

Automate the boring stuff

Belevingsdagen 2024

Wie ben ik?

- Elke Boonen
- Informaticus
- Docent programmeren
- Opleidingscoördinator ADI/IMS
- Praktijklector educatieve master Informatica KUL
- Moeder van 4
- Dol op avontuur
- Fervente lezer (& schrijver)
- Met een hoofd vol ideeën of chaos ;)





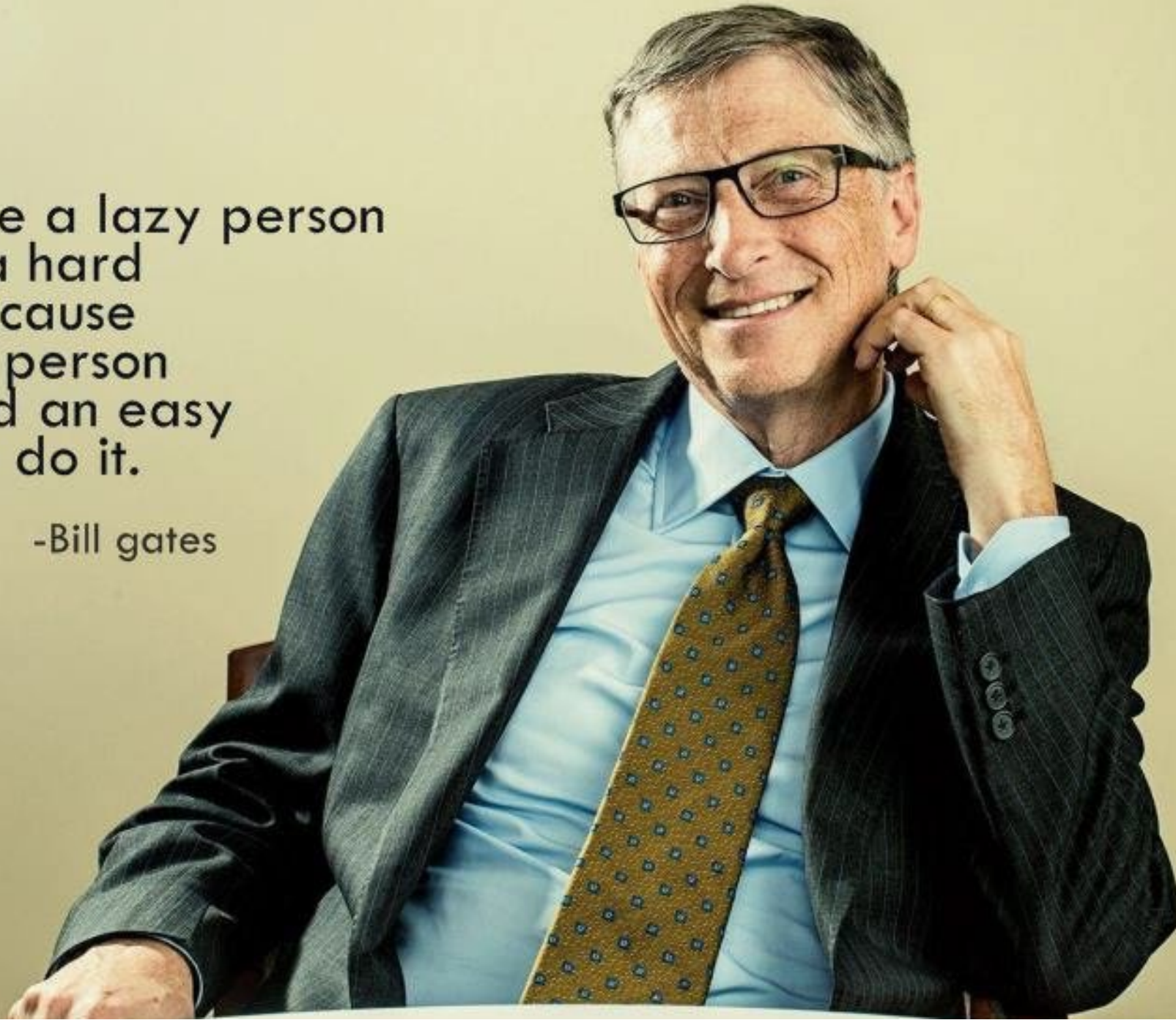
Wat is automatiseren?

Automatisering is het uitvoeren
van taken met minimale
menselijke tussenkomst, maw we
reduceren het aantal kliks!



I choose a lazy person
to do a hard
job. Because
a lazy person
will find an easy
way to do it.

-Bill gates

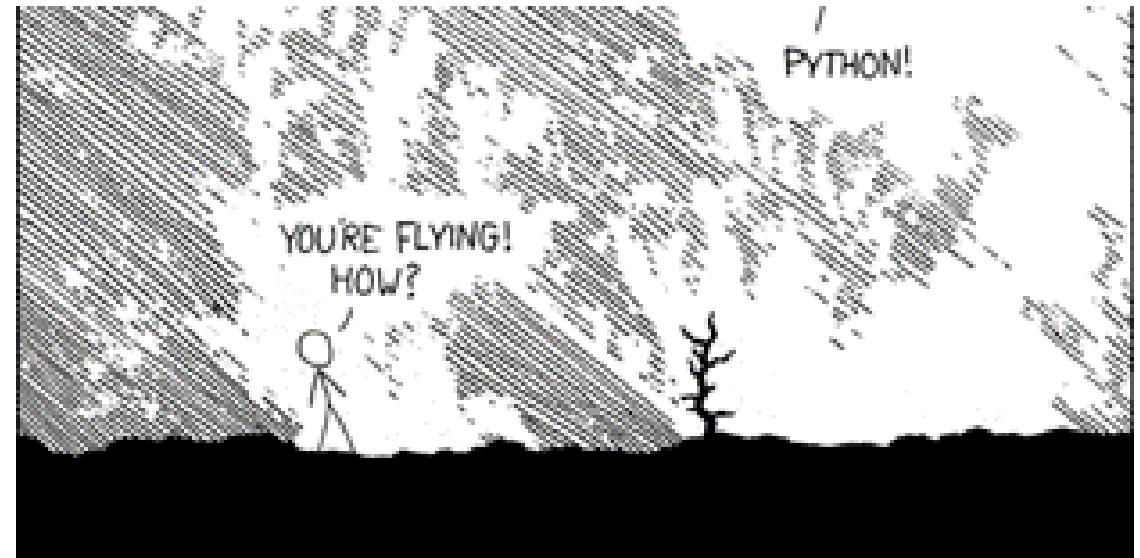


Waarom zou je automatiseren?

- **Efficiëntie verhogen**
 - *Repetitieve taken automatiseren*
 - *Meer tijd voor waardevolle taken*
- **Minder fouten**
 - *Machines werken consistent (geen vermoeidheid of afleiding)*
- **Veiligheid verbeteren**
 - *Machines kunnen taken uitvoeren die gevaarlijk zijn*

Automatiseren met Python

- Scripttaal met **eenvoudige** syntax
- **Veelzijdig** door bibliotheken voor bestandsbeheer, web scraping, AI & ML, gegevensanalyse, e-mail, lasercutting enz....
- **Open source**
- Python is dus perfect voor automatiseren!



Wat gaan we automatiseren?

- In één klik...
 - Informatie van Wikipedia
 - Samenvatting van de tekst
 - Grootste/beste foto
 - E-mail samenvatting & foto



WIKIPEDIA Doorzoek Wikipedia

Einstein

Duits-Zwitserse-Amerikaanse theoretisch natuurkundige en uitvinder (1879–1955)

PDF downloaden Volgen Brontekst

Albert Einstein (doorverwijspagina) voor andere betekenissen van Albert Einstein.

Einstein (Ulm, 14 maart 1879 – Princeton (New Jersey), 18 april 1955) was een Duits-Zwitserse-Amerikaanse theoretisch natuurkundige van Joodse afkomst. Hij wordt algemeen gezien als een van de belangrijkste natuurkundigen uit de geschiedenis, naast Isaac Newton en James Clerk Maxwell. Zelf noemde hij altijd Newton als een veel belangrijker natuurkundige dan zichzelf omdat Newton, anders dan Einstein, ook grote experimentele ontdekkingen deed. In het dagelijks leven is de naam Einstein synoniem geworden met intelligentie.^[1]

oud

Levenswerk

Einstein werd vooral bekend vanwege de twee relativiteitstheorieën: de speciale relativiteitstheorie van 1905 en de algemene relativiteitstheorie van 1915 en volgde de speciale relativiteitstheorie uitbreidt door ook plaats in te ruimen voor de zwaartekracht. Hij publiceerde meer dan 300 wetenschappelijke en meer dan 150 niet-wetenschappelijke werken.^[2] In zijn latere jaren schreef Einstein uitvoerig over filosofische kwesties onderwerpen.^[3] Hij wordt vaak samen met Max Planck beschouwd als de belangrijkste figuren van de moderne natuurkunde.

Albert Einstein
14 maart 1879 - 18 april 1955

A portrait of Albert Einstein, showing him from the chest up, wearing a suit and tie, with his characteristic wild hair and mustache.

Wat gaan we automatiseren?



Informatie downloaden



Mappen & bestanden doorzoeken



Informatie samenvatten (met API)



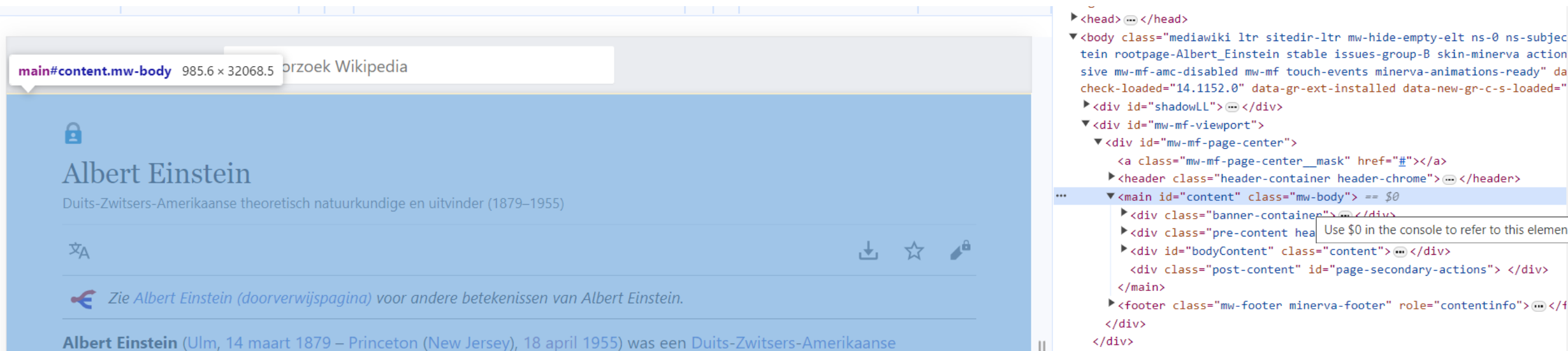
Automatisch e-mailen

Informatie downloaden

- **Webscrapen**
- **Automatisch verzamelen** van informatie van websites
 - Informatie = tekst, afbeeldingen en tabellen
 - Verschillende doeleinden, bv marktonderzoek, gegevensanalyse...
- **Wat moet je hiervoor kennen?**
 - HTML: hoe zit een webpagina in mekaar?
 - Python

De bron van een webpagina

CTRL+Shift+I/J om pagina bron te inspecteren - Elements



De bron van een webpagina

IMG-tag duidt altijd op afbeeldingen, zowel voor foto's als ondersteunende afbeeldingen

```
<!DOCTYPE html>
<html>
<head>
  <title>Images Example</title>
</head>
<body>
  <h1>My Images</h1>
  <div id="content">
    <!-- JPG Image -->
    
    <!-- GIF Image -->
    
  </div>
</body>
</html>
```

id/class in een tag maken het makkelijk om bepaalde informatie te zoeken

Tekst scrapen

```
1 import requests
  from bs4 import BeautifulSoup

url = "url naar de webpagina" 2

response = requests.get(URL)
soup = BeautifulSoup(response.content, "html.parser") 3

text = soup.find(id="mw-content-text").get_text()
text = text.replace("\n", " ") 4

title = soup.title.string 5

print(title)
print(text) 6
```

1. Importeer alle packages
2. Adres van webpagina en naam van map voor afbeeldingen
3. Request naar webpagina én haal alle html binnen
4. Zoek de tag met alle info én verwijder alle lege lijnen
5. Lees de titel
6. Print alle info

Afbeeldingen scrapen

```
1 import requests
  from bs4 import BeautifulSoup
  from urllib.parse import urljoin, urlparse
  import os

url = "url naar de webpagina"
folder = "./images/"

3 response = requests.get(url)
  soup = BeautifulSoup(response.content, "html.parser")

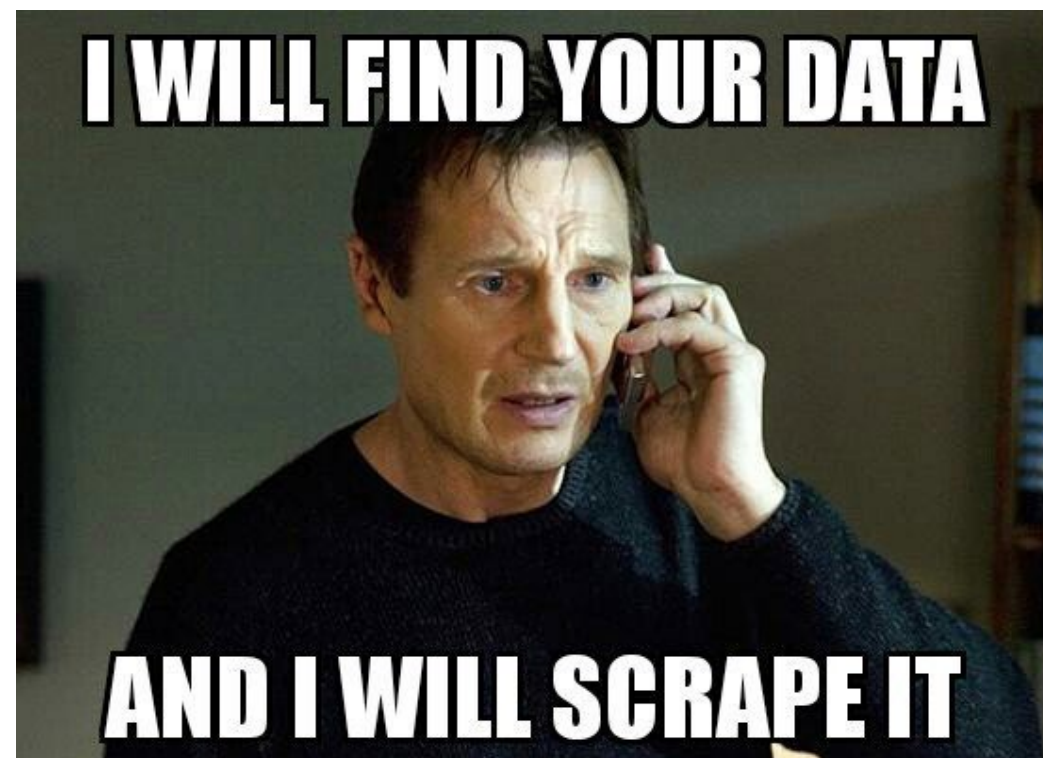
  for img_tag in soup.find_all("img"):
4      img_src = img_tag.get("src")
5      filename = os.path.basename(urlparse(img_src).path)

      with open(os.path.join(folder, filename), "wb") as f:
6          img_url = urljoin(url, img_src)
          response = requests.get(img_url)
          f.write(response.content)
          print(f"Downloaded {filename}")
```

1. Importeer alle packages
2. Adres van webpagina en naam van map voor afbeeldingen
3. Request naar webpagina én haal alle html binnen
4. Loop over alle img-tags
5. Lees source en maak er bestandsnaam van
6. Download afbeelding

Is webscrapen legaal?

- **Grijze zone**, ingewikkeld onderwerp!
- Kan in strijd zijn met
 - **Regels** van website-eigenaren (toestemming vragen!)
 - **Auteursrechten**
 - Privacywetgeving (**GDPR**)
 - enz...



Wat gaan we automatiseren?



Informatie downloaden



Mapperen & bestanden doorzoeken



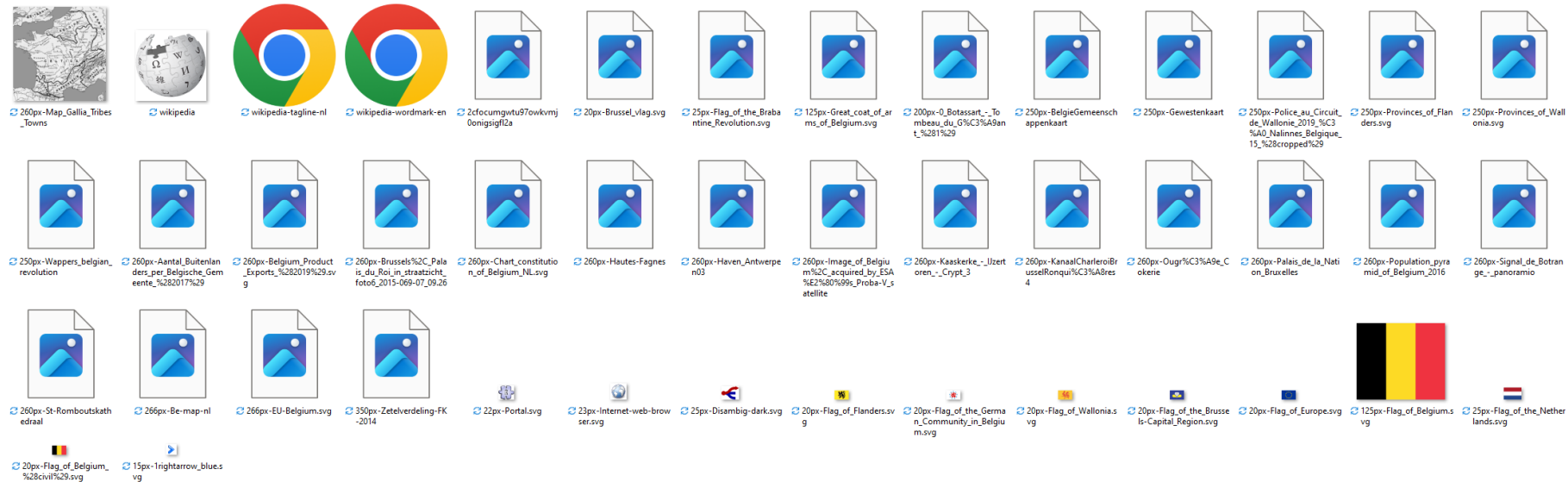
Informatie samenvatten (met API)



Automatisch e-mailen

Grootste afbeelding zoeken

- Niet alle afbeeldingen zijn interessant!
- Meestal is de grootste afbeelding die je hebt gedownload ook de meest interessante!



Grootste afbeelding zoeken

```
import os
```

1

```
FOLDER = "./images/"
```

```
afbeelding = None
```

```
max_size = 0
```

2

```
for file in os.listdir(FOLDER):
```

```
    file_path = os.path.join(FOLDER, file)
```

3

```
    size = os.path.getsize(file_path)
```

```
    if size > max_size:
```

```
        afbeelding, max_size = file_path, size
```

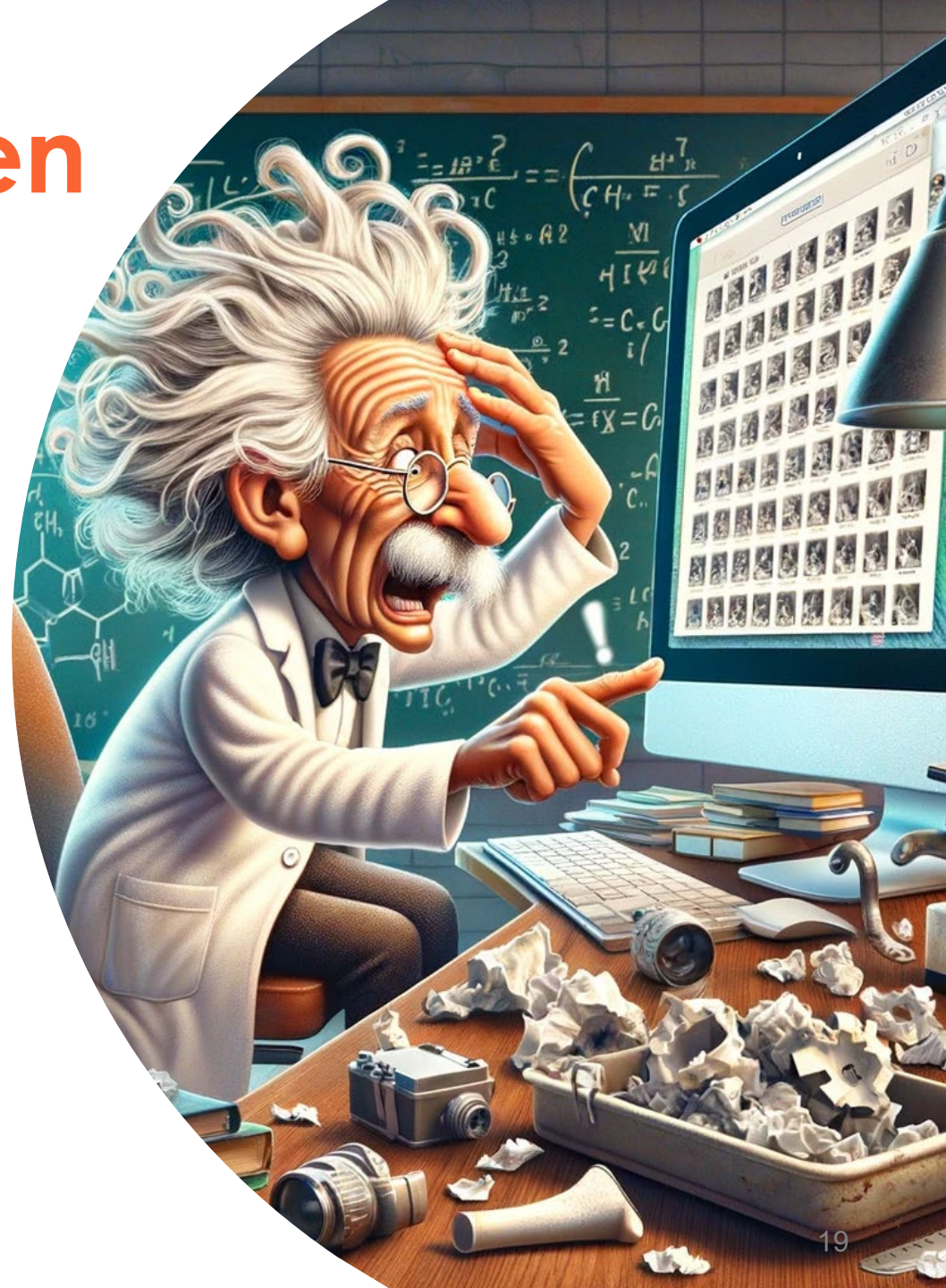
```
print(f"Grootste bestand '{afbeelding}' behouden.")
```

4

1. Importeer alle packages
2. Naam van map met afbeeldingen, variabelen maken waarin we naam afbeelding bijhouden
3. Loop over alle files, neem de locatie van elke file
4. Vergelijk de grootte van de afbeelding met de grootste tot nu toe én als die groter is, heb je een nieuwe grootste!

Afbeeldingen verwijderen

- Je kan ook snel alle afbeeldingen ineens verwijderen!
- Pas doen nadat e-mail is verstuurd natuurlijk ;)



Afbeeldingen verwijderen

```
import os ①  
FOLDER = "./images/" ②  
  
for file in os.listdir(FOLDER): ③  
    file_path = os.path.join(FOLDER, file)  
    ④ os.remove(file_path)  
  
print("Alle bestanden zijn verwijderd.")
```

1. Importeer alle packages
2. Naam van map met afbeeldingen
3. Loop over alle files, neem de locatie van elke file
4. Verwijder elk bestand in map

Wat gaan we automatiseren?



~~Informatie downloaden~~



~~Mappen & bestanden doorzoeken~~

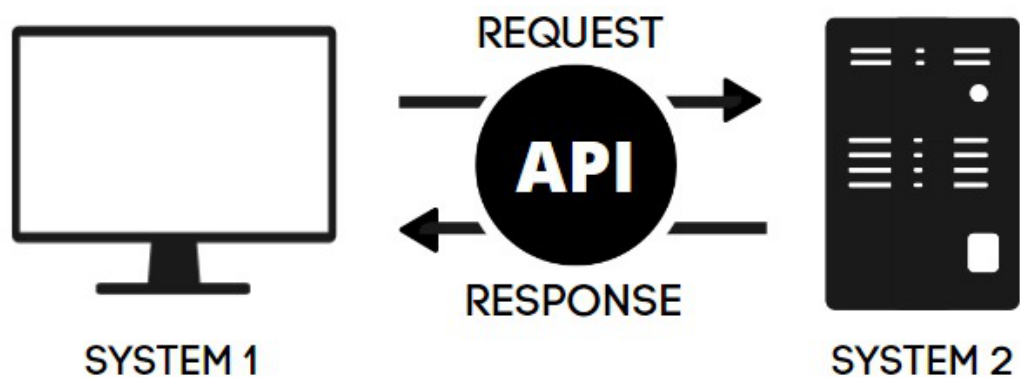


Informatie samenvatten (met API)



Automatisch e-mailen

Wat is een API?



How API Works?

- **Application Programming Interface**
- Een manier om **informatie & diensten** uit te **wisselen** over het web
- Je stuurt een **request (vraag)** naar een API en krijgt altijd een **response (antwoord)** terug
- Zijn er veel API's? Ja! Iedereen heeft er: google, spotify, twitter, ChatGPT... er zijn API's om memes te maken, locaties op te vragen...

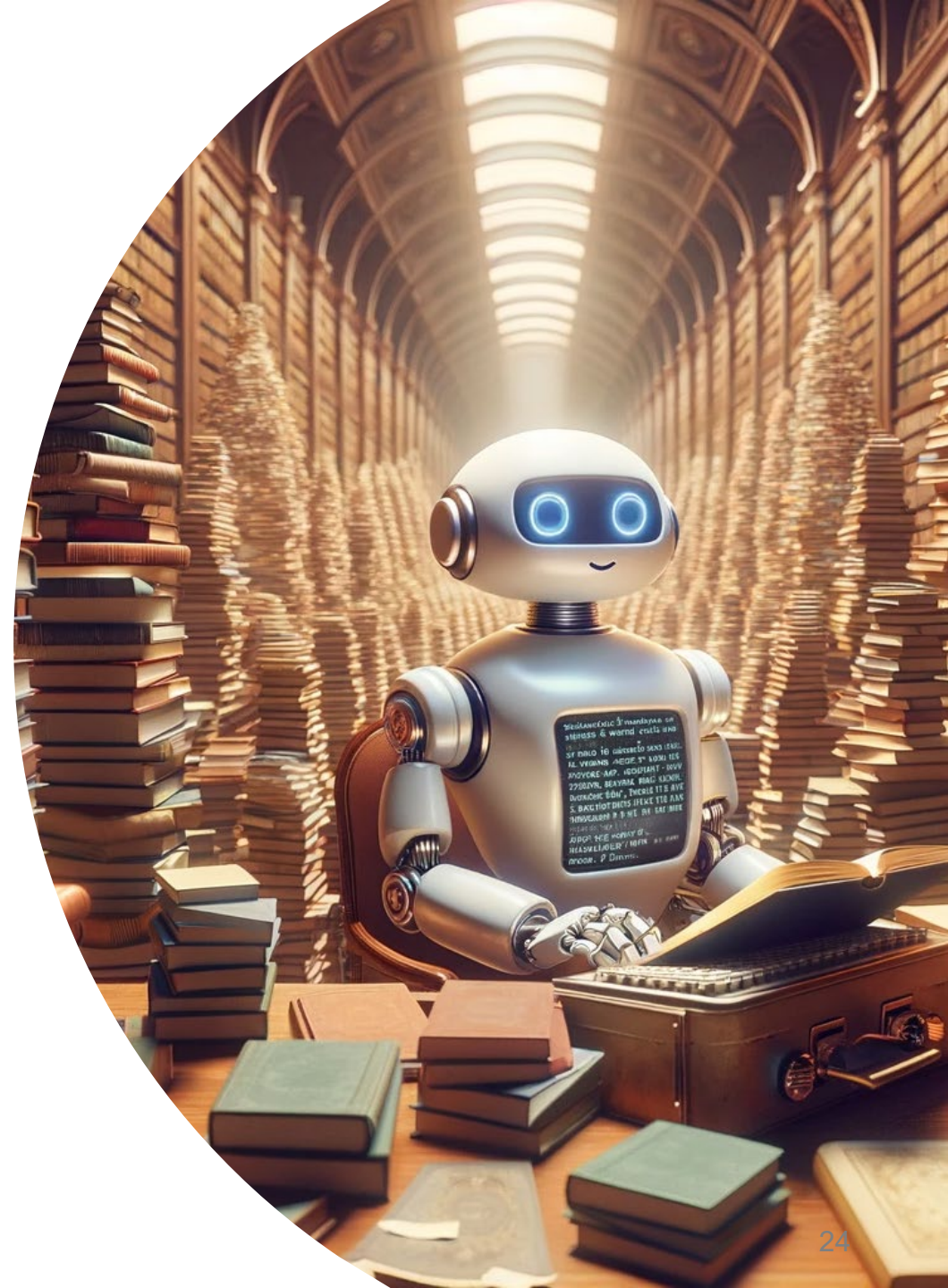
Wat is een API?

- Er zijn heel veel API's: betalend of gratis (of gelimiteerd)
- Meestal altijd een key aanvragen om toegang te krijgen
 - Beveiliging van de API
 - Toegangscontrole
 - Facturatie
 - Statistieken
- Werken met API van ChatGPT!
 - *Whoop whoop!*



Wat is ChatGPT?

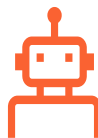
- Groot taalmodel: begrijpt en voorspelt tekst.
 - Autocompletion 'on speed'
- Getraind met 300 miljard woorden (\approx 3,75 miljoen boeken)
- Analytisch, niet alwetend
 - Geeft mogelijkheden, geen zekerheden!
- We kunnen een API van ChatGPT in Python-code gebruiken (zie <https://rapidapi.com>)



Tekst samenvatten

spagina) voor andere betekenissen van Albert Einstein. Albert Einstein (1879-1955) was een Duitse theoretische natuurkundige van Joodse afkomst. Hij won de Nobelprijs voor de Natuurkunde in 1921 voor zijn verklaring van het foto-elektrisch effect. Einstein werd in 1879 in Ulm geboren en groeide op in München. Hij trouwt in 1903 met Mileva Marić en samen kregen ze twee zoons. In 1919 scheidden Einstein en Marić. Datzelfde jaar trouwde Einstein met zijn nicht Elsa Löwenthal. Einstein verloor zijn Duitse staatsburgerschap door zijn weigering om zijn militaire dienstplicht uit te voeren. Hij werd lid van de Pruisische academie van Wetenschappen in Berlijn in 1914 en werd directeur van het Keizer Wilhelm Instituut voor Natuurkunde in Berlijn. Verder werk omvat onder meer onderwerpen in de kwantummechanica, de relativiteitstheorie, de kosmologie, de theorie van straling (fotonen, dualiteit, de Compton-effect, de theorie achter de laser) en de veldentheorie. Een chemisch element einsteinium is ook naar hem vernoemd, net als de Einsteinring. Albert Einstein werd in een liberaal-joodse familie in het Duitse Keizerrijk geboren. Hij emigreerde naar de VS in 1953. Hij overleed op 18 april 1955 in Princeton aan een hartinfarct. Hij werd begraven op de Princeton University Chapel. Hij werd in 1999 opgenomen in de 19e eeuw Jeugd Cijferlijst van Einsteins eindexamen middelbare school. Hij kreeg 6. waarbij 6 het hoogste cijfer. en 4 een voldoende is. Einstein had vijf kinderen: twee zoons en drie dochters.

Schrijf een zeer korte Nederlandstalige heldere samenvatting van de gegeven tekst met enkel de belangrijkste feiten



Albert Einstein (1879-1955) was een Duitse theoretische natuurkundige van Joodse afkomst. Hij werd bekend vanwege zijn bijdragen aan de speciale en algemene relativiteitstheorie. In 1921 won hij de Nobelprijs voor de Natuurkunde voor zijn verklaring van het foto-elektrisch effect. Einstein werd in 1879 in Ulm geboren en groeide op in München. Hij trouwt in 1903 met Mileva Marić en samen kregen ze twee zoons. In 1919 scheidden Einstein en Marić. Datzelfde jaar trouwde Einstein met zijn nicht Elsa Löwenthal. Einstein verloor zijn Duitse staatsburgerschap door zijn weigering om zijn militaire dienstplicht uit te voeren. Hij werd lid van de Pruisische academie van Wetenschappen in Berlijn in 1914 en werd directeur van het Keizer Wilhelm Instituut voor Natuurkunde in Berlijn.

Tekst samenvatten

```
1 import requests
url = "https://chatgpt-42.p.rapidapi.com/conversationgpt4"
payload = {
    "messages": [
        {
            "role": "user",
            "content": "Schrijf een zeer korte Nederlandstalige heldere
2                      samenvatting van de gegeven tekst met enkel
3                      de belangrijkste feiten. \n" + text
        },
        {
            "system_prompt": "",
            "temperature": 0.9,
            "top_k": 5,
            "top_p": 0.9,
            "max_tokens": 250,
            "web_access": False
4        }
    ],
    "system_prompt": "",
    "temperature": 0.9,
    "top_k": 5,
    "top_p": 0.9,
    "max_tokens": 250,
    "web_access": False
5 }
headers = {
    "content-type": "application/json",
    "X-RapidAPI-Key": APIKEY,
    "X-RapidAPI-Host": "chatgpt-42.p.rapidapi.com"
6 }

response = requests.post(url, json=payload, headers=headers)
samenvatting = response.json()["result"]
print(samenvatting)
```

1. Importeer alle packages
2. Adres van API (voor request)
3. Stel vraag aan ChatGPT
4. Parameters voor ChatGPT
5. Voeg geheime API-Key toe
6. Stuur request/vraag naar API, print antwoord uit

Parameters ChatGPT

```
system_prompt": ""  
temperature": 0.9,  
top_k": 5,  
top_p": 0.9,  
max_tokens": 250,  
web_access": False
```

- "system_prompt": "" - Begin van het verhaal.
- "temperature": 0.9 - Hoe creatief de antwoorden zijn.
- "top_k": 5 - Top 5 woorden kiezen.
- "top_p": 0.9 - Beste 90% woorden gebruiken.
- "max_tokens": 150 - Maximale lengte van antwoord.
- "web_access": False - Geen internettoegang.

Wat gaan we automatiseren?



~~Informatie downloaden~~



~~Mappen & bestanden doorzoeken~~



~~Informatie samenvatten (met API)~~



Automatisch e-mailen

Afbeelding e-mailen

- Vaak geautomatiseerd, waarom?
 - Tijd besparend & meer efficiëntie: e-mails nieuwe klanten, nieuwsbrieven...
 - Personalisatie: e-mails personaliseren op basis van interesses
 - Schaalbaarheid
 - Analyse & tracking



Gmail gebruiken

- E-mails uitsturen: vaak betalend
- Kan ook werken met Google e-mailserver = Gmail!
- Gmail-account instellen
 - Verificatie in twee stappen aanzetten
 - App-wachtwoord maken
- Gegevens gebruiken in Python-script en zo automatische mails versturen!

← App-wachtwoorden

Met app-wachtwoorden kun je inloggen op je Google-account vanuit apps op apparaten die geen ondersteuning bieden voor verificatie in 2 stappen. Je hoeft het app-wachtwoord maar één keer op te geven en hoeft het niet te onthouden. [Meer informatie](#)

Je app-wachtwoorden

Naam	Gemaakt	Laatst gebruikt	
Python automate	13:11	–	

Selecteer de app en het apparaat waarvoor je het app-wachtwoord wilt genereren.

App selecteren ▼ Apparaat selecteren ▼

GENEREREN



Afbeelding e-mailen

```
1 import smtplib
  from email.mime.multipart import MIMEMultipart
  from email.mime.text import MIMEText
  from email.mime.image import MIMEImage

  recipients = [""]

2
3 msg = MIMEMultipart()
  msg["Subject"] = title
  msg["To"] = ", ".join(recipients)
  msg["From"] = 'f"{GMAIL}@gmail.com"'

4 tekstdeel = MIMEText(samenvatting, 'plain', 'utf-8')
  msg.attach(tekstdeel)

  with open(afbeelding, 'rb') as f:
    afbeeldingsinhoud = f.read()

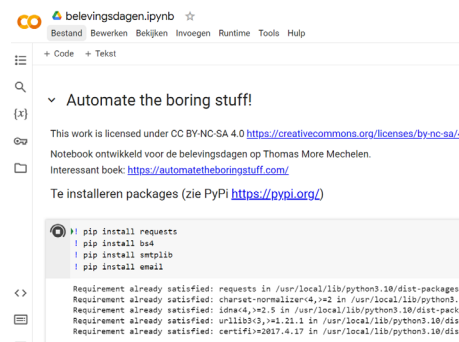
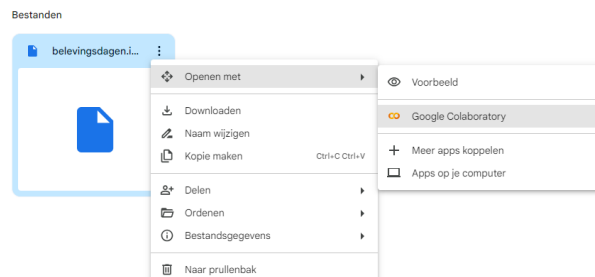
5 mime_subtype = 'jpeg' if afbeelding.lower().endswith('.jpg') or
  afbeelding.lower().endswith('.jpeg') else 'png'
  afbeelding = MIMEImage(afbeeldingsinhoud, _subtype=mime_subtype)
  msg.attach(afbeelding)

6 smtp_server = smtplib.SMTP_SSL('smtp.gmail.com', 465)
  smtp_server.login(GMAIL, GMAIL_APP_PASSWORD)
  smtp_server.sendmail(msg["From"], recipients, msg.as_string())
  smtp_server.quit()
```

1. Importeer alle packages
2. Ontvangers selecteren
3. E-mail gegevens instellen
4. Tekst toevoegen aan e-mail
5. Afbeelding toevoegen als attachment
6. E-mail versturen!
Let op Gmail-adres & app-paswoord

Kan jij dit allemaal zelf?

- Ja, je hebt enkel een Gmail-account nodig!
- Download mijn materiaal en ga aan de slag!
 - <https://github.com/ElkeBoonen/Belevingsdagen>
- Zet de demo-notebook in je Gmail-drive, open met Google Colaboratory
- Maak een 'images'-folder
- Vul de juiste waarden in én experimenteer!



Bedrijf, Organisatie & Toerisme

Mechelen - Antwerpen

Bachelors

Bedrijfsmanagement

- Accountancy – Fiscaliteit
- Financiën en Verzekeringen
- Internationaal ondernemen
- Marketing (Sport, Digital, Fashion)
- Supply Chain Management (Airfreight Logistics, Logistiek Management)
- Automotive Management
- Business en Management
- Immobiliën en Verzekeringen
- **Applied Data Intelligence (NL)**

Organisatie en Management

- Eventmanagement
- Human Resources
- Intercultural Relations Management (Diplomatie, Social Engagement)

Toerisme en recreatiemanagement (Aviation, Reisarchitect, Cultuur & Sporttoerisme)

International Business Management

- **Data Science, Protection and Security (ENG)**
- Global Supply Chain Management
- International Business and Trade

International Tourism and Leisure

Applied Data Intelligence (NL)

Data Science, Protection & Security (EN)



- **Data:** Beheer, transformeer en analyseer data én ontdek Machine Learning!
- **Technische vaardigheden:** Leer programmeren, begrijp AI-algoritmen en automatiseer taken.
- **Beveiliging:** Bescherm en beveilig data, hack AI-systemen voor een verbeterde beveiliging.
- **Bedrijfskennis:** Vertaal data naar zakelijke inzichten, ondersteun besluitvorming en communiceer effectief over je bevindingen.
- **Praktische ervaring:** Pas kennis toe in projecten en stage!

Met dank aan:

- <https://automatetheboringstuff.com/>
(beste Python boek!)
- <https://support.google.com/mail/answer/185833?hl=nl>
- <https://chat.openai.com/chat> ;)
- <https://rapidapi.com/rphrp1985/api/chatgpt-42>

