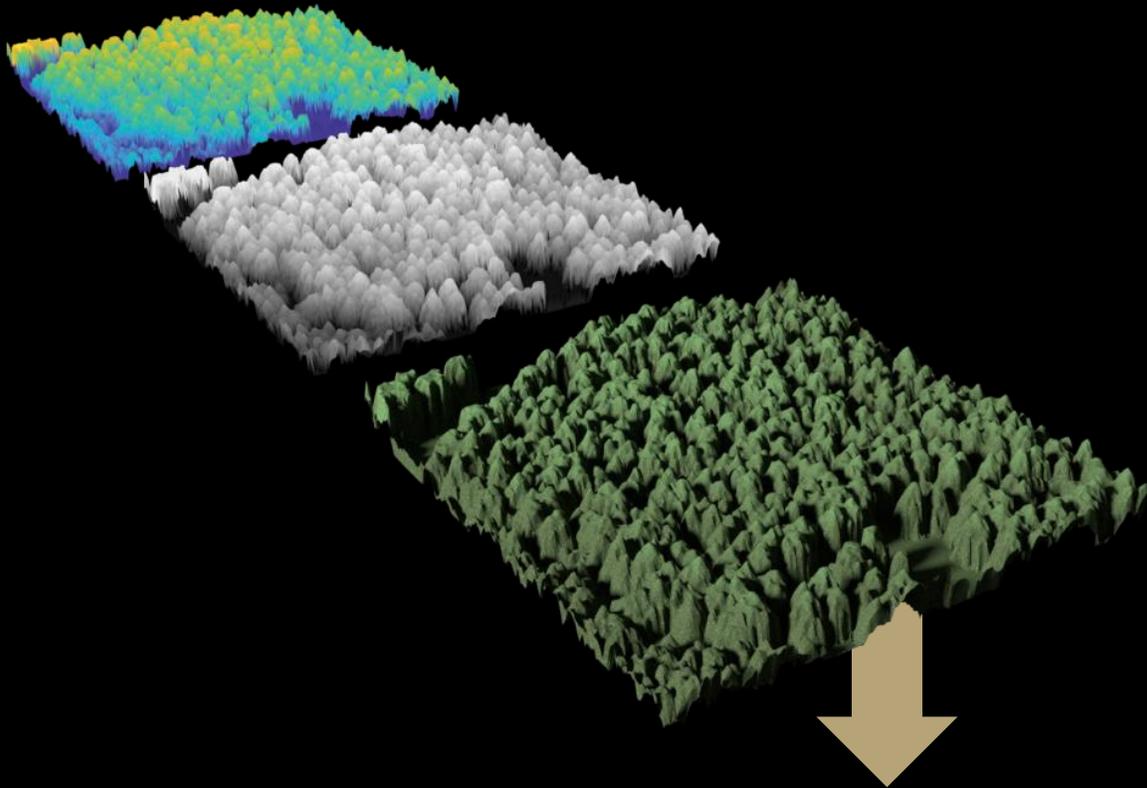


POC - Was haben wir gelernt

Doch woher kommen gelabelte Trainingsdatenmassen?
Teils von Satelliten, teils können wir sie künstlich erzeugen lassen.

Synthetic Data Sets nennt sich die Methode.
Beim autonomen Fahren schon groß in Mode.

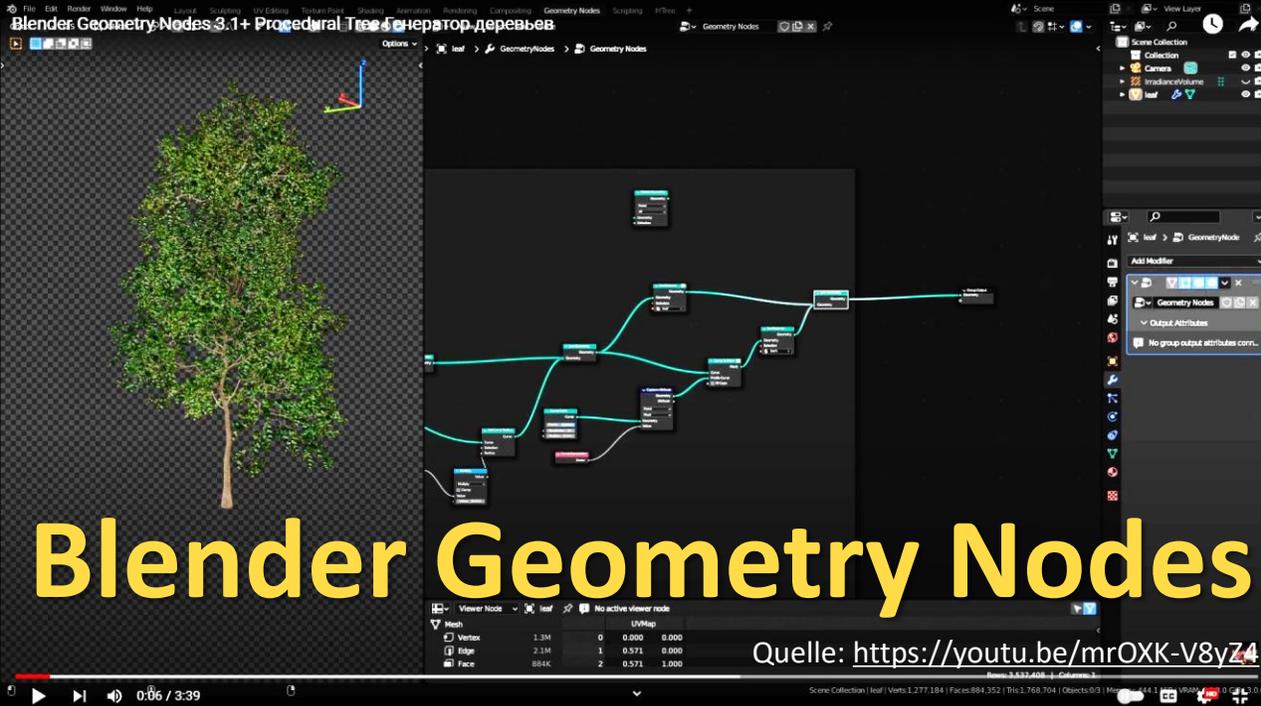
Doch eignet sich das auch für Biotope hier?
Den KI-Erkennungs-Nachweis lieferten wir.



Das ML-Modell identifizierte dies als Wald

Problem: Die Trainingsdatenmenge war deutlich zu gering.
Idee: synthetic Data Sets to the rescue > proven 3d data set built.
Illustration © by Jörg Osarek – 3d model and lighting created in blender





POC - Was haben wir gelernt

Die freie 3D Software Blender kann prozedural Vegetation erzeugen in unendlicher Zahl.*
 Das Labeln dabei automatisch abfällt:
 Trainingsdatenmassen für viel weniger Geld.

**Blender Addon
 Botaniq –**

Quelle der
 Bilder:
<https://blendermarket.com/products/botaniq-trees>



Wichtig ist umfassendes 3D Know How
 damit die Abbildung der Natur ausreichend genau
 Multispektral und im sichtbaren Licht
 doch das schreckt den 3D Enthusiasten nicht.

Problem: Die Trainingsdatenmenge war deutlich zu gering.

* Mit Blender Geometry Nodes lassen sich komplexe 3D-Szenarien prozedural generieren und inzwischen auch simulieren. Für Vegetation gibt inzwischen es auf Geometry Nodes basierende teils kostenpflichtige Plugins, um eine beliebige Synthetic Data Set Pipeline zu bauen, die neben optischen und multispektralen Bildern weitere Daten wie Labels, und Höheninformationen liefern kann - so, wie es für die spezielle Aufgabenstellung benötigt wird. Wir haben experimentell nachgewiesen: Die AI/KI akzeptiert diese synthetisch erzeugten Bilder und erkennt Wald vom annähernd fotorealistischen gerenderten Bild.



POC - Was haben wir gelernt



Midjourney Herbstwald
Seed: 1359038443



Midjourney Frühlingwald
Seed: 17121117

Diejenigen, die sich mit Synthetischen Trainingsdaten befassen sollten Generative AI nicht außer Acht lassen.

Also mit Text Prompts erzeugte Bilder, die werden wöchentlich immer besser und wilder.

Sind solche GANs zum Helfen in der Lage?

In Klärung ist gerade diese Frage. *

Was wir beim POC lernten war wirklich spitze.

Ich hoff, dass es Euch auch ein wenig nütze.

Habt Ihr Fragen, sprecht mich gerne an.

Ich freu mich, wenn ich weiterhelfen kann.

Problem: Die Trainingsdatenmenge war deutlich zu gering.

* Ob am Markt befindliche GANs ausreichend präzise steuerbar über einen längeren Zeitraum eine hinlängliche Realitätstreue erzeugen können, ist heute unklar. Trainiert man ein solches GAN selbst und liefert ihm nur akkurate, echte und relevante Trainingsdaten, sollte nichts dagegen sprechen, dass ein solches GAN präzise endlos viele Biotop-Bilder erstellen kann. Vielversprechend erscheint eine Aufteilung der Trainings-Pipeline in 1. einen ersten generativen Teil (Synthetic Data Sets), 2. das Trainieren eines speziellen GANs damit (Generative adversarial network) und 3. das Erzeugen der synthetic Datasets prompt based (Text to Image) direkt mit diesem trainierten GAN und 4. das Trainieren des eigentlichen neuronalen Netzwerks mit den durch das GAN erzeugten syntetischen Datasets. Dabei wird es hilfreich sein nicht nur Bilder zu erzeugen, sondern im gleichen Zug auch Depth Maps sowie weitere Spektren. Hinweis: Was zu beweisen wäre. Wenn sich dies als praktikabel erweist, lassen sich damit sehr effektive und effiziente KI-Trainingsdatenpipelines für hochkomplexe biologische, geologische oder auch Menschgemachte Umgebungen sowie für die Tierwelt (Fauna) oder auch die Medizin erstellen. Großer Vorteil: Ich habe den Bau der Trainingspipeline vollständig unter Kontrolle (wie das Klima in einem Gewächshaus) und die Trainingspipeline für das jeweilige ML-Projekt wird dadurch für den ML-Trainer deutlich vereinfacht.





Nächster Schritt: skaliertes System

Ein automatisches System, das über die globale Biosphäre wacht,
wäre das nicht wunderbar und eine Pracht?

Um ein solch edles System zu skalieren,
brauchen wir Menschen, die in das Projekt investieren.

47 Seiten POC Bericht lesen dauert eine Ewigkeit.
Vergeuden will ich nicht der Investoren kostbare Zeit.
Wie kann ich es den Damen und Herrn
schnell und einfach erklären?
Vielleicht indem mein Gedicht sie erfreut.

Illustration zu „Die Sterntaler“ von Viktor Paul Mohn, 1882 (1842-1911), Public Domain – gemeinfrei, Urheberrecht abgelaufen –
https://commons.wikimedia.org/wiki/File:%D0%92-%D0%9F-%D0%9C%D0%BE%D0%BD-%D0%97%D0%B2%D0%B5%D0%B7%D0%B4%D0%BD%D1%8B%D0%B5_%D1%82%D0%B0%D0%BB%D0%B5%D1%80%D1%8B.jpg





Sherlock Holmes im Orbit: Künstliche Intelligenz erkennt Biotope auf Satellitenbildern, unterstützt durch Synthetic Data Sets

Download dieses Slide Decks:
DE: sherlock-ki.climatehackerz.com
EN: sherlock-ai.climatehackerz.com

Jörg Osarek, May 2023
m:contact at climatehackerz.com
T: +49-151-23-0-24-333



2023 - 2035



Images (where not mentioned otherwise) mostly generated by Midjourney Basic plan with General commercial terms 2023 or © by Joerg Osarek and Chuck Ian Gordon

