

# GitHub als Daten-Repository

BMLD Informatik 2, Woche 4

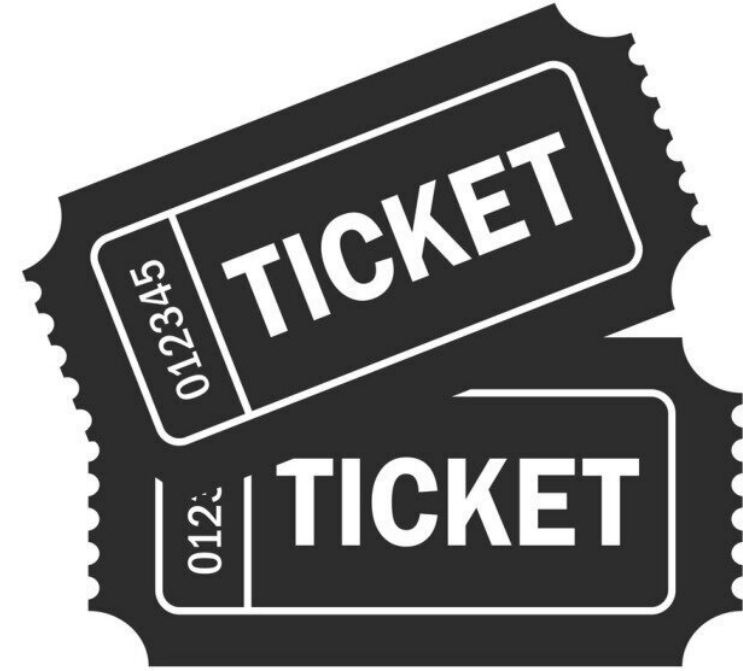
Samuel Wehrli

# Ziele Woche 4

- Du lernst, wie du Daten auf GitHub speichern kannst
- Du weisst, wie man einen Sicherheitsschlüssel (Token) für Github generiert
- Du kennst du verschiedenen read/write Methoden von GithubContents
- Du kennst das Konzept von Streamlit secrets
- Du weisst, wie der GitHub Token sicher gespeichert werden

# Zugriffstoken

Die Fahrkarte für deinen Datenzugriff



*myloview*

# Tokenbasierte Authentifizierung



Die tokenbasierte Authentifizierung ist ein Protokoll, mit dem Benutzer ihre Identität verifizieren können und im Gegenzug ein eindeutiges Zugriffstoken erhalten. Während der Nutzungsdauer des Security Tokens haben die Benutzer Zugriff auf die Website oder Anwendung, für die das Token herausgegeben wurde. Sie müssen dann nicht jedes Mal ihre Anmeldedaten erneut eingeben, wenn sie auf dieselbe Webseite, Anwendung oder Ressource zugreifen, die durch das betreffende Token geschützt wird.

**Authentifizierungstoken funktionieren wie ein gestempelter Fahrschein.** Der Benutzer hat Zugriff, solange das Token gültig ist. Sobald sich der Benutzer abmeldet oder eine Anwendung beendet, wird das Token entwertet.

# Erstellen eines Tokens (Repetition)

- **Privates** GitHub Repo für die Daten
- Massgeschneiderter Token (Fahrkarte)
  - Nur für das Daten-Repo
  - Nur für lesen und schreiben von Dateien
  - Zeitlich beschränkt



# Der Token

```
github_pat_11APCiWMY0jsz5AzMh50j5_  
lrQBVJuk1NL7pW2um4GAScPRxZFT4V77  
henEMj1FomjGQLH7QW6u1gRihwn
```

- pat = „personal access key“
- 80 Zeichen mit ca. 60 Kombination (Klein-, Grossbuchstaben, Ziffern)  
 $80^{60} \rightarrow 10^{114}$  Kombination
- Anzahl Atome im Universum ca.  $10^{80}$



# Tokens sind „Secrets“

bzw. wie sichergestellt werden kann, dass  
die Fahrkarte nicht in falsche Hände kommt.

**Achtung! Taschendiebe**



**Attention! Pickpockets**

# Umgang mit Tokens und Login-Daten

- Nie im Code hinterlegen
- Nie auf Github hinterlegen
- Als Secret hinterlegen
- Dokumentation: Streamlit [Secrets Management](#)



# Login-Daten im TOML-Format

- TOML:  
Tom's Obvious, Minimal Language
- Check out: <https://toml.io/en/>

```
# This is a TOML document
```

```
title = "TOML Example"
```

```
[owner]
```

```
name = "Tom Preston-Werner"
```

```
dob = 1979-05-27T07:32:00-08:00
```

```
[database]
```

```
enabled = true
```

```
ports = [ 8000, 8001, 8002 ]
```

```
data = [ ["delta", "phi"], [3.14] ]
```

```
temp_targets = { cpu = 79.5, case = 72.0 }
```

```
[servers]
```

```
[servers.alpha]
```

```
ip = "10.0.0.1"
```

```
role = "frontend"
```

```
[servers.beta]
```

```
ip = "10.0.0.2"
```

```
role = "backend"
```

# TOML und Streamlit

## TOML

```
[github]
owner = "samuelwehrli"
repo = "MyContactsApp-Data"
token="github_pat_11APCVWM...
```

## Zugriff in Streamlit

```
import streamlit as st

owner=st.secrets["github"]["owner"]
repo=st.secrets["github"]["repo"]
token=st.secrets["github"]["token"]
```

# Secrets auf deinem Rechner

... werden in Verzeichnis „.streamlit“ in der Datei secrets.toml gespeichert.

```
your-LOCAL-repository/  
├── .streamlit/  
│   ├── config.toml  
│   └── secrets.toml # Make sure to gitignore this!  
├── your_app.py  
└── requirements.txt
```

# Secrets auf der Streamlit Cloud

The image shows the Streamlit Cloud interface for a user named 'samuelwehrli'. The main area lists two applications: 'addressbook · main · addressbook\_app.py' and 'digipall · main · digipall\_app.py'. A red circle highlights the three-dot menu icon next to the first application, and a blue arrow points from this menu to the 'Settings' option at the bottom. The 'Settings' option is also circled in red. To the right, an inset shows the 'App settings' modal with three tabs: 'General', 'Sharing', and 'Secrets'. Blue arrows point to the 'General' and 'Secrets' tabs, with the text 'Passender URL wählen' (Choose appropriate URL) and '«Secrets» eingeben' (Enter «Secrets») respectively. The 'App URL' section shows a custom subdomain 'samuelwehrli-addressbook-addressbook-app-38v7c' followed by '.streamlit.app'.

Analytics Settings

samuelwehrli's apps

New app

addressbook · main · addressbook\_app.py

digipall · main · digipall\_app.py

Copy URL

Favorite

Analytics

Reboot

Delete

Settings

App settings

General

Sharing

Secrets

Passender URL wählen

«Secrets» eingeben

App URL

Pick a custom subdomain for your app's URL. The default URL is based on the app's location in GitHub.

samuelwehrli-addressbook-addressbook-app-38v7c .streamlit.app

Save

Lesen und schreiben auf Github  
mit `github_contents.py`

# How to use the package

```
from github_contents import GithubContents
github = GithubContents("your username", "your repo", "your token")

# writing to github
github.write_text("github/filepath", "your text", "your commit message")
github.write_json("github/filepath", {"your": "json"}, "your commit message")
github.write_df("github/filepath", df, "your commit message")

# reading from github
text = github.read_text("github/filepath")
dict_or_list = github.read_json("github/filepath")
df = github.read_df("github/filepath")
```

# Commit Message

Use good commit messages

## Activity

 All branches ▾

 All activity ▾

 All users ▾

 All time ▾

### Add contact 'Name' to the file MyContactsTable.csv

 samuelwehrli pushed 1 commit to [main](#) • 525f28f...1672cbe • 9 hours ago

### Add Samuel Wehrli to the MyContactsTable.csv

 samuelwehrli pushed 1 commit to [main](#) • 044d0b6...525f28f • 10 hours ago

### subfolder test

 samuelwehrli pushed 1 commit to [main](#) • 2248ec5...044d0b6 • 11 hours ago

### csv test

 samuelwehrli pushed 1 commit to [main](#) • 753cede...2248ec5 • 11 hours ago

### csv test

 samuelwehrli pushed 1 commit to [main](#) • 9d5b960...753cede • 11 hours ago

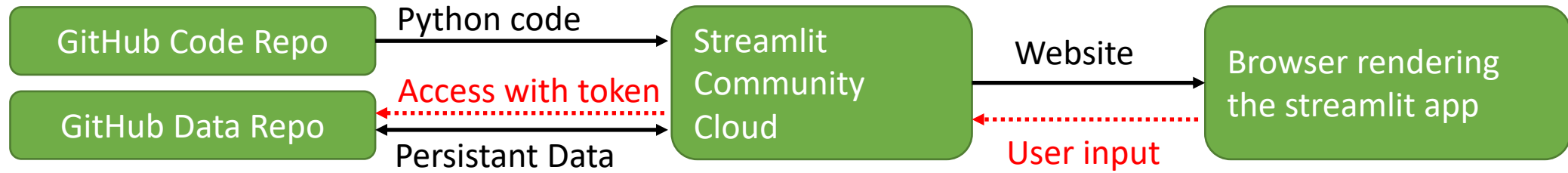
### json test

 samuelwehrli pushed 1 commit to [main](#) • 6003e08...9d5b960 • 11 hours ago

### text test

 samuelwehrli pushed 1 commit to [main](#) • 44f7d14...6003e08 • 11 hours ago

# Streamlit App in der Cloud



<https://mycontactsapp-w04.streamlit.app>



# Disclaimer

- GitHub ist eigentlich nicht als Datenbank gedacht
- Für kleinere Datensätze ist es okay
- Vorteil: Einfach und transparent

Für grössere Datenmengen → Datenbanken

<https://docs.streamlit.io/knowledge-base/tutorials/databases>

Wir unterstützen Mutige gerne!

# Vorbereitung Woche 5

# Vorbereitung Woche 5

1. App aus der Vorlesung in die Cloud laden mit eigenem Daten-Repo  
Die App darf gerne modifiziert werden – oder eine eigene App darf hochgeladen werden, welche aber Datenspeicherung enthalten muss.
2. Frage zum Code: `mycontacts_app_w04.py`

Kein Frageblatt, siehe Quiz!