# Einführung in Data Science und maschinelles Lernen mit R

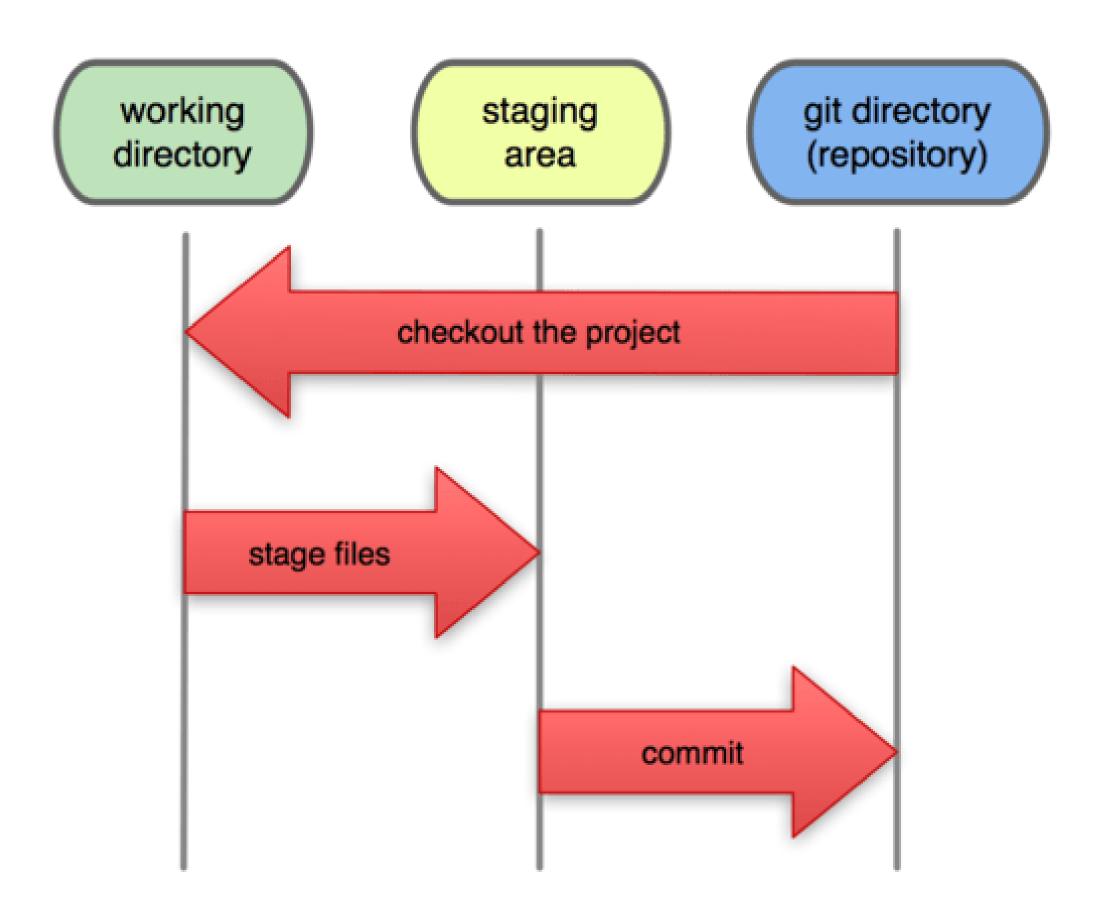
Versionierung mit git (Teil 2) und Start des Auswertungsprojektes



- Einführung in git Teil 2
- Zusammenstellung der Teams
- Erste Arbeiten im Team
- "Al for Everyone"



#### VERSIONIERUNG MIT GIT

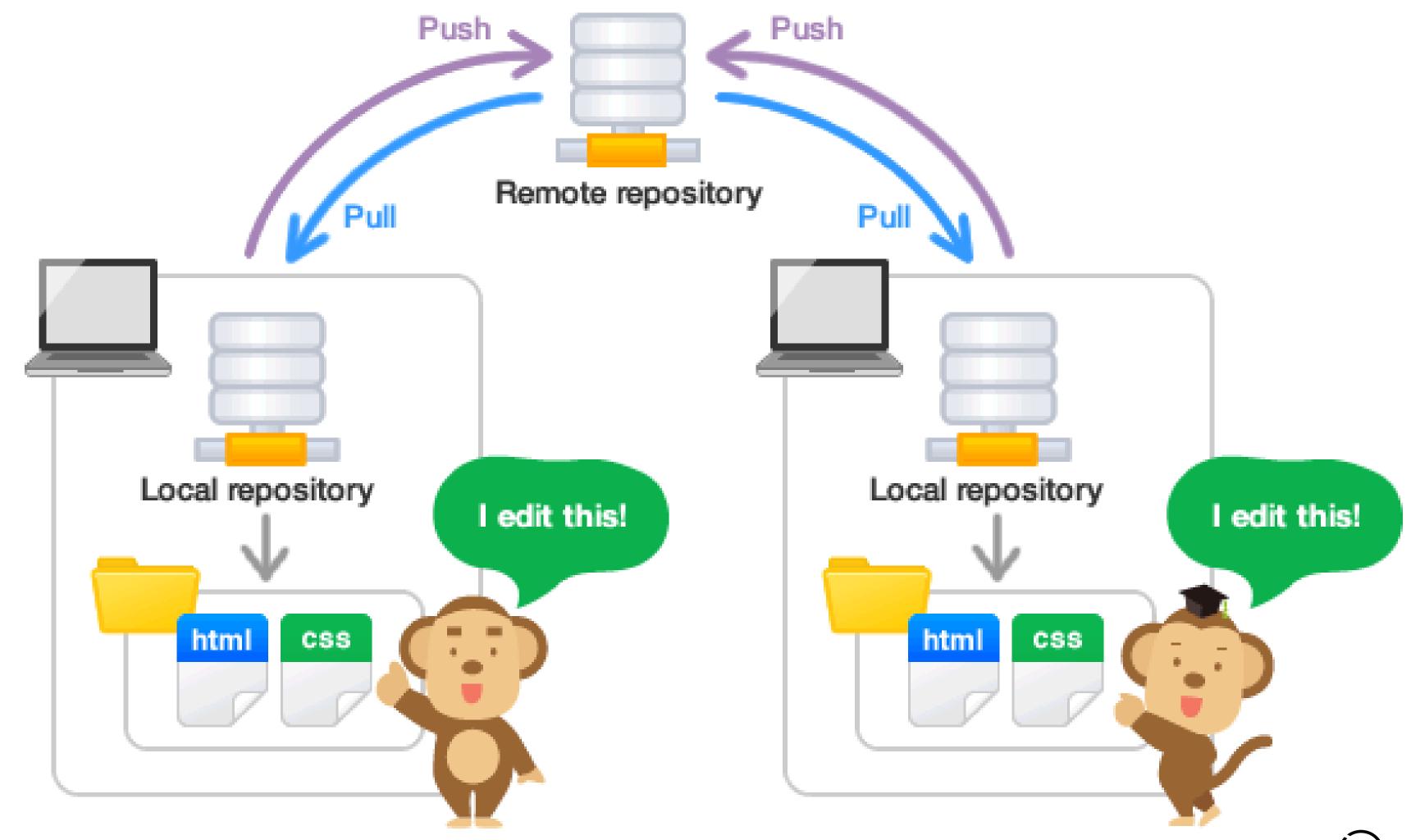


# Eine Datei kann drei möglich Zustände haben:

- modified ("geändert")
- staged ("vorgemerkt")und
- committed ("versioniert").



#### LOKALE UND REMOTE REPOSITORIES



# AUTHENTIFIZIERUNG BEI GITHUB

Zur Verbindung Eures RStudio mit Eurem GitHub Account führt folgende zwei Schritte durch:

#### (1) RStudio Menu

(Generierung eines Schlüssels in RStudio, um sich mit diesem bei GitHub zu authentifizieren)

Tools → Global Options → Git SVN

Auf "Create RSA Key..." klicken und über "View public key" den angezeigten Schlüssel in die Zwischenablage kopieren.

#### (2) GitHub

(Einfügen des generierten Schlüssel in GitHub) Unter Account Settings gehe zu *Edit profile*  $\rightarrow$  *SSH key*  $\rightarrow$  *Add SSH key* 



#### **PROJEKT**

# Präsentation (Powerpoint / R Markdown)

- Euren Namen auf der Titelseite
- Vorstellung der selbst erstellten Variablen
- Balkendiagrammen mit Konfidenzintervallen für zwei selbst erstellte Variablen
- Optimierung einer SVM:
  - (a) Source code
  - (b) MAPE je Warengruppe
  - (c) Warengruppenumsätze für den 05.06.2019
- Optimierung eines neuronalen Netzes:
  - (a) Source Code
  - (b) MAPE je Warengruppe
  - (c) Warengruppenumsätze für den 05.06.2019

#### Repository

- Datensatz
- Programmcode
- Präsentation



# Zusammenstellung der Teams



#### Dokumentation des Teams

- Vor- und Nachname
- Nummer des Teams

Ausfüllen des im Zoom Chat verlinkten Fragebogens.

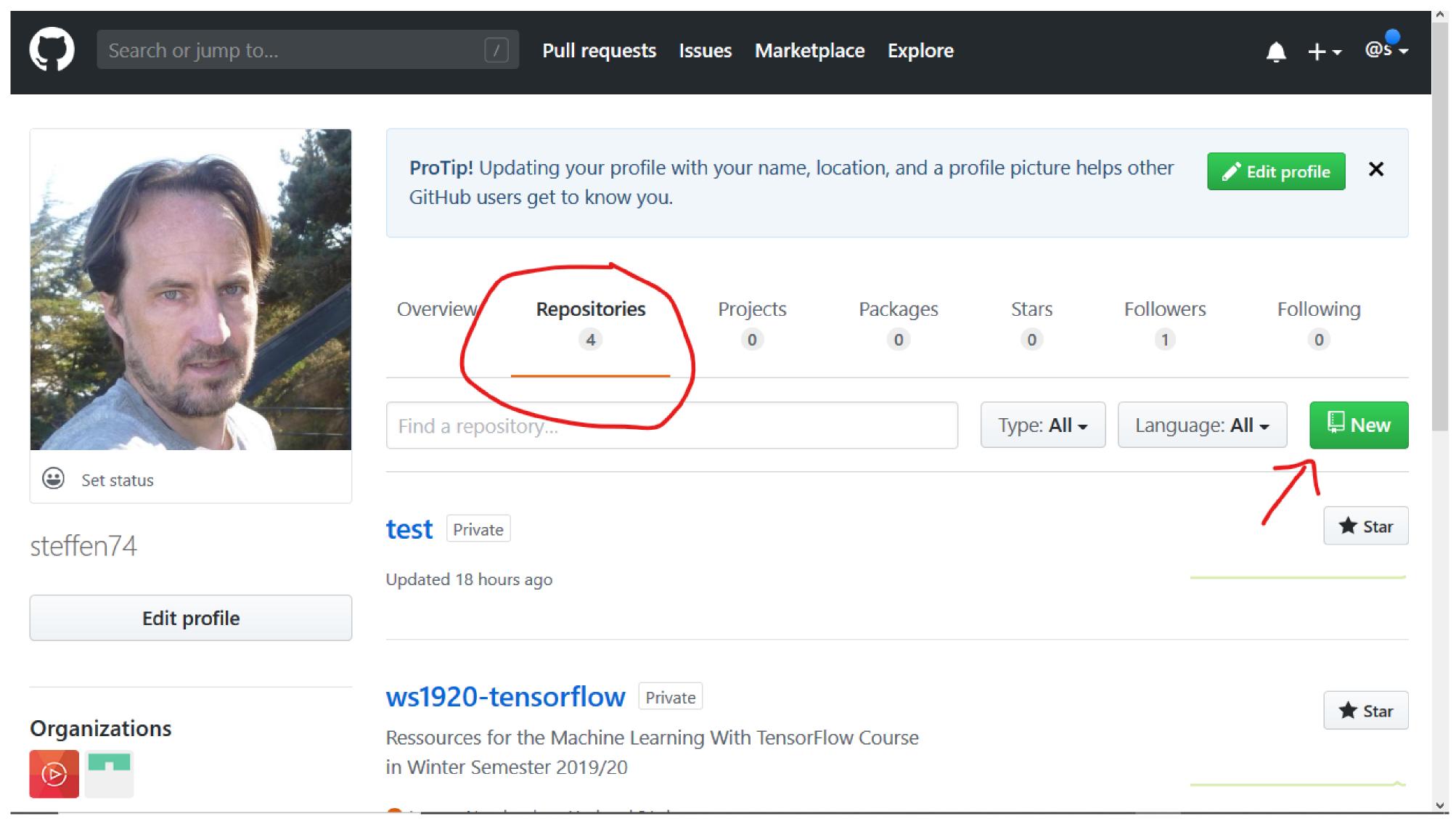


#### TEAMAUFGABEN

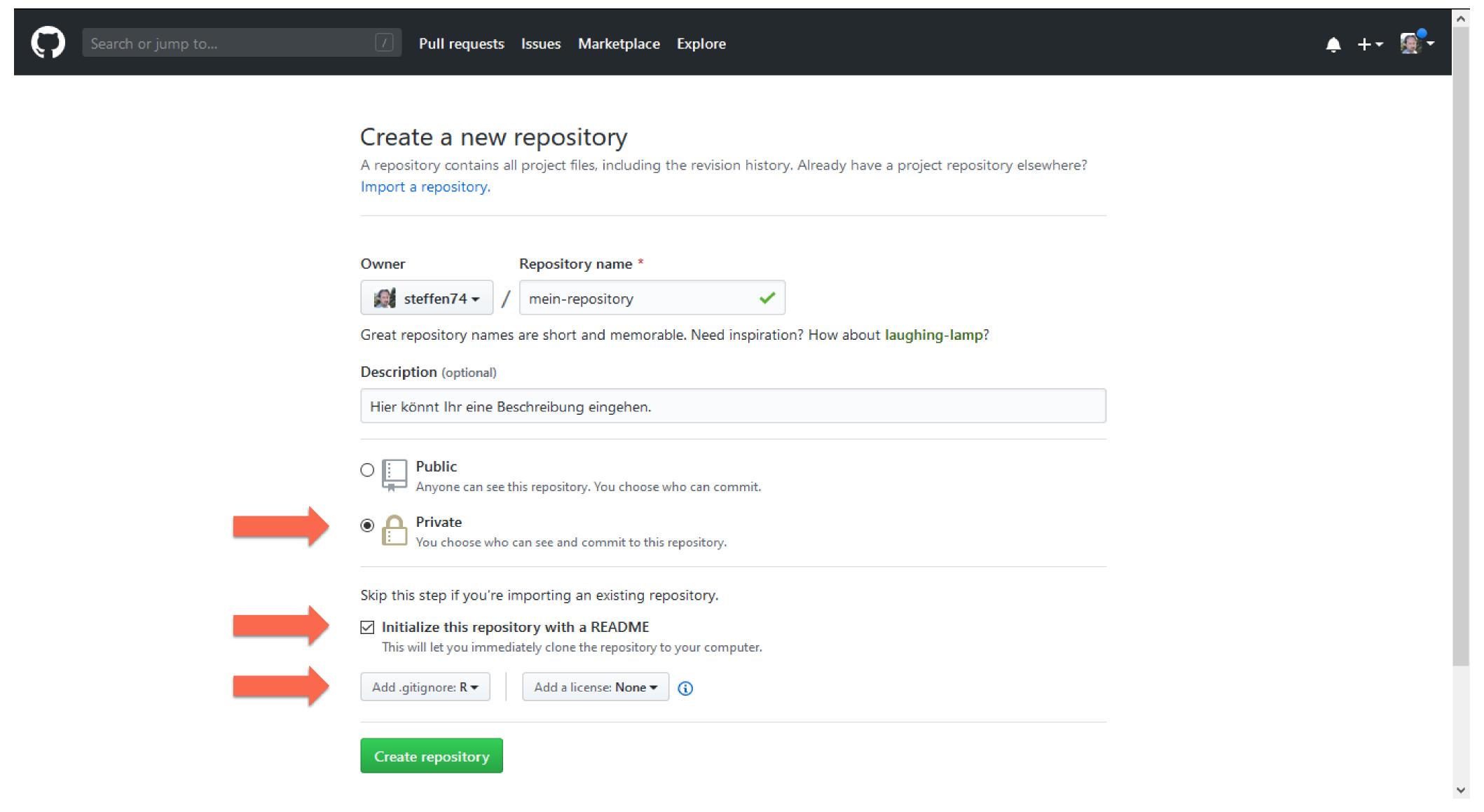
- (1) Anlegen eines GitHub Repository (Remote Repository) im GitHub durch eines der Team-Mitglieder
- (2) Einrichten lokaler Repositories, die mit dem GitHub Repository verbunden sind.
- (3) Anlegen eines Team Channels im Chat ("Data Science Team 1", ...)
- (4) Entwicklung von Ideen zur Konstruktion zusätzlicher Variablen, die für die Vorhersage des Umsatzes in der Bäckerfiliale relevant sein können.
- (5) Verteilung von Aufgaben im Team



# NEUES GITHUB REPOSITORY ANLEGEN (1)

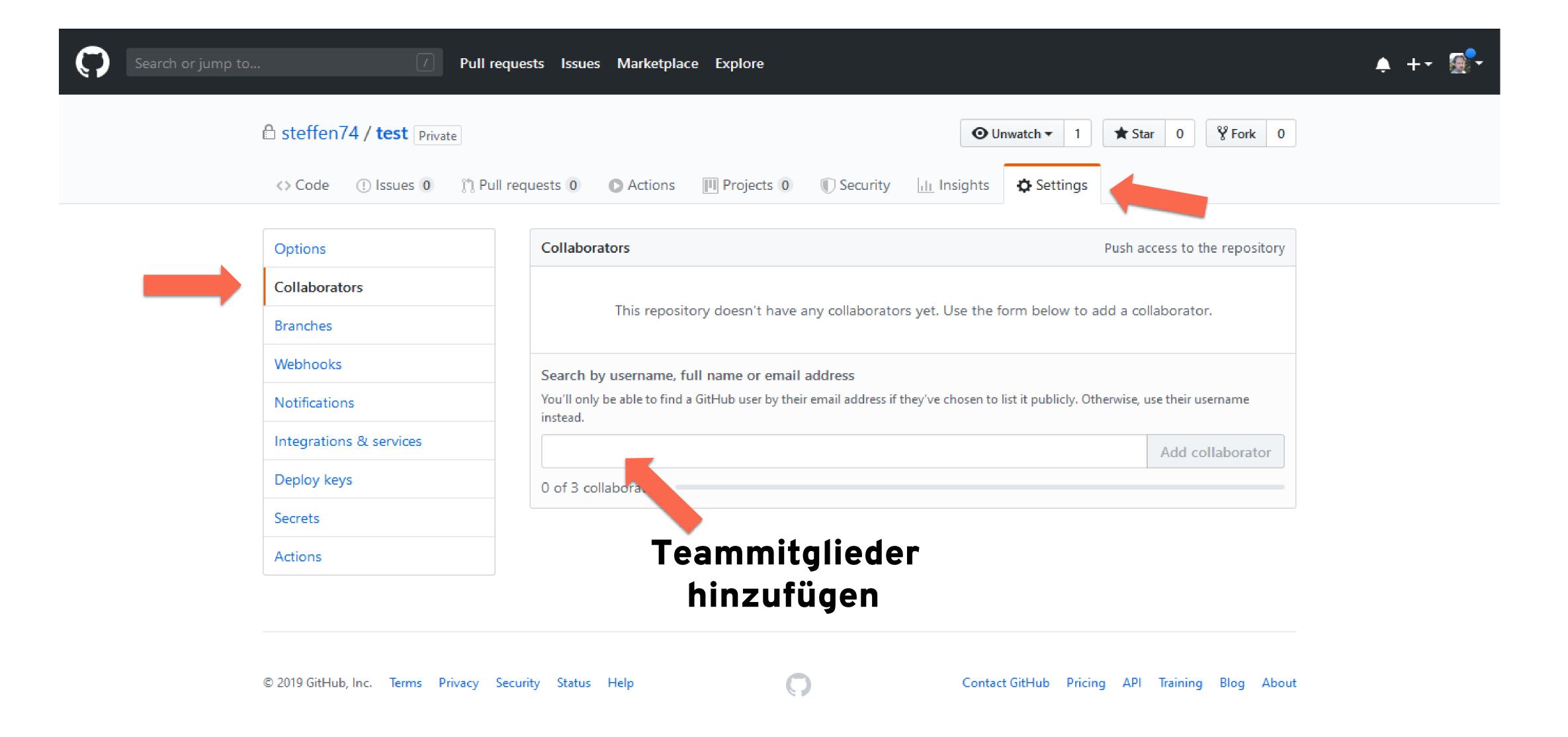


# NEUES GITHUB REPOSITORY ANLEGEN (2)



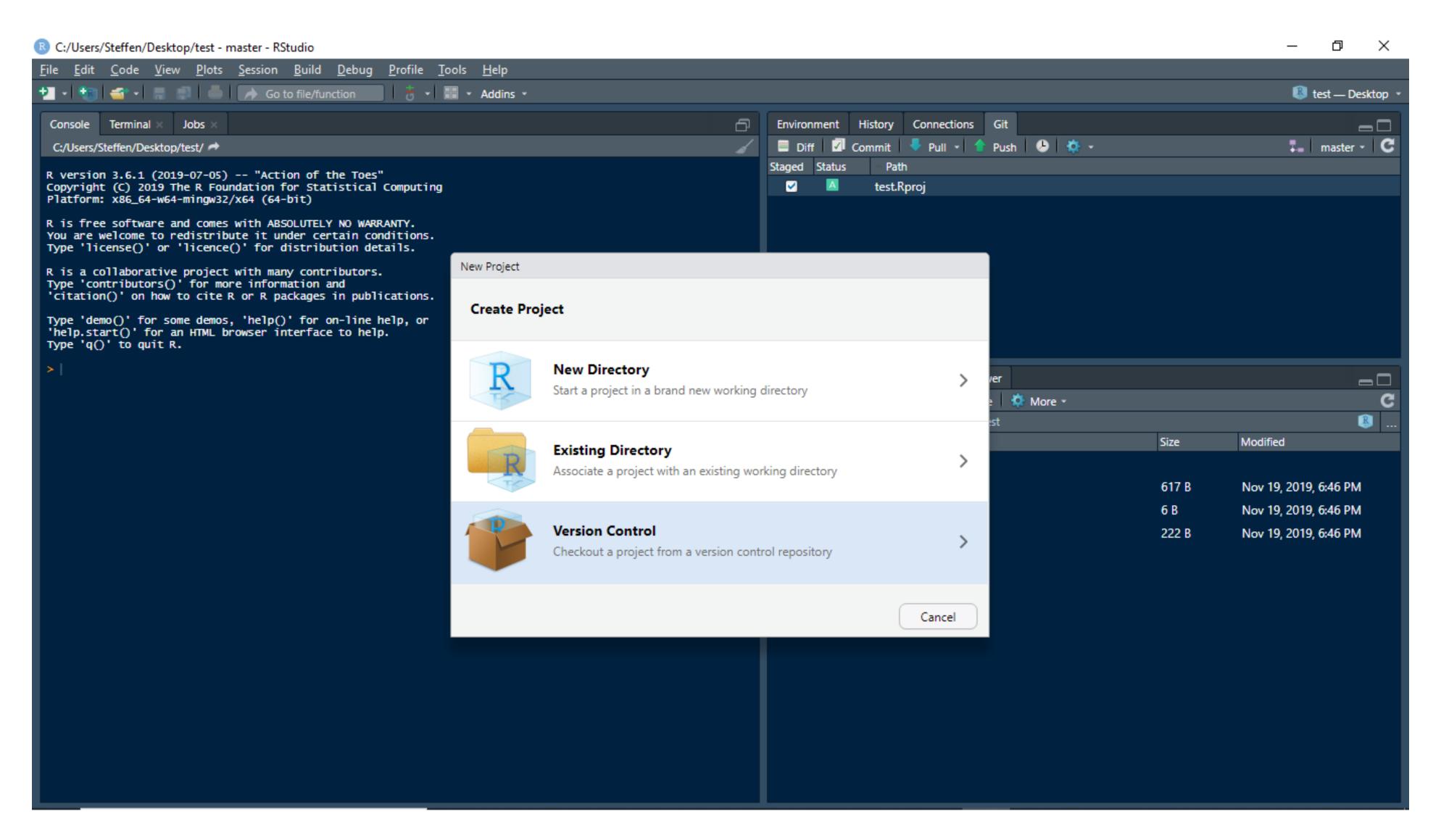


# NEUES GITHUB REPOSITORY ANLEGEN (3)



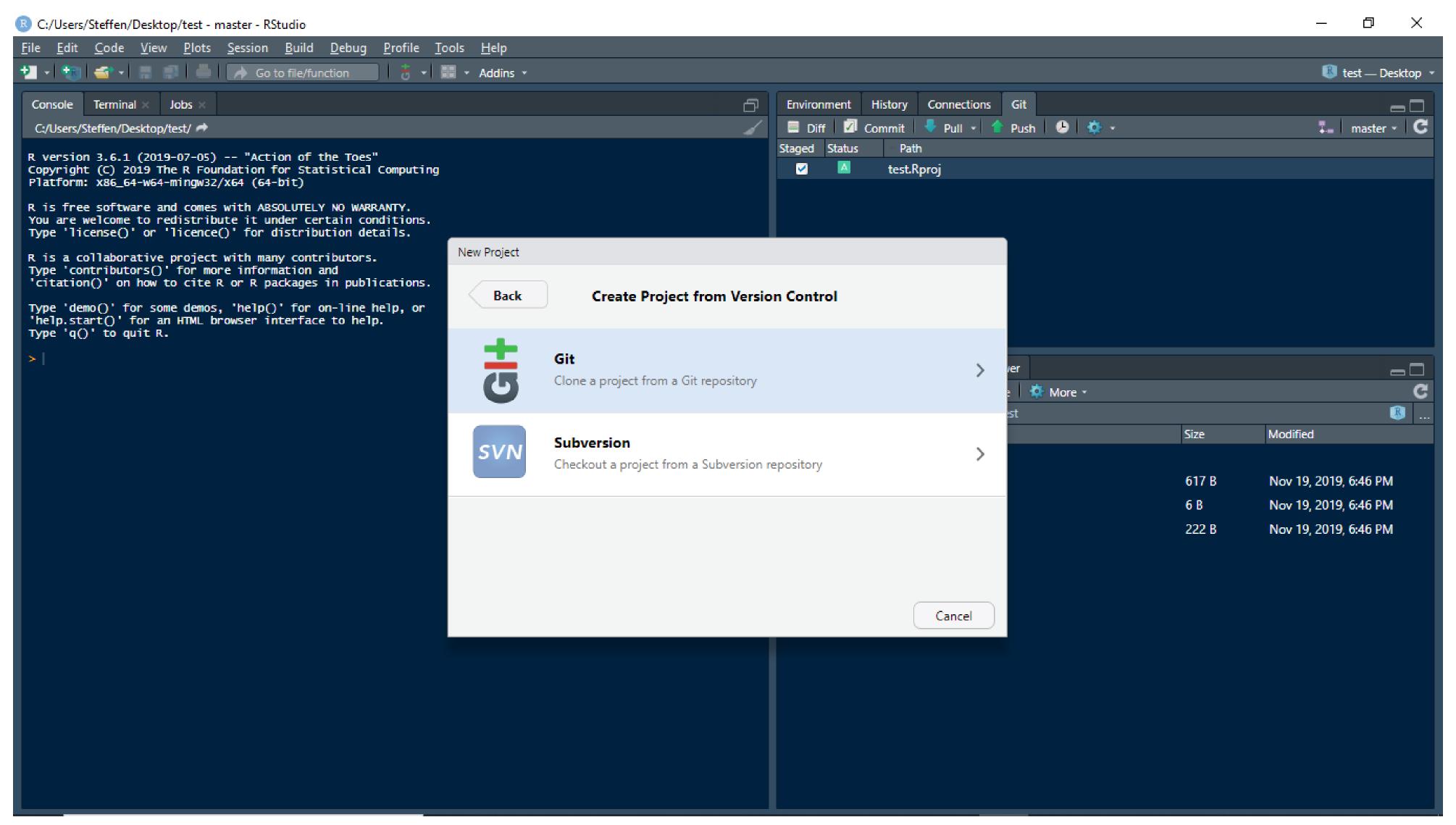


#### GITHUB REPOSITORY ALS PROJEKT IMPORTIEREN (1)

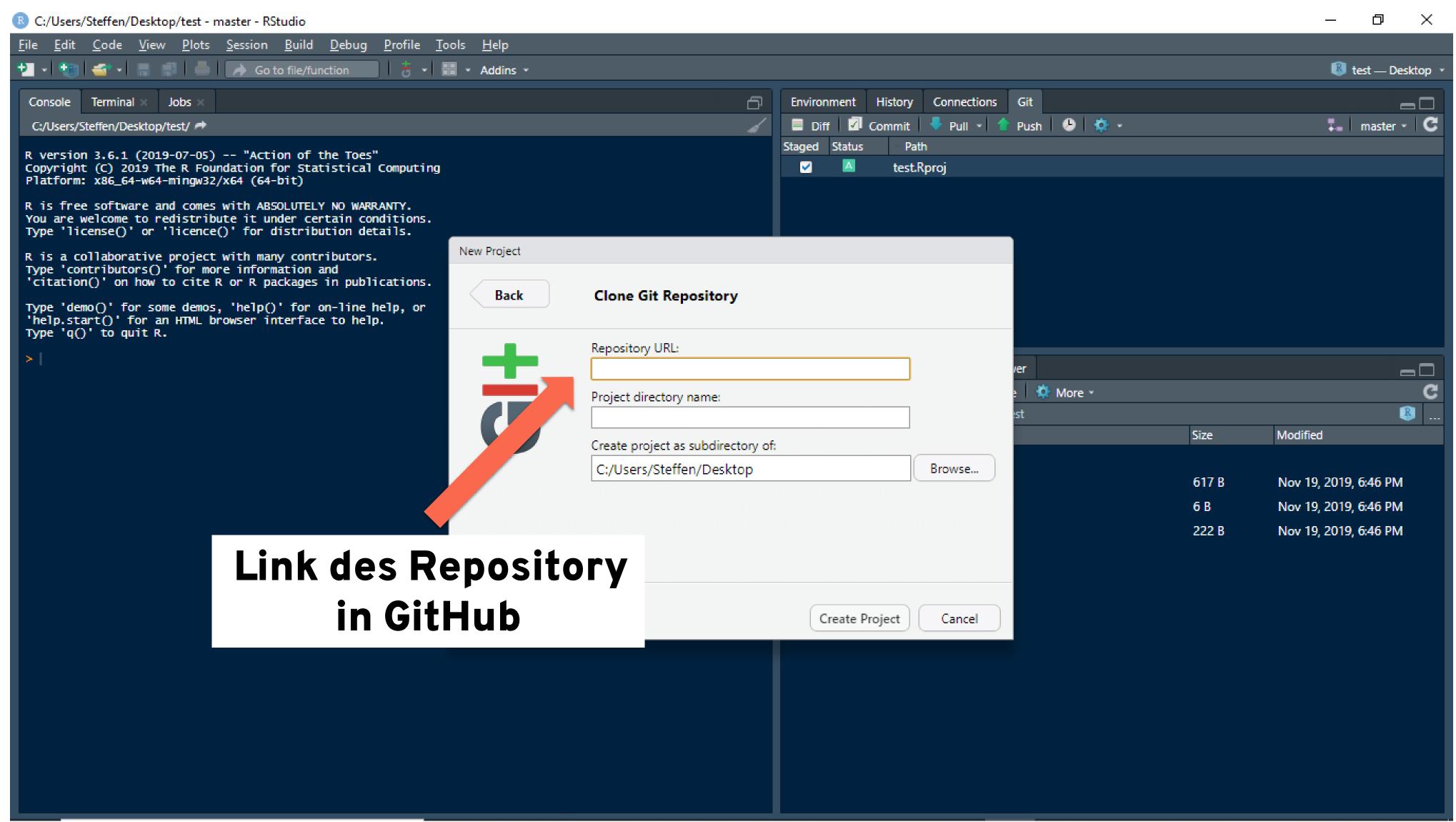


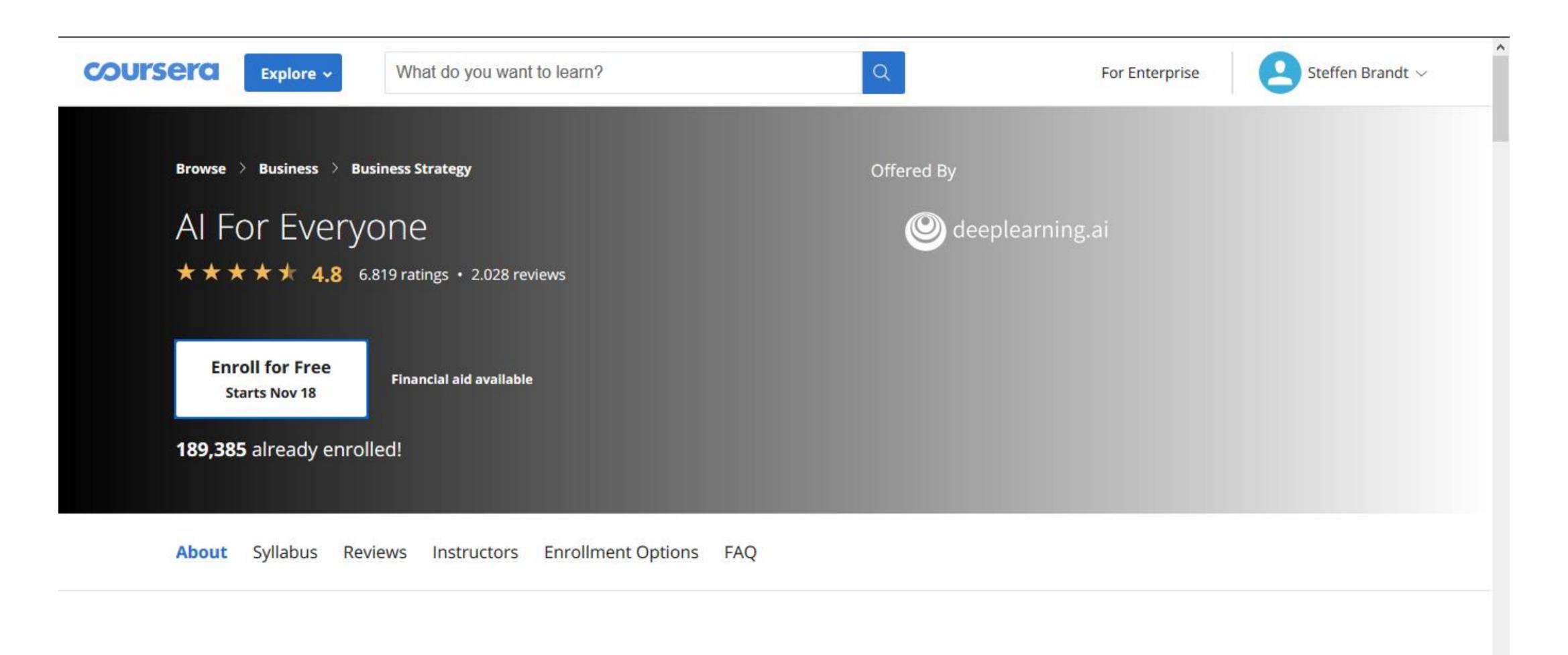


# GITHUB REPOSITORY ALS PROJEKT IMPORTIEREN (2)



# GITHUB REPOSITORY ALS PROJEKT IMPORTIEREN (3)





#### About this Course

624,390 recent views

All is not only for engineers. If you want your organization to become better at using Al, this is the course to tell everyone--especially your non-technical colleagues--to take.

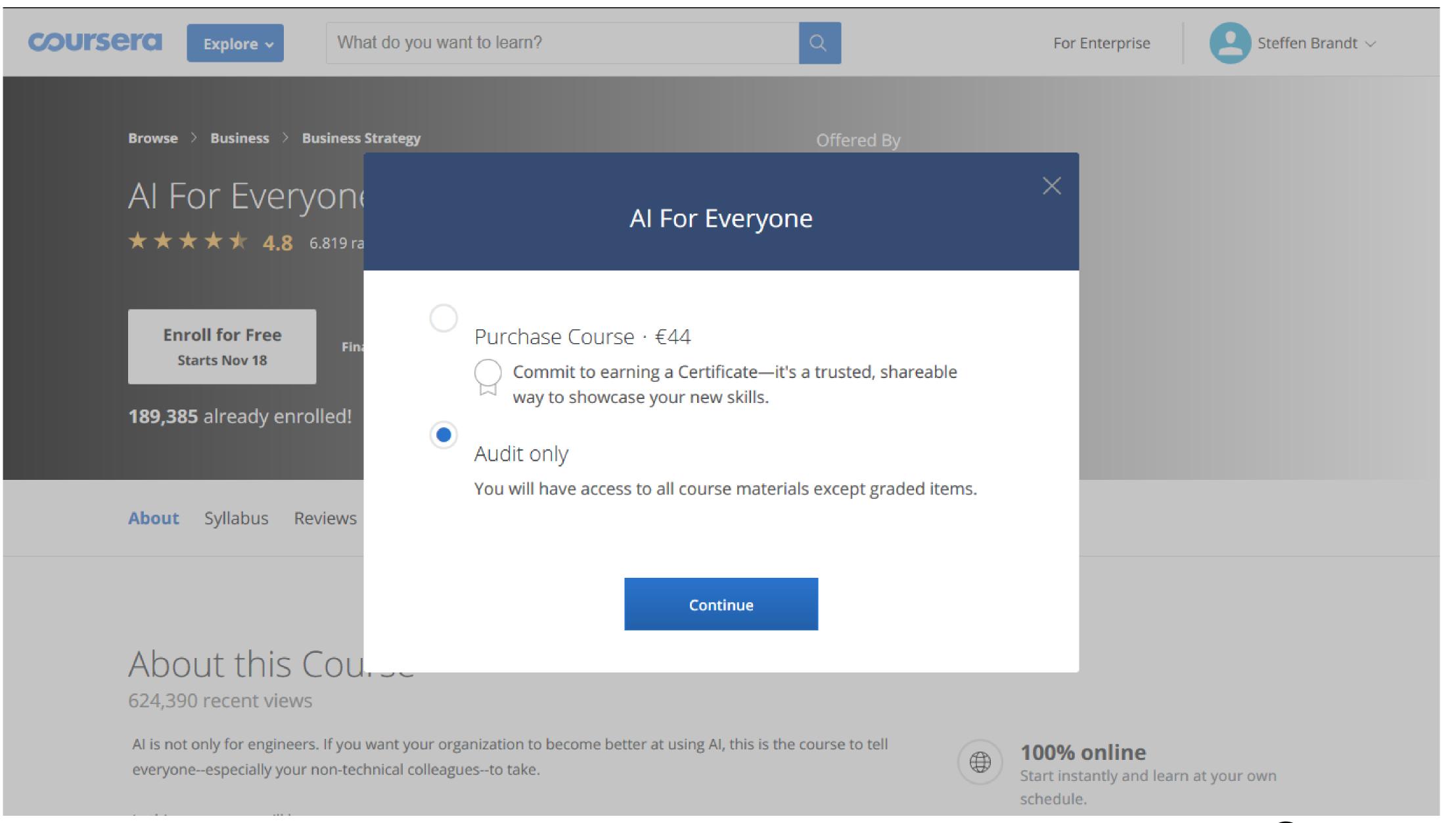


100% online

Start instantly and learn at your own schedule.

2 2 2







#### TYPEN DES MASCHINELLEN LERNENS

Supervised Learning

Unsupervised Learning

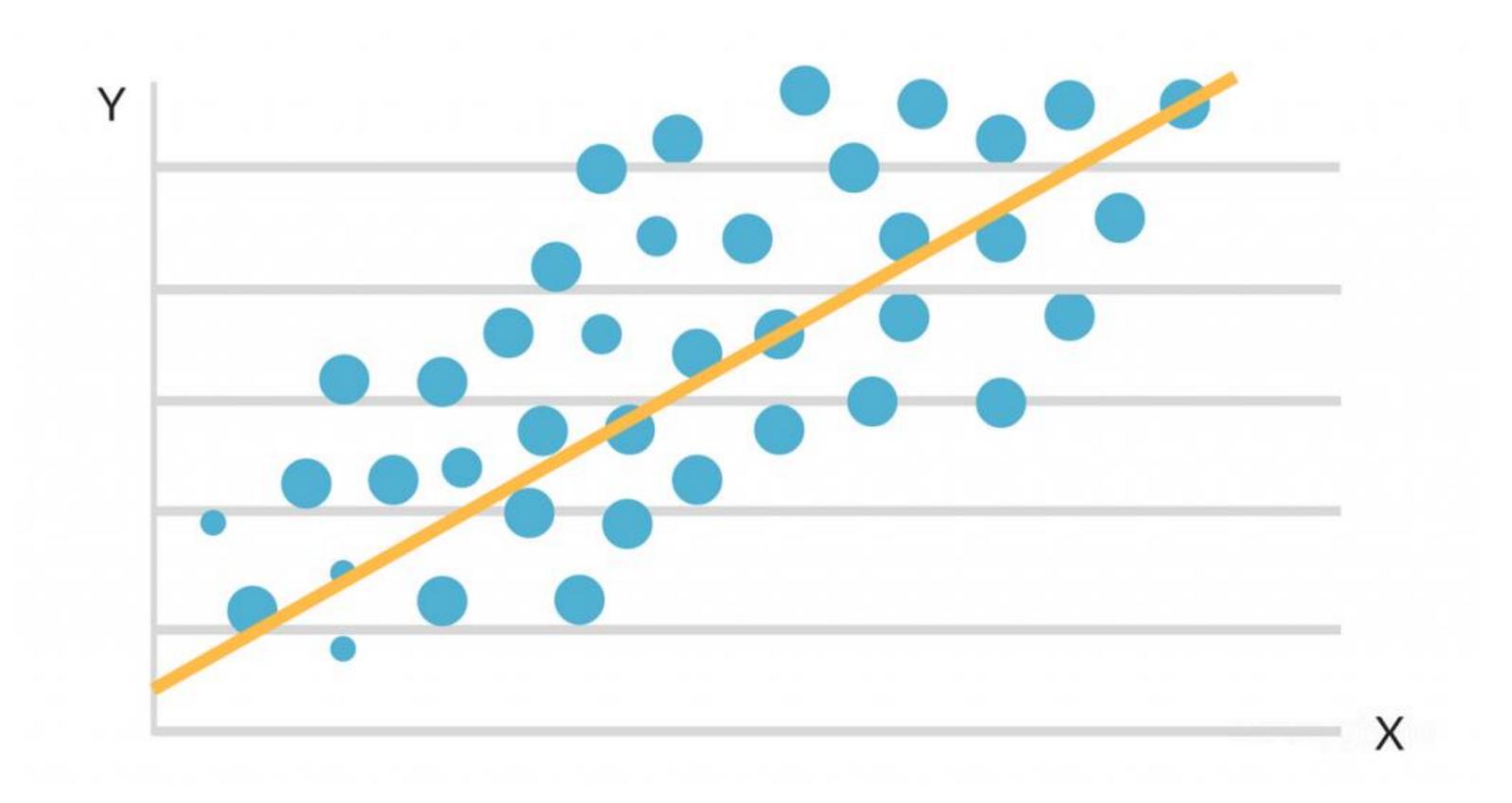
Self-Supervised Learning

Semi-Supervised Learning

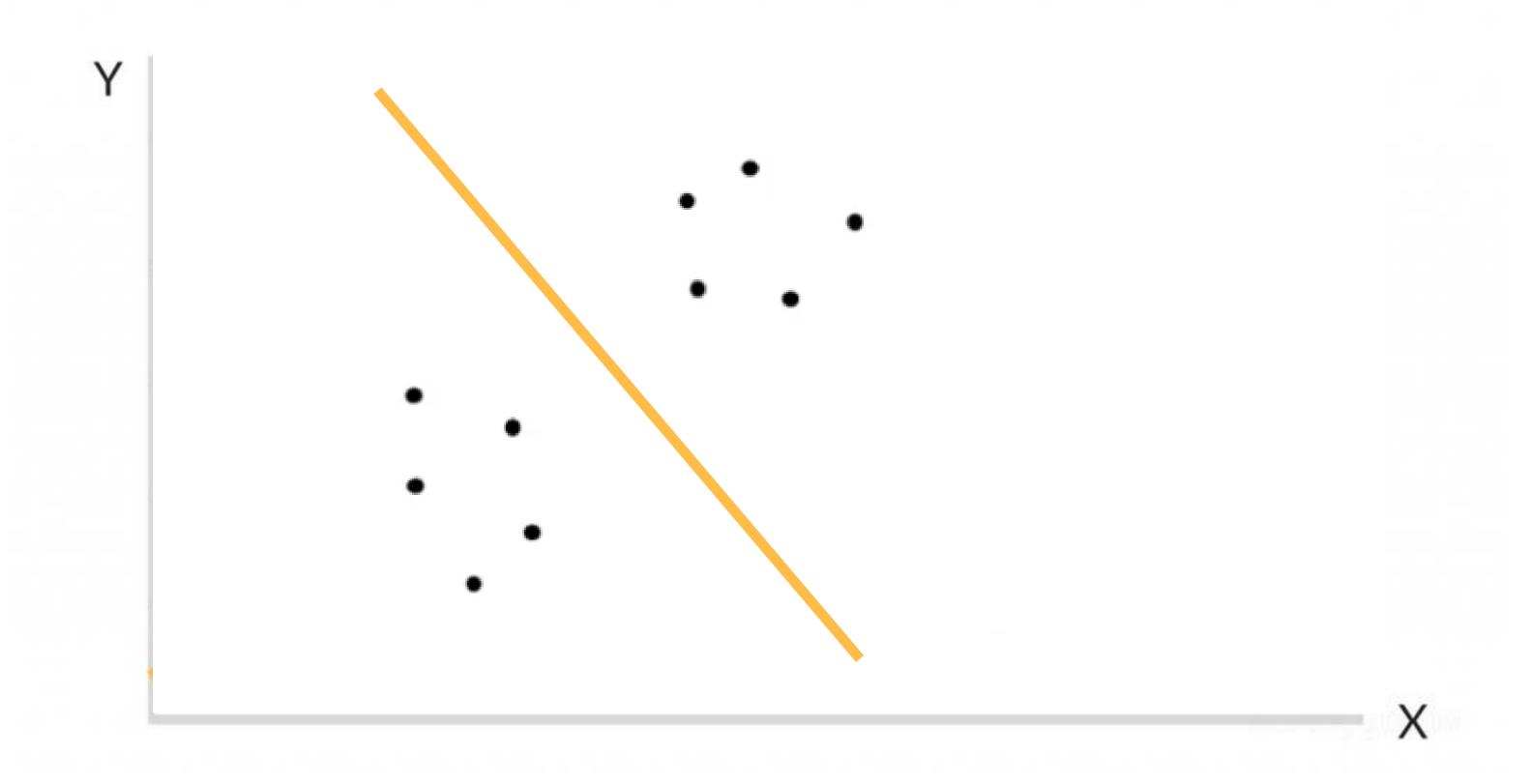
Reinforcement Learning



# SUPERVISED LEARNING



#### UNSUPERVISED LEARNING



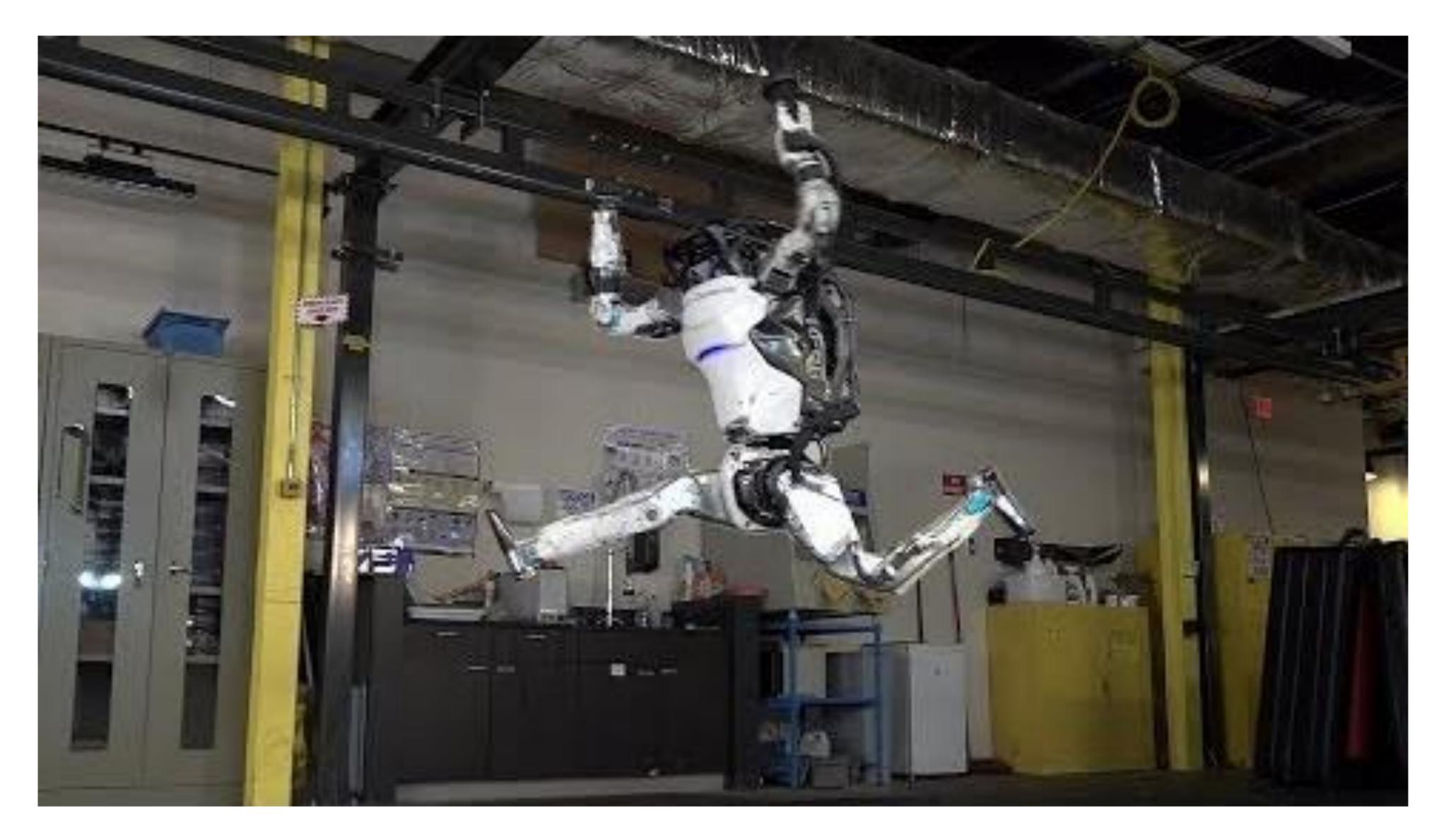
# BEISPIELE KÜNSTLICHER INTELLIGENZ



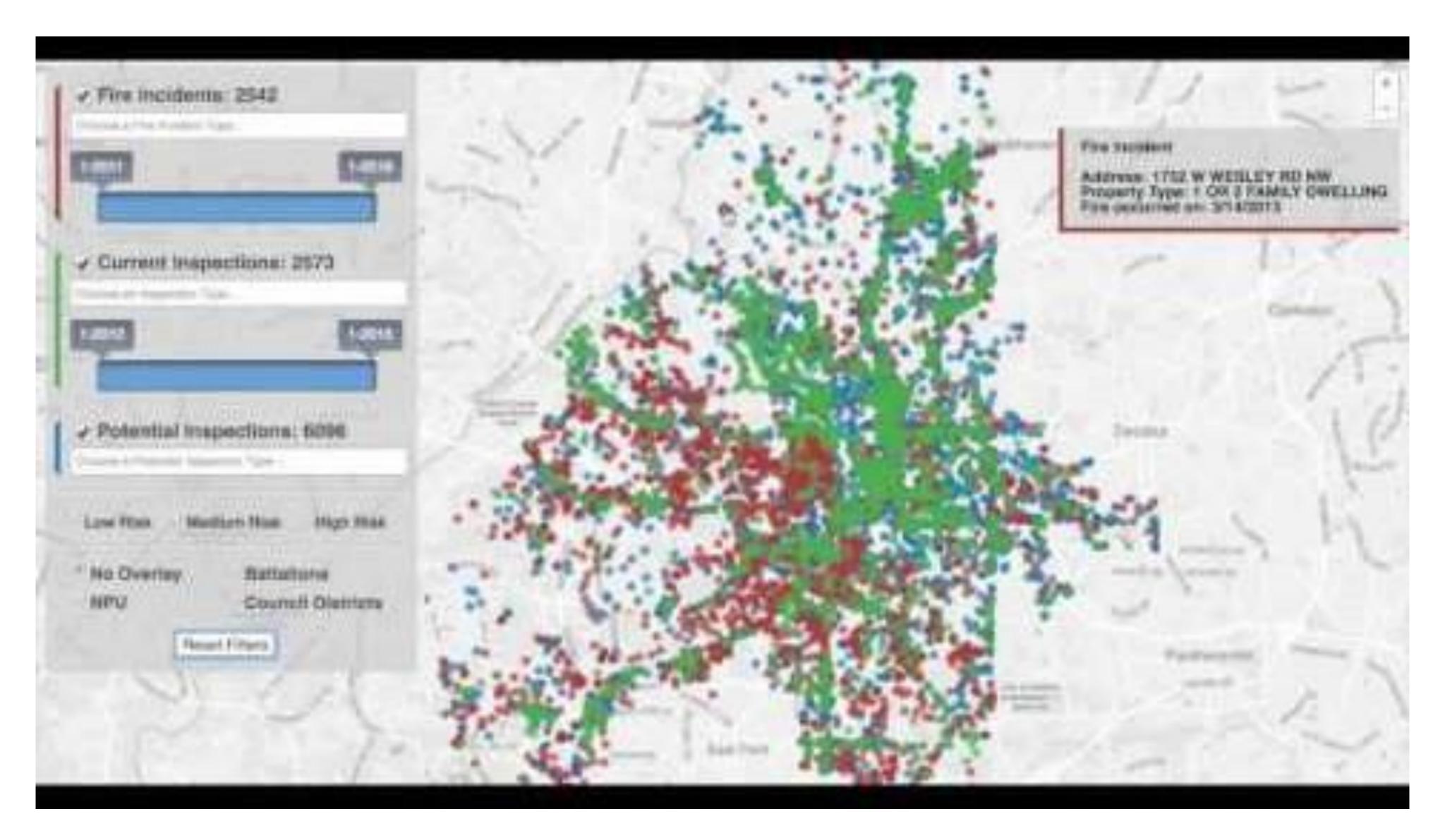
#### SCHACHCOMPUTER



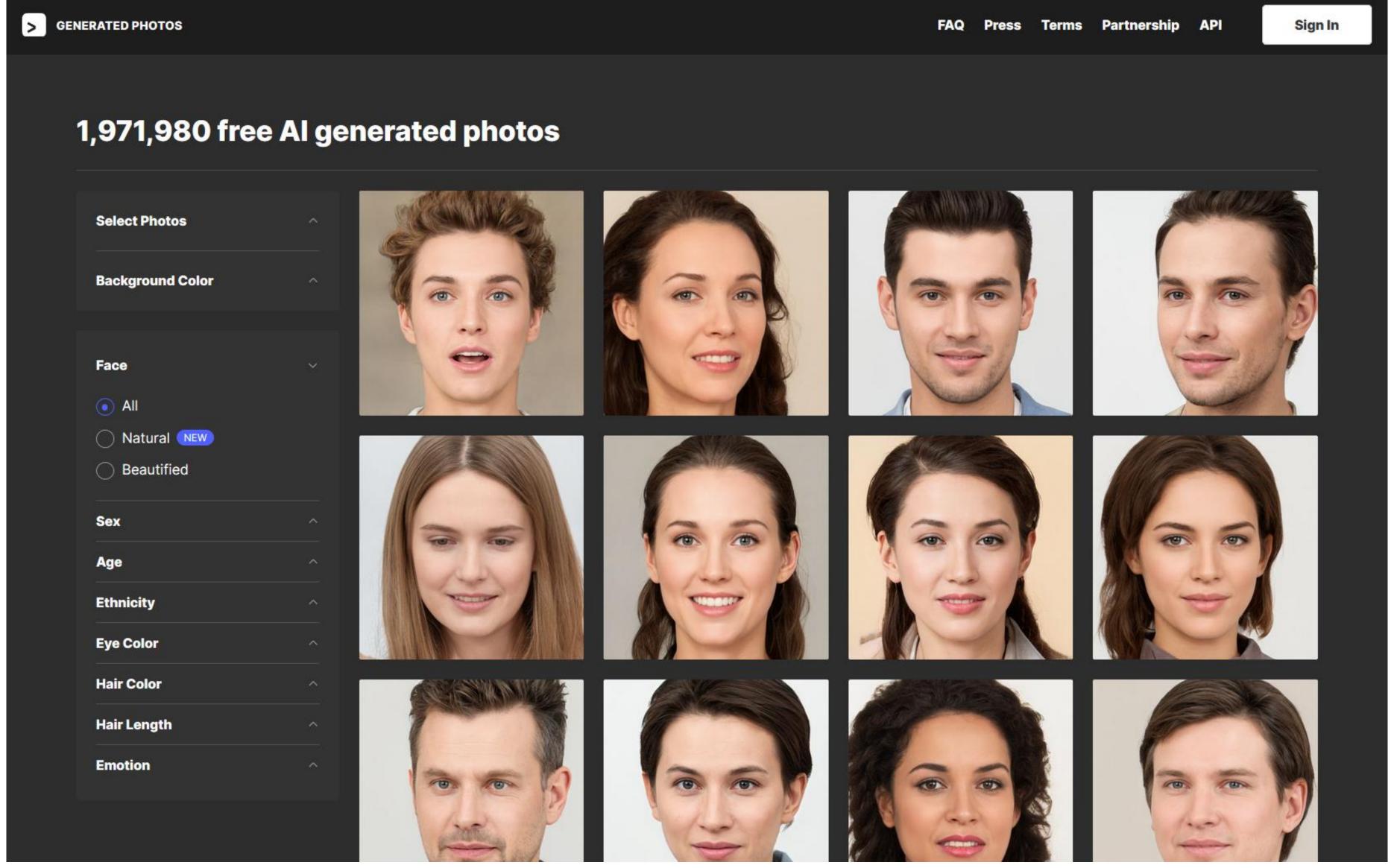
# ROBOTERSTEUERUNG



#### RISIKOPROGNOSEN



#### IMITATION



#### NATURAL LANGUAGE PROCESSING

GPT-3

https://mixed.de/gpt-3-diese-fuenfapps-zeigen-das-ki-potenzial/



#### AUFGABEN

- Trefft Euch einmal im Team, um die neu erstellten Variablen zur Umsatzvorhersage zu diskutieren und mögliche zusätzliche Ideen zu besprechen.
- Fügt alle ggf. neuen Daten (z.B. zusätzliche heruntergeladene oder selbst erstellte Listen) Eurem Team Repository hinzu, fügt die Programme und Funktionen, die neue Daten erstellen Eurem Team Repository hinzu, und fügt alle vorhandenen und erstellten Daten in einem Datensatz zusammen.
- Führt als Einführung zum Thema Künstliche Intelligenz die erste Woche des Kurses "Al for Everyone" durch:
  - https://www.coursera.org/learn/ai-for-everyone

