



Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Swiss Confederation

Bundesamt für Statistik BFS
Office fédéral de la statistique OFS
Ufficio federale di statistica UST
Federal Statistical Office FSO

Arealstatistik Deep Learning (ADELE)

Swiss Insights TALK 2024

16. Mai 2024
Gillian Milani

Ablauf der Präsentation

1. Das BFS in Kürze
2. Die Arealstatistik
3. Automatisierung
 1. Ziele
 2. Integration in den Betrieb
 3. Aktuelle Grenzen
 4. Neue Möglichkeiten

Wozu braucht es Statistik?

Statistiken ...

- **objektive** Fakten
- **Grundlage** für die Steuerung und Entscheidungsfindung
- kollektive **Gedächtnis** einer Gesellschaft
- objektive Informationen zur **Staatsführung**
- **unverzichtbaren Beitrag** für einen modernen, demokratischen Staat

Themenfelder

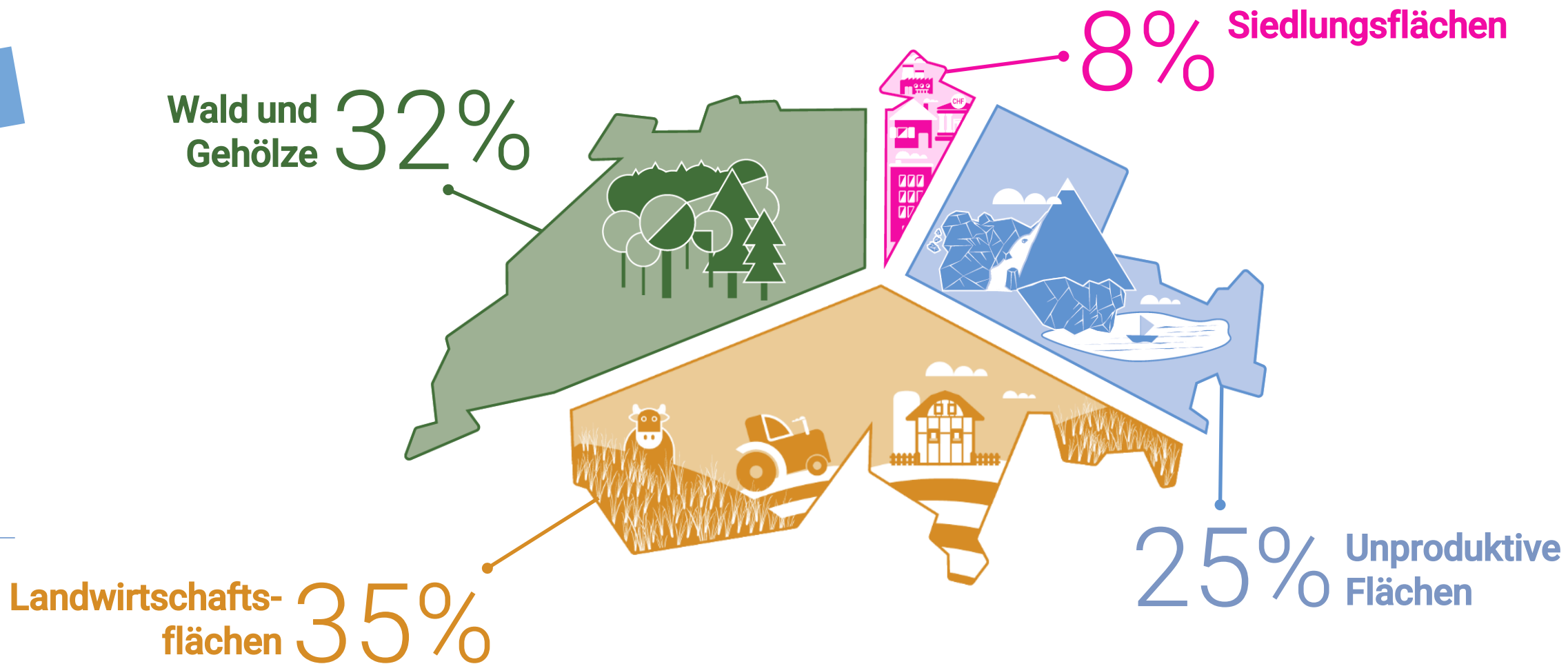
- Bevölkerung
- Bildung und Wissenschaft
- Energie
- Gesellschaft
- Gesundheit
- Mobilität und Verkehr
- Raum, Umwelt und nachhaltige Entwicklung
- Soziale Sicherheit
- Wirtschaft und öffentliche Finanzen
- Register

Diffusion

- 18,7 Millionen Seitenaufrufe des BFS-Angebots im Internet pro Jahr
- 229 Publikationen pro Jahr (sprachübergreifend)
- 150 Medienmitteilungen pro Jahr (in 4 Sprachen)
- 30 000 Mal in den Medien erwähnt
- 3769 Posts auf den Social Media-Kanälen
- 10 000 *Followers* auf LinkedIn

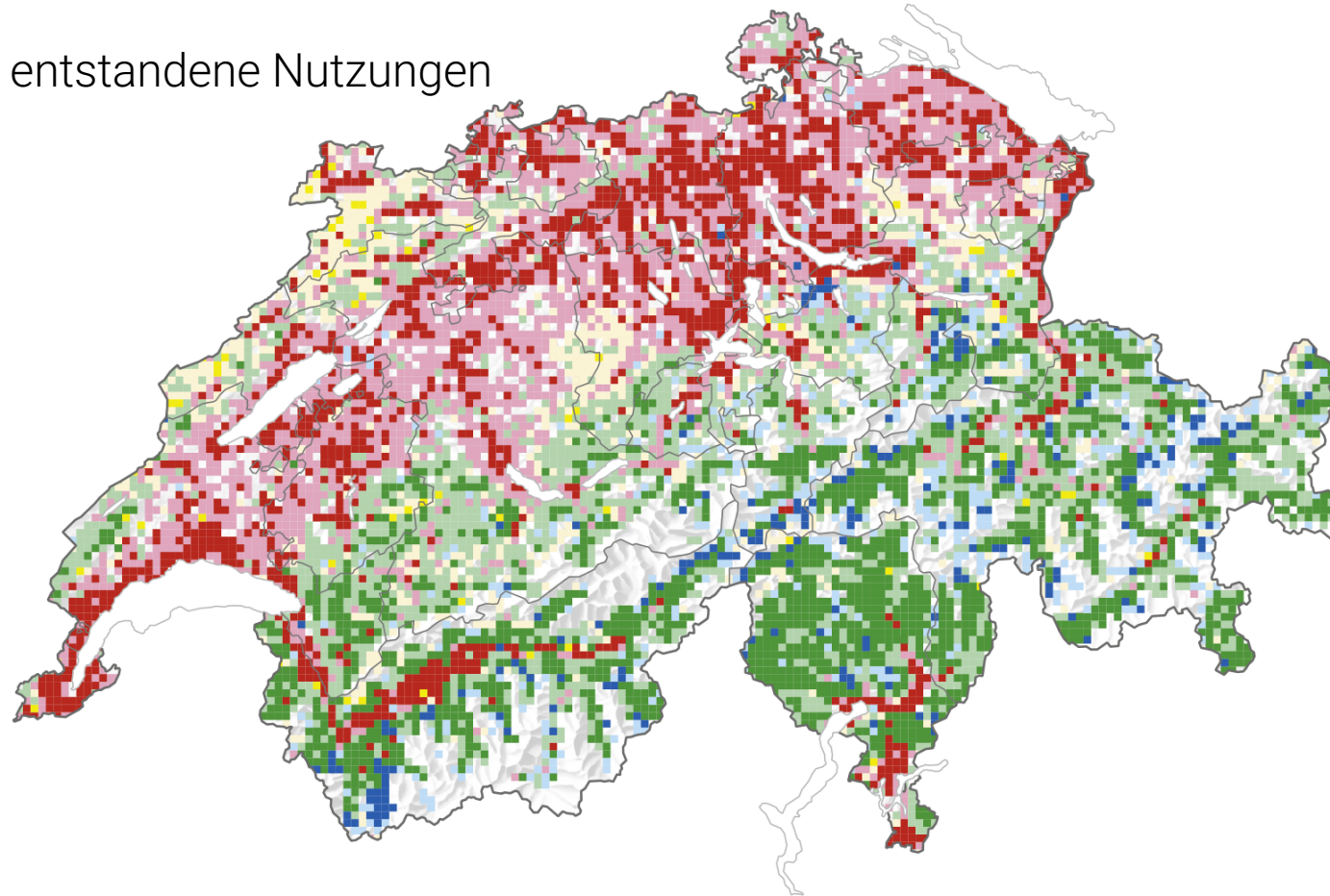


Eine vielfältige Schweiz



Die Bodennutzung verändert sich

Wichtigste neu entstandene Nutzungen
1985–2018



Nutzungsart mit der grössten Zunahme

- Siedlungsflächen
- Landwirtschaftsflächen
- bestockte Flächen¹
- unproduktive Flächen

Intensität der Zunahme, in %

- Nicht gefärbt < 1,0
- Hell gefärbt 1,0 – 4,9
- Dunkel gefärbt ≥ 5,0

¹ Wald und Gehölze

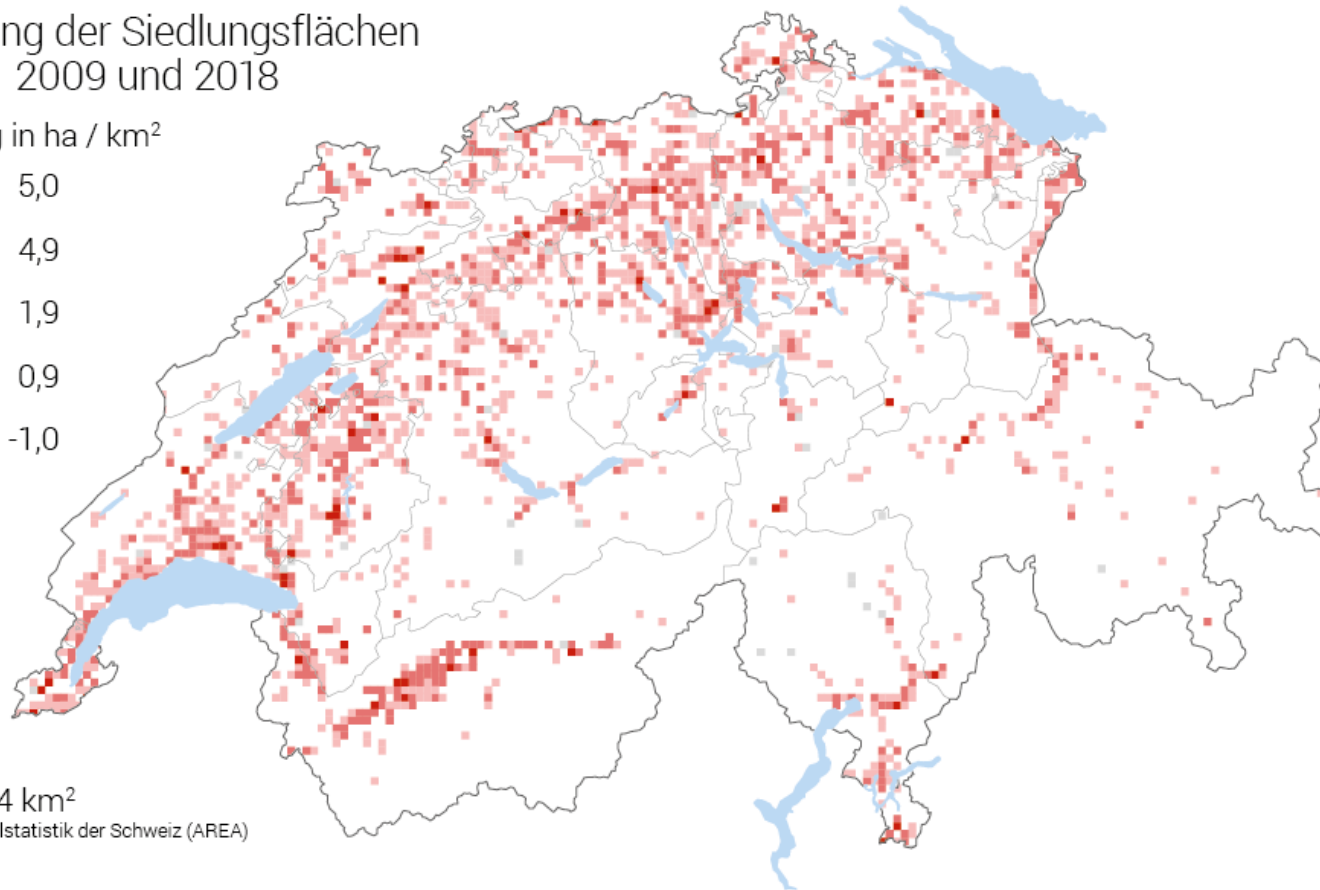
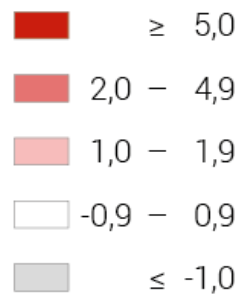
0 25 50 km

Auflösung: 4 km²

Eine zunehmend urbane Schweiz

Entwicklung der Siedlungsflächen
zwischen 2009 und 2018

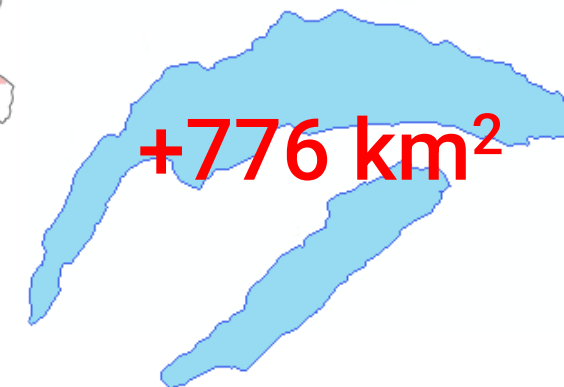
Entwicklung in ha / km²



Auflösung: 4 km²

Quelle: BFS – Arealstatistik der Schweiz (AREA)

Bilanz seit 1985

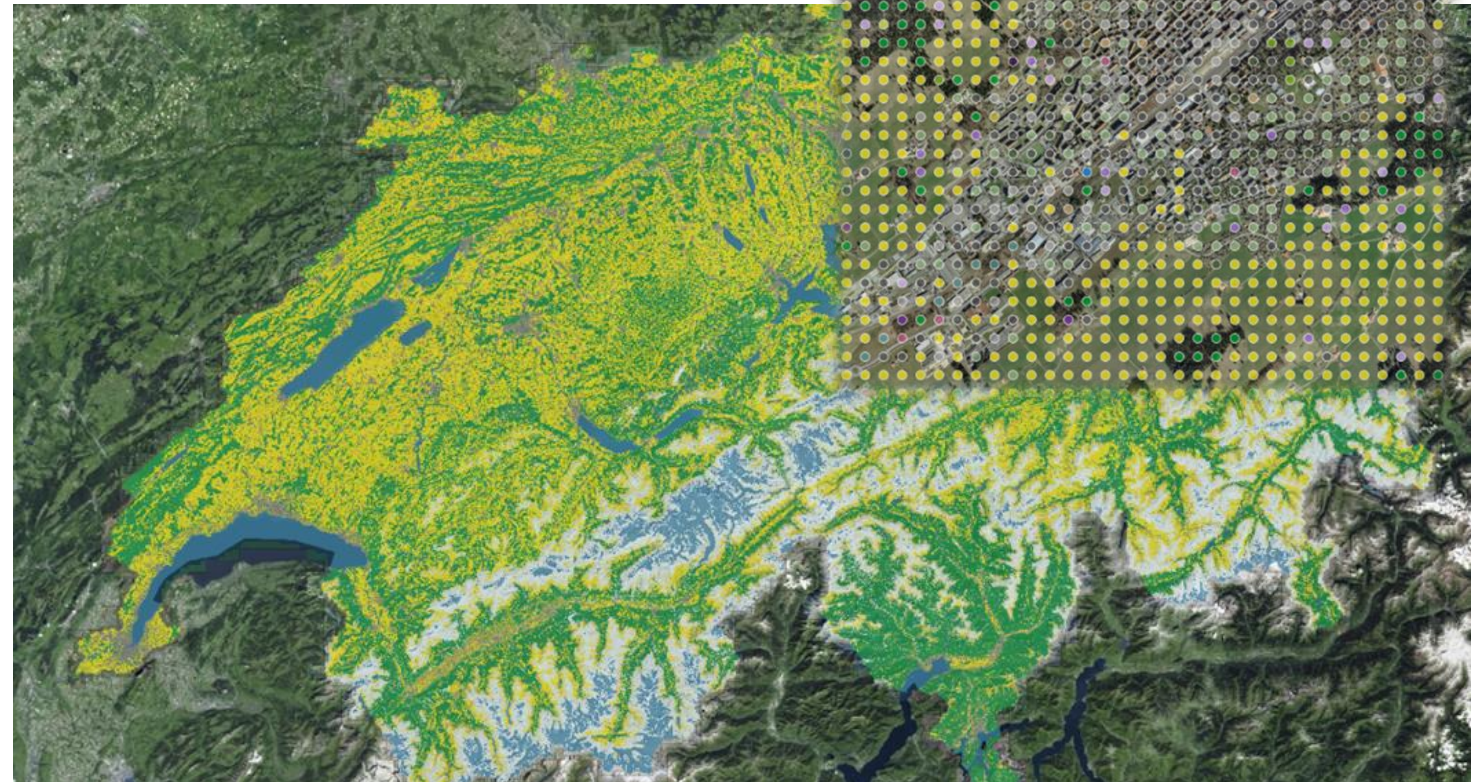


Etwa die Fläche des
Genfer- und des
Neuenburgersees

Was ist genau die Arealstatistik?

Klassifikation **Bodennutzung** (46 Klassen)
und **Bodenbedeckung** (27 Klassen)

Langfristige Zeitreihe
(seit 1979) mit 4.1
Mio. fixen
Stichprobepunkten, in
fünf Erhebungen
untersucht.



Die Arealstatistik

Beispiel für einen interpretierten Stichprobenpunkt



Stichprobenpunkt in der Mitte:

Bodenbedeckung (LC):

21 – Grass-/Krautvegetation

Bodennutzung (LU):

163 - Golfplätze



Wie entstehen die Statistiken?

Fünf Erhebungsperioden, basierend auf Luftbildern von...



...1979/85



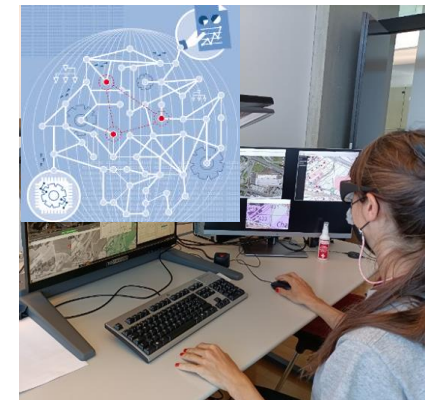
...1992/97



...2004/09



...2013/18



...2020/25

Intervalle

12

12

9

6



abgeschlossen

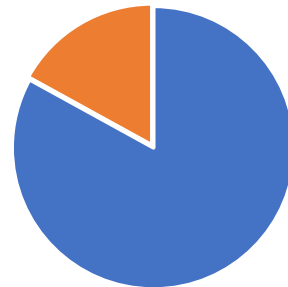


in Arbeit

Ziele des ADELE-Projekts

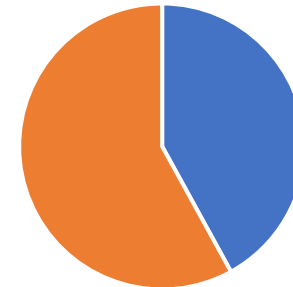
Verkürzung des Erhebungszeitraums von 9 auf 6 Jahre
ohne zusätzliche Ressourcen
Beibehaltung der Qualität der Zeitreihen (Nomenklatur)

Punkte ohne KI
Automatisierung



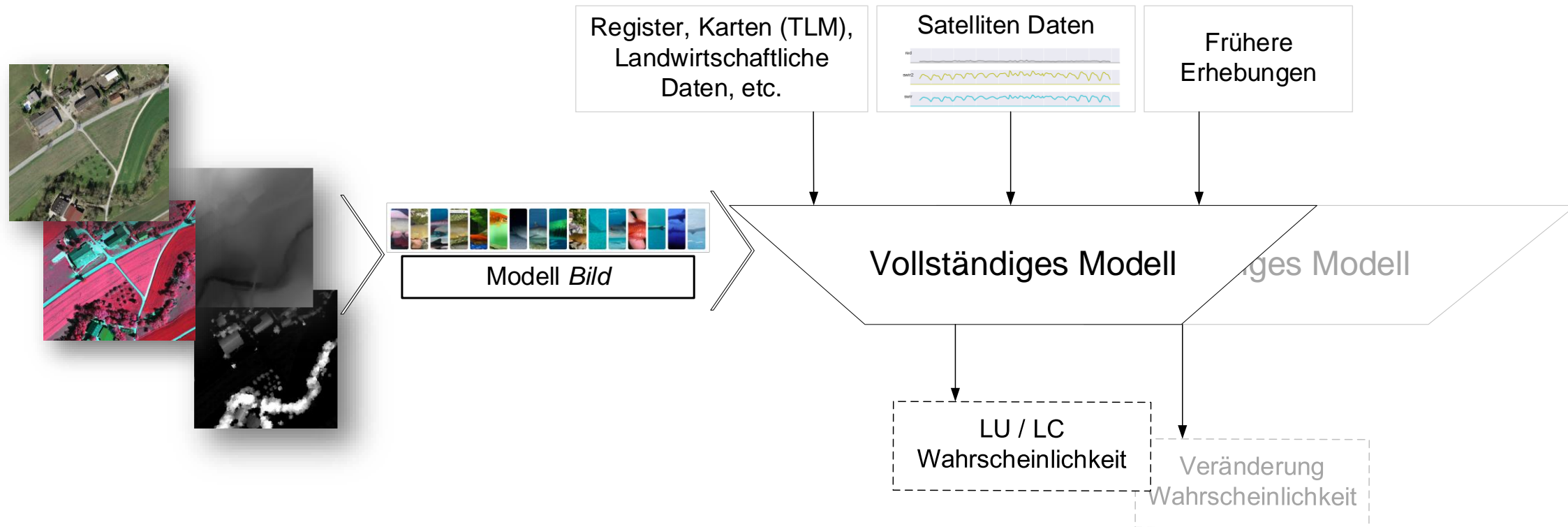
■ interpretiert ■ nicht interpretiert

Punkte mit ADELE
(Ziel)



■ interpretiert ■ nicht interpretiert

ADELE: Schema des Modells



Implementierung KI

Beispiel von Vorhersage

2013

LC 21 (Gras-, Krautvegetation)

LU 223 (Dauerweide)



2020

LC 52 (Lockergestein)

LU 146 (Baustellen)



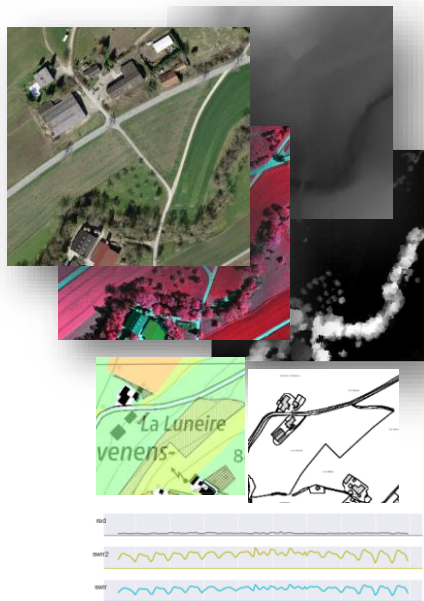
Vorhersage

Typ	Resultat (Wkt. [%])
Veränderung	Ja (88%)
LC	52 (94%)
LU	146 (95%)

Integration in den Interpretationsprozess

Kompromisse zwischen Automatisierung und Qualität der Änderungserkennung

Daten



Maximierung der automatisierten Punkte

Veränderung erkennt ?

Automatische Klassifizierung

Maximierung der Erkennung von Veränderungen

Visuelle Interpretation

Herausforderung

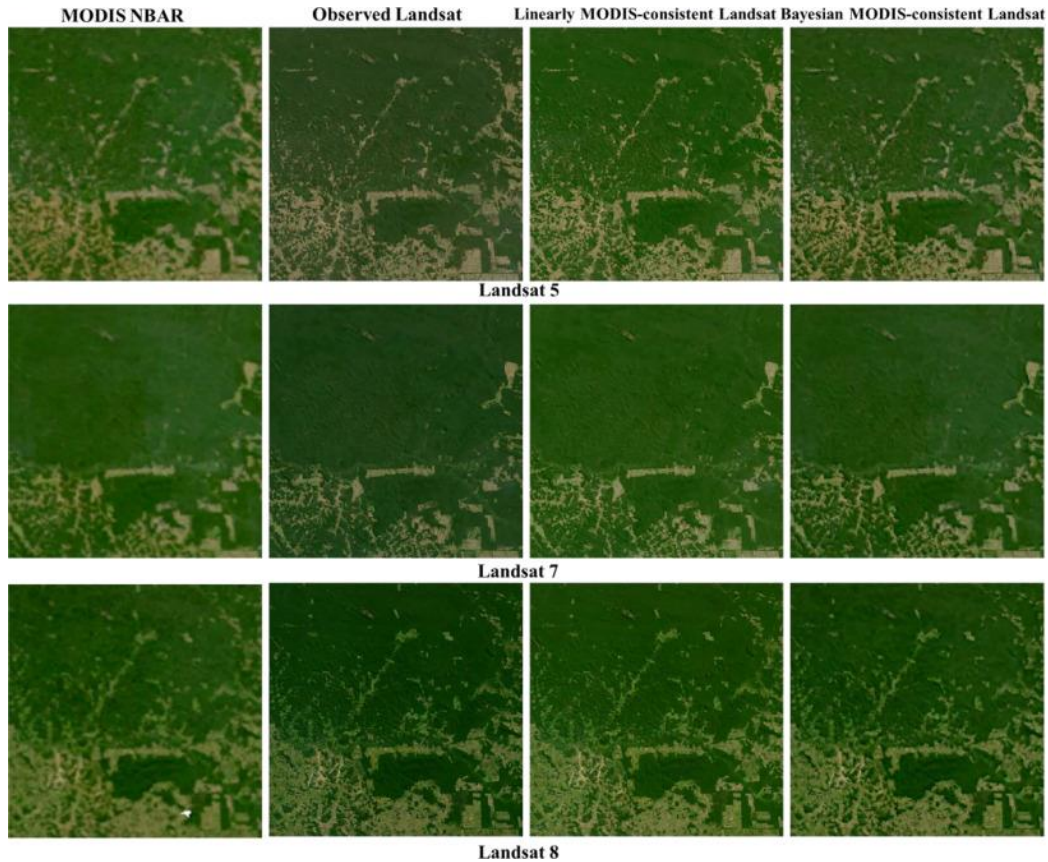
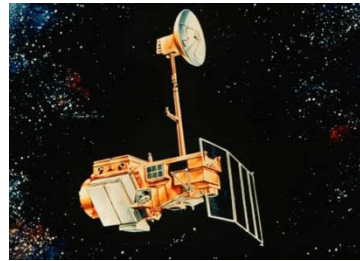
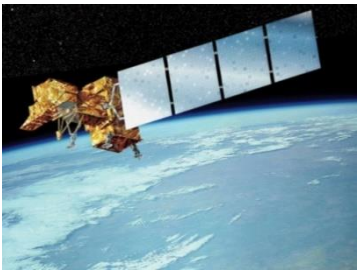
Heterogenität der verfügbaren Daten über einen längeren Zeitraum



XKCD

Herausforderung

Heterogenität der verfügbaren Daten über einen längeren Zeitraum

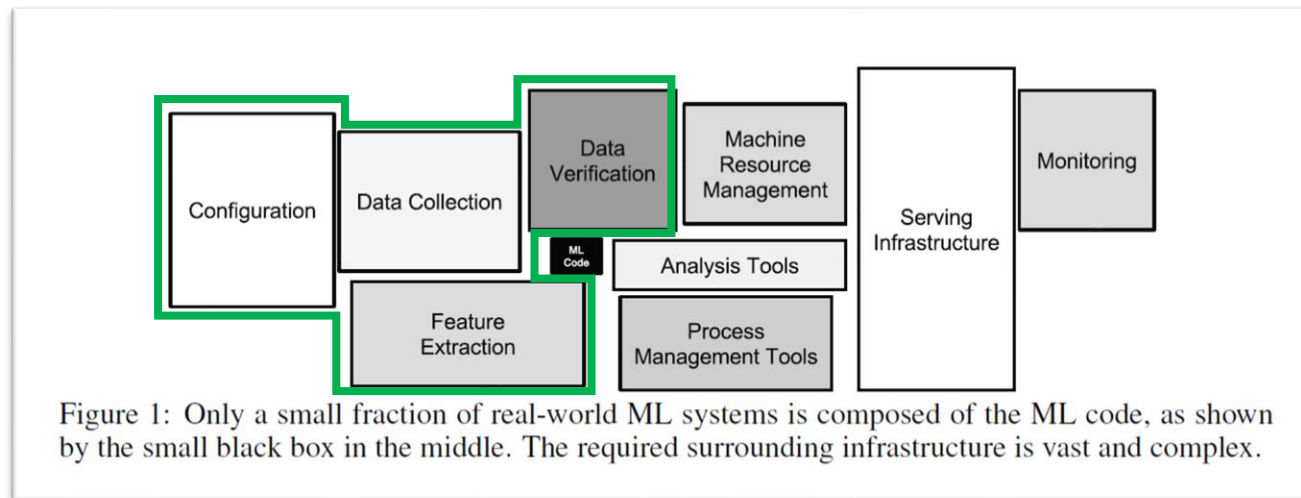


Che, X., Zhang, H. K., & Liu, J. (2021). Making Landsat 5, 7 and 8 reflectance consistent using MODIS nadir-BRDF adjusted reflectance as reference. *Remote Sensing of Environment*, 262, 112517.



Herausforderung

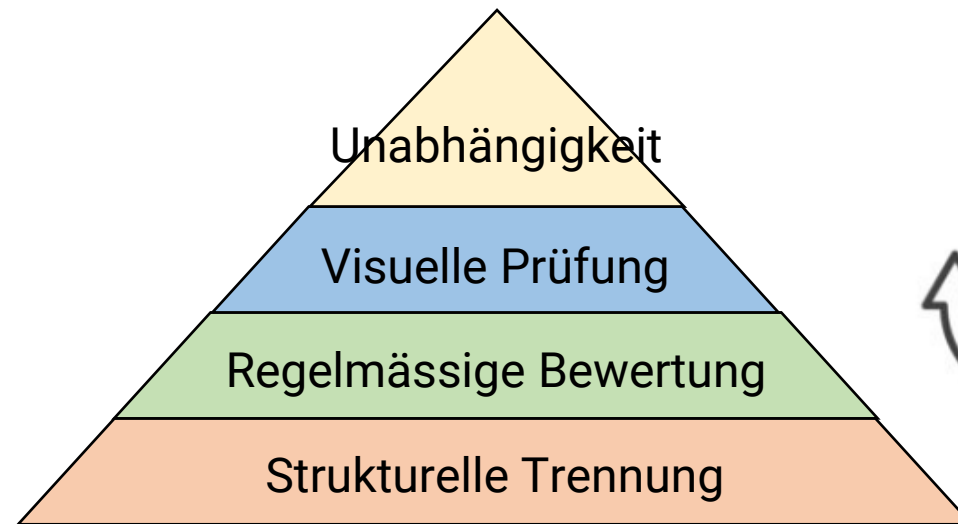
- Heterogenität der verfügbaren Daten über einen längeren Zeitraum
- Grosse Datenmengen bei begrenzten Ressourcen



Sculley, D., Holt, G., Golovin, D., Davydov, E., Phillips, T., Ebner, D., ... & Dennison, D. (2015). Hidden technical debt in machine learning systems. *Advances in neural information processing systems*, 28.

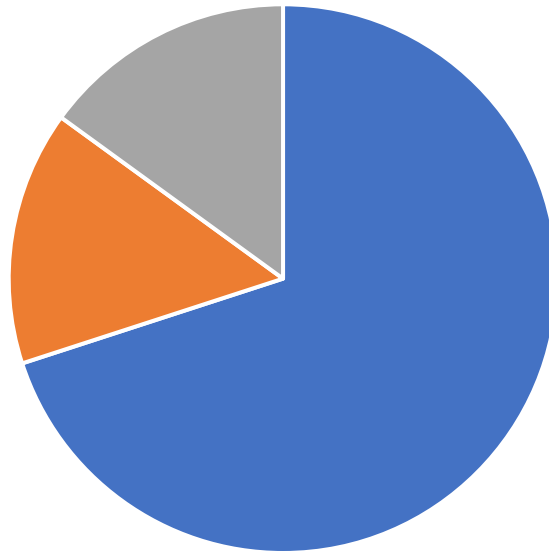
Herausforderung

- Heterogenität der verfügbaren Daten über einen längeren Zeitraum
- Grosse Datenmengen bei begrenzten Ressourcen
- Kohärenz mit hohen Qualitätsstandards



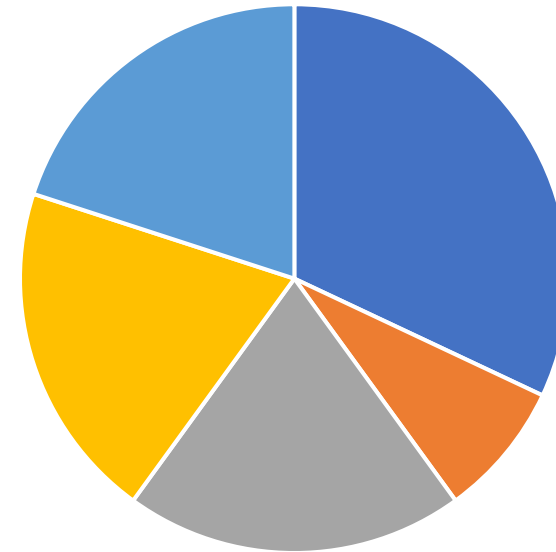
Kontroll-Praktiken

Standard-Trennung



■ Training ■ Validierung ■ Test

ADELE-Trennung



■ Training ■ Validierung ■ Test ■ Kalibrierung ■ Kontrolle

Vom Experimentellen zur Produktion

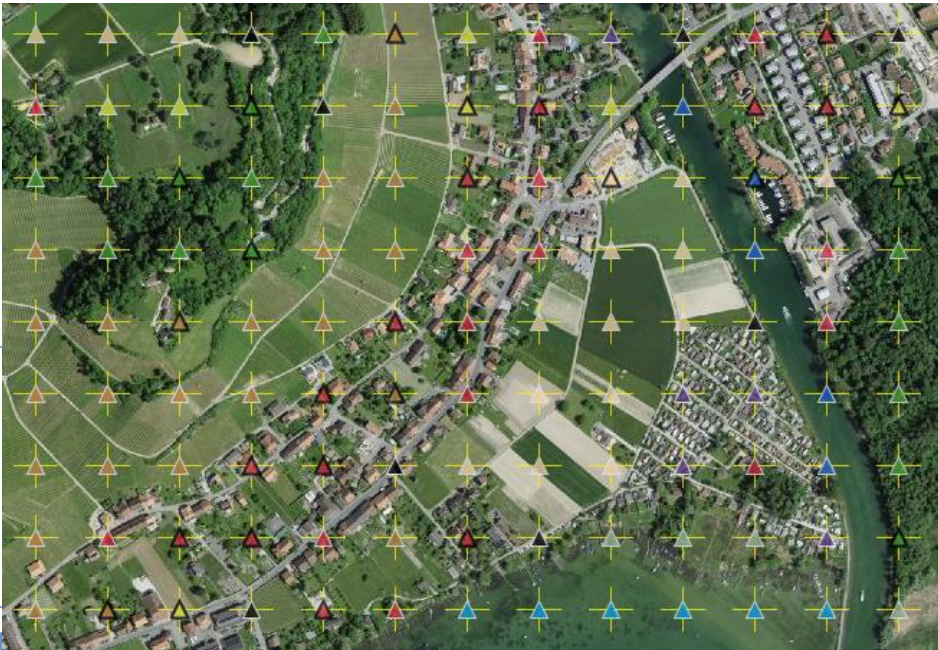
- Dokumentation und Entscheid des Transfers in die Produktion
 - **34 spezifische Kriterien für die experimentelle Statistik, z.B.:**
 - Beitrag zur Innovationsstrategie des BFS
 - Allgemeine Relevanz und Mehrwert
 - Reproduzierbare Beschreibung der Validierung der Ergebnisse
 - Auswirkungen auf den Produktionsprozess

Beobachtungen der Inbetriebnahme

- Das Potenzial zur Verringerung des Arbeitsaufwands bei der visuellen Interpretation ist gross, erfordert jedoch hohe Kosten für die Umsetzung
- Datenbeschaffung und -aufbereitung: Zeit und Ressourcen hoch
- Analyse der Resultate: Zeit und Ressourcen hoch
- Folgen auf benötigte Stellenprofile. Neu 50%-Stelle für Entwicklung und Integration der KI-Modelle (anstelle visuelle Luftbildinterpretation)

Neue Möglichkeiten: Verdichtung der Klassifikation

25 cm Bilder als Basis
5 zu 10 Klassifizierung pro Segment
zusätzliche Schritte für kleine Strukturen



Ein Blick in die Zukunft

Vollständige Automatisierung, Neuzuteilung von Ressourcen

Heute

- Beschleunigung
- Neue Produkte
- Kostensenkung

in 10 Jahren ?

- Veränderungen in der Gesellschaft und der Industrie werden nicht berücksichtigt
- Statistische Verzerrung
- Verlust an Genauigkeit

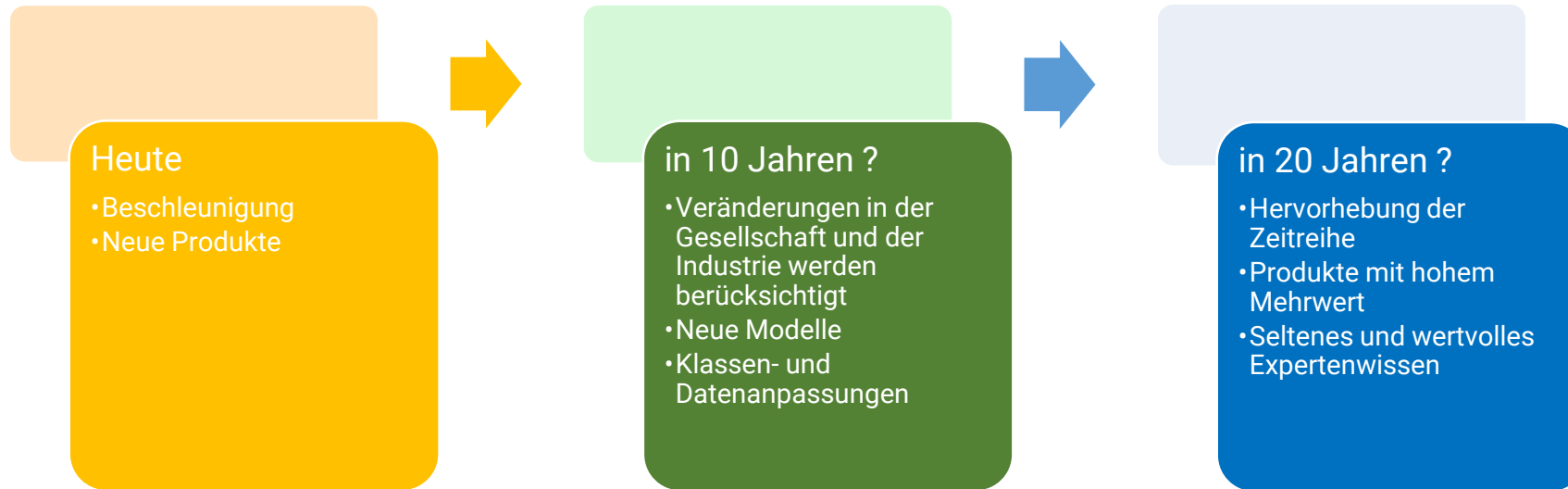
in 20 Jahren ?

- Ungenauigkeit: Neue Statistik unbrauchbar
- Expertenwissen und Organisation nicht mehr verfügbar
- Daten für Automatisierung nicht mehr verfügbar



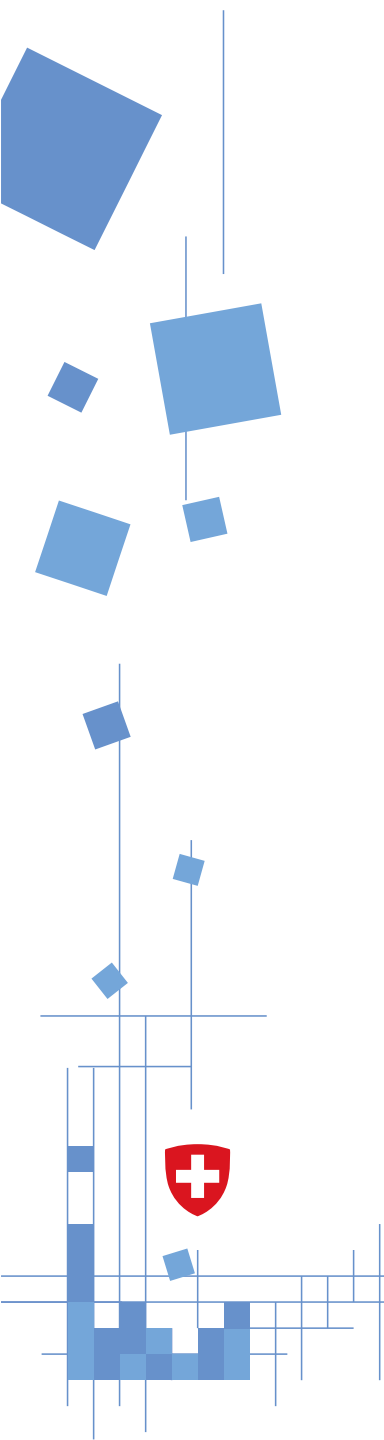
Ein Blick in die Zukunft

Gewissenhafte Automatisierung, langfristige Perspektive



Fazit

- **KI bildet eine Hilfe** für die Bearbeitung repetitiver Aufgaben
- in der Praxis ist KI **ressourcenintensiv**
- **KI ist Statistik:** Der Mensch muss ein zentrales Element in diesem Prozess bleiben, um Fehlentwicklungen zu vermeiden. Die nächsten Jahre werden entscheidend sein, um die nachhaltige Nutzung wertvoller Daten sicherzustellen.



geostat@bfs.admin.ch

[Aktuell BFS](#) (07.09.2023) "Künstliche Intelligenz:
Erster Einsatz in der Arealstatistik".