## Kurzbeschreibung:

Git ist ein Versionierungstool bzw. ein DVCS (Distributed Version Control System) und dient dazu, bei längeren Projekten alle Versionen zu verschiedenen Zeitpunkten sauber zu dokumentieren.

Um Teamarbeit zu ermöglichen oder verschiedene Ideen parallel zu entwickeln, kann man bei Git – ähnlich wie bei einem Baum – am Stamm (dem *main*- oder *master*-Branch) arbeiten und sich von dort aus verzweigen. Das bedeutet, man kann an unterschiedlichen Ästen bzw. Branches arbeiten (z. B. jedes Teammitglied an einem eigenen) und die Ergebnisse, die man behalten möchte, später wieder zusammenführen (mergen).

Das typische Vorgehen sieht folgendermaßen aus:  
Jedes Teammitglied arbeitet in einem eigenen Branch (Zweig). Änderungen und neue Dateien werden dabei zunächst lokal über Git gespeichert.  
Regelmäßig sollte der eigene Branch zur Einsicht und Zusammenarbeit in ein zentrales, gemeinsam genutztes Verzeichnis (Remote-Repository) hochgeladen werden. So können andere Teammitglieder Verbesserungsvorschläge einbringen oder notwendige Änderungen anregen.

Wenn der Branch ausreichend geprüft und für gut befunden wurde, kann er mit dem Hauptzweig (*main* oder *master*) zusammengeführt (gemergt) werden.  
Damit alle Teammitglieder auf dem aktuellen Stand bleiben, sollten sie regelmäßig den neuesten Stand des Haupt-Branches abrufen (pull).

Auf diese Weise entsteht im Laufe der Zeit ein gut strukturiertes Projekt. Es bleibt jederzeit möglich, auf ältere Versionen zuzugreifen oder sie wiederherzustellen.

Arbeitsweise:

Git speichert bei jedem *Commit* (Speichervorgang) eine Art Snapshot des gesamten Dateiordners bzw. des zu überwachenden Projekts. Diese Snapshots werden chronologisch geordnet im Repository abgelegt.

Um sinnvoll mit Git zu versionieren, ist es essenziell, jeden Commit mit einer kurzen, aussagekräftigen Beschreibung zu versehen. Diese sollte idealerweise alle Änderungen seit dem letzten Stand sowie ggf. die dahinterstehende Intention enthalten. So wird es möglich, später auf eine bestimmte Version zurückzublicken und nachzuvollziehen, was sich verändert hat und warum.

Wie bereits erwähnt, können in Git mehrere Personen – jeweils in ihrem eigenen Branch – unabhängig voneinander arbeiten. Auch in diesen Branches wird bei jedem Commit ein Snapshot erstellt und im Repository „abgelegt“. Im gemeinsamen Remote-Repository befinden sich somit die Snapshots aller Branches sowie des Haupt-Branches (*main* oder *master*) nebeneinander.

Beim Merge-Vorgang wird ein bestimmter Branch mit dem *main*-Branch zusammengeführt.

Jede\*r Mitarbeitende verfügt – nach einem *Pull* vom Remote-Repository – über eine vollständige, lokale Kopie des gesamten Projekts. Dadurch kann das Repository jederzeit auf den aktuellen Stand gebracht oder auf einen früheren Zustand zurückgesetzt werden.

## GIT einrichten mit GITHUB etc:

1. Melde dich bei GITHUB an. Das ist eine Internetseite, die es ermöglicht öffentlich zugängliche Repositorys anzulegen und zu verwalten. Benötigt wird dafür ein Username und eine Emailadresse.
2. Installiere GIT und öffne GITBASH (einfach unten in Suchleiste eingeben)
3. Nun Username und email (zwingend die gleichen wie in GITHUB angeben)
   1. git config --global user.name "Dein Name"
   2. git config --global user.email [deine@email.de](mailto:deine@email.de)
4. Prüfen ob alles passt
   1. git config –list
5. Nun verknüpfe dein lokales GIT mit GITHUB und ssh
   1. ssh-keygen -t ed25519 -C [deine@email.de](mailto:deine@email.de)
   2. enter drücken
   3. enter drücken
   4. cat ~/.ssh/id\_ed25519.pub
   5. alles inklusive mail kopieren
6. In GITHUB auf Profil > Einstellungen > SSH.. > add new SSH > Text einfügen und speichern
7. In GITBASH prüfen
   1. ssh -T [git@github.com](mailto:git@github.com)
   2. „yes“ eingeben

## Erstes Repo bauen und in GITHUB anzeigen lassen

Im allgemeinen möchte man mit GIT nur bestimmte Dateien tracken. Deshalb benötigen wir dafür einen Ordner in dem alle Dateien liegen. Darin werden wir nun unser Repo also einen (.git) Ordner legen. Das heißt auch, dass nur Dateien die in dem Ordner liegen getrackt werden können.

1. Optionen:
   1. Ziehe den Ordner per Drag and Drop in GITBASH
   2. cd dateipfad
2. git init
3. Im Ordner oben bei Ansicht ausgeblendete Elemente anzeigen und prüfen ob ein .git Ordner erscheint
4. git add . // „.“ Fügt alles aus dem Ordner zu GIT hinzu
5. git commit -m "Notiz" // erster Commit fügt den Ordner nun lokal zur Datenbank
6. Prüfe ob alles hinzugefügt wurde
   1. cd dateipfad
   2. ls -a //list files all zeigt also alle Dateien aus dem Repo

Jetzt erstellen wir das Repo in GITHUB

1. In GITHUB oben rechts auf „+“ dann „New repository“
2. Namen des Projektes angeben. Evtl selben wie den Ordnernamen
3. Privat oder public
4. Das README-Zeug nicht ändern
5. Create Repository
6. In GITBASH:
   1. git remote add origin [git@github.com:deinBenutzername/dein-repo-name.git](mailto:git@github.com:deinBenutzername/dein-repo-name.git)

Erster PUSH

1. git branch -M main
2. git push -u origin main // das -u verlinkt nun den lokalen Branch mit GITHUB

// In Zukunft reicht es git push zu schreiben

## Wichtige Begriffe:

**Modified:**

Dateien die geändert, aber noch nicht in lokaler Datenbank eingecheckt wurden.

**Repository**  
Ein Projektordner, der von Git verwaltet wird. Enthält alle Dateien und die gesamte Versionshistorie.

**Commit**  
Ein gespeicherter Versionsstand mit einer Nachricht, die beschreibt, was geändert wurde. Wird vorerst lokal gespeichert.

**Branch**  
Ein Entwicklungszweig, in dem unabhängig gearbeitet werden kann. Ermöglicht parallele Entwicklung an verschiedenen Funktionen oder Fehlerbehebungen.

**Merge**  
Das Zusammenführen von Änderungen aus einem Branch in einen anderen (z. B. in den Hauptbranch).

**Main (früher Master)**  
Der Hauptentwicklungszweig eines Projekts. Standardmäßig heißt er heute oft „main“.

**Clone**  
Herunterladen einer Kopie eines Remote-Repositories auf den lokalen Rechner.

**Push**  
Übertragen der lokalen Commits in das Remote-Repository, damit andere darauf zugreifen können.

**Pull**  
Aktualisieren des lokalen Repositories mit den neuesten Änderungen aus dem Remote-Repository.

**Remote Repository**  
Ein zentrales Repository auf einem Server (z. B. GitHub, Bitbucket), das von mehreren Personen genutzt wird.

**Staging Area**  
Ein Zwischenspeicherbereich, in dem Änderungen gesammelt werden, bevor sie committet werden.

**Conflict (Merge-Konflikt)**  
Ein Versionskonflikt, der entsteht, wenn zwei Änderungen dieselbe Stelle im Code betreffen und nicht automatisch zusammengeführt werden können.

**Checkout**  
Wechseln zu einem anderen Branch oder zu einem bestimmten Commit.

**Fork**  
Eine Kopie eines Repositories, meist um unabhängig daran arbeiten zu können (z. B. in Open-Source-Projekten).

**Pull Request (PR)**  
Ein Antrag zur Zusammenführung von Änderungen aus einem Branch in einen anderen, meist mit der Möglichkeit zur Überprüfung und Diskussion.

## Arbeitsablauf?

1. Ändere eine Datei (im Verzeichnisbaum)
2. Stelle selektiv eine Änderung in Staging Area um sie später zu comitten
3. Commit: Snapshot von Datei in Staging Area wird in GIT-Verzeichnis aufgenommen

Das heißt: wenn eine Datei verändert wurde, ist sie erst modified.

1. git pull //Nur nötig um Dateien vom Remote zu ziehen
2. git add . //bzw ein anderer add Befehl
3. git status //zeigt: neue/geaänderte Dateien, offene Commits
4. git commit -m „sinnvolle Notiz“ //staged
5. git push //pushed die Dateien auf das GITHUB-Repo
6. git log //zeigt commit-history

Sollte Git in einer Umgebung verwendet werden, die GIT bereits integriert, findet man die einzelnen Befehle idR als Symbole recht schnell über das git-Symbol.

Branchen und Mergen

## GIT-Bash/CMD-Befehle im Überblick:

|  |  |
| --- | --- |
| Befehl | Beschreibung |
| git init | Erstellt ein neues Git-Repository im aktuellen Verzeichnis |
| git clone <url> | Klont ein entferntes Repository (z. B. von GitHub) lokal |
| git status | Zeigt den aktuellen Status der Arbeitskopie (z. B. Änderungen, Staging-Bereich) |
| git add <datei> | Fügt eine Datei dem Staging-Bereich hinzu |
| git add . | Fügt alle Änderungen im Verzeichnis dem Staging-Bereich hinzu |
| git commit -m "Nachricht" | Speichert die Änderungen mit einer Commit-Nachricht |
| git log | Zeigt eine Historie aller Commits |
| git diff | Zeigt die Unterschiede zwischen Dateien/Commits |
| git branch | Zeigt alle lokalen Branches |
| git branch <name> | Erstellt einen neuen Branch |
| git checkout <name> | Wechselt zu einem anderen Branch |
| git checkout -b <name> | Erstellt und wechselt zu einem neuen Branch |
| git merge <branch> | Führt den angegebenen Branch in den aktuellen zusammen |
| git pull | Holt neue Änderungen vom Remote-Repository und merged sie |
| git push | Schickt lokale Commits ins Remote-Repository |
| git remote -v | Zeigt die verbundenen Remote-Repositories |
| git reset --hard <commit> | Setzt das Projekt hart auf einen früheren Commit zurück |
| git stash | Speichert ungesicherte Änderungen zwischen |
| git stash pop | Stellt gespeicherte Änderungen wieder her |
| git rm <datei> | Entfernt eine Datei aus dem Arbeitsverzeichnis und dem Staging-Bereich |
| git config --global user.name "Name" | Setzt den globalen Git-Benutzernamen |
| git config --global user.email "E-Mail" | Setzt die globale Git-E-Mail-Adresse |

## Arbeiten mit VS-Code

Visual Studio Code ist ein von Microsoft bereitgestellter Editor, welcher das arbeiten in verschiedensten Programmiersprachen ermöglicht. Anwendung findet es insbesondere um einen kurzen Überblick über Code zu bekommen oder schnelle Änderungen vorzunehmen.

Über GIT ist es dann noch möglich an einem Projekt über Pycharm oder Intellij etc. und VSCode zu arbeiten.

1. Falls man sich bei VSCode anmelden muss, am besten mit den gleichen Daten wie GITHUB.
2. In VSCode einfach links auf das GIT-Symbol klicken und der Anleitung folgen.

## Jupyter -Notebooks und R in VS-Code

Wie gesagt bietet VS-Code verschiedenste Möglichkeiten. So kann man auch Jupyter-Notebooks oder auch in R coden, wenn man die Plugins dafür installiert.

Um jetzt ein Jupyter-Notebook in VS-Code zu öffnen, brauchen wir zunächst einmal R oder R-Studio. (<https://cran.r-project.org/>)

Außerdem Python, am besten die neueste Version. (<https://www.python.org/downloads/windows/>)

In VS-Code:

1. R installieren
2. Python installieren

In R:

1. install.packages("IRkernel")
2. IRkernel::installspec()

In VS-Code:

1. pip install jupyter
2. pip install notebook

# CML mit GIT

CML verfügt bereits über integrierte Git-Funktionalitäten, erkennbar an der linken Seitenleiste (Sidebar).  
Die einfachste Möglichkeit, den eigenen Fortschritt in CML zu versionieren, ist das Anlegen eines Remote-Repositories, z. B. bei **Bitbucket** oder **GitHub**.

**Vorgehen in Bitbucket:**

1. Ein neues Repository anlegen
2. Die Repository-URL kopieren
3. Optional: Über ein Terminal (z. B. Git Bash) prüfen, ob Name und E-Mail korrekt hinterlegt sind:
   * git config --user.name
   * git config --user.email

**In CML:**

1. In das Verzeichnis wechseln, das versioniert werden soll
2. In der linken Sidebar auf das Git-Symbol klicken und **„Clone Repository“** auswählen
3. Die zuvor kopierte URL einfügen, Benutzername und Passwort (bzw. Token) eingeben

Daraufhin wird auf der gewählten Ebene ein neuer Ordner mit dem Namen des Repositories erstellt. Dieser Ordner wird nun von Git überwacht (getrackt).

Innerhalb dieses Ordners kann man über das Git-Symbol in der Sidebar alle Änderungen einsehen. Dort erscheinen alle Dateien als **staged**, **changed** oder **untracked**. Eine **History-Ansicht** aller bisherigen Commits ist ebenfalls verfügbar. Nicht commitete Dateien werden durch Warnhinweise hervorgehoben.

**Commit:**  
Im Feld **„Summary“** sollte eine kurze Beschreibung der vorgenommenen Änderungen eingetragen werden (Pflicht für den Commit, da Git sonst einen Fehler ausgibt).  
Optional kann im Feld **„Description“** genauer erläutert werden, warum bestimmte Änderungen vorgenommen wurden – besonders hilfreich bei komplexeren Projekten oder für eine saubere Dokumentation.

**Push & Pull:**  
In der oberen Menüleiste von CML finden sich unter dem Menüpunkt **Git** die Optionen **Push** und **Pull** sowie weitere Git-Funktionen.

Branch & Merge