

GETIJDETHERMOMETER

BOUWEN MET WEBSCRAPING

Weten wanneer een strandwandeling erin zit of dat het strand juist onbegaanbaar is vanwege de vloed: hoe handig zou het zijn om een led-verlichte meter met de actuele getijdestanden te hebben. Die gaan we in deze workshop bouwen met behulp van webscraping, Python 3 én een 3d-printer.



Ook doen?

Op pactive.nl/workshops/6957-getijdethermometer-bouwen-met-webscraping vind je dit volledige artikel met werkende URL's (die hier zijn onderstreept), plus de download van Zeno's webscraping-programma in Python 3.

Een groot deel van Nederland ligt onder de zeespiegel. Duinen en dijken houden het water veilig buiten. Toch is de informatie over de waterstanden in Nederland belangrijk en wordt door de opwarming van de aarde steeds significanter. Waterstanden of getijdedata worden nu bijvoorbeeld veel opgevraagd door zee- en waddenhavens, zeilverenigingen, strandtenten, vissers, strandwandelaars en reddingsbrigades, maar in de toekomst zal iedereen erin geïnteresseerd zijn. Het idee van deze getijdethermometer ontstond tijdens een strandwandeling bij het Zeeuwse Westkapelle. Als ik op dat moment had geweten dat het op dat moment volledig vloed was en het strand dus bijna niet meer te belopen viel, was ik de wandeling op een ander tijdstip gestart. Er is dus behoefte aan *real-time* data.

Getijde-informatie is op veel websites beschikbaar, maar om iedere keer de browser op te starten en te zoeken naar de juiste website en de gewenste informatie is erg onpraktisch. Vandaar deze thermometer die de status van eb of vloed en de actuele hoogte van de waterstand ten opzichte van (N)ormaal (A)msterdams (P)eil optisch weergeeft door middel van een lichtbalk met zes meer-kleurenleds. De data verkrijgen we met behulp van *webscraping*.

WEBSCRAPING

Webscrapen is een computertechniek waarbij een microcontroller of microcomputer wordt gebruikt om informatie van webpagina's op het internet te lezen en eventueel te interpreteren zonder dat een gebruiker er invloed op heeft. In PC-Active

Weerstation Westkapelle ZLD			
getijden - Gezeiten - tides - marées			
Datum	Tijd in uur	Waterstand	t.o.v. NAP
Dinsdag 30-11-2021	10:06	hoog	143 cm
	16:30	laag	-143 cm
	22:41	hoog	179 cm
Woensdag 01-12-2021	05:00	laag	-119 cm
	11:05	hoog	171 cm
	17:25	laag	-157 cm
Donderdag 02-12-2021	23:33	hoog	206 cm
	05:50	laag	-136 cm
	11:51	hoog	200 cm
Vrijdag 03-12-2021	18:20	laag	-171 cm
	00:18	hoog	227 cm
	06:36	laag	-152 cm
Zaterdag 04-12-2021	12:36	hoog	224 cm
	19:05	laag	-180 cm
	01:02	hoog	239 cm
Zondag 05-12-2021	07:25	laag	-166 cm
	13:20	hoog	241 cm
	19:50	laag	-183 cm
Maandag 06-12-2021	01:46	hoog	243 cm
	08:10	laag	-175 cm
	14:05	hoog	249 cm
Dinsdag 07-12-2021	20:35	laag	-179 cm

Getijde-informatie op Westkapelle.info

316 vanaf pagina 74 staat een uitgebreid artikel over dit onderwerp. Het *scrapen* van websites die zijn geschreven in HTML-formaat, is relatief eenvoudig. Met een softwaretool zoals BeautifulSoup (crummy.com/software/BeautifulSoup/bs4/doc) vind je al snel de gewenste informatie op een website (zie ook het voorbeeld Webscraping met Python [op mijn eigen site](#)).

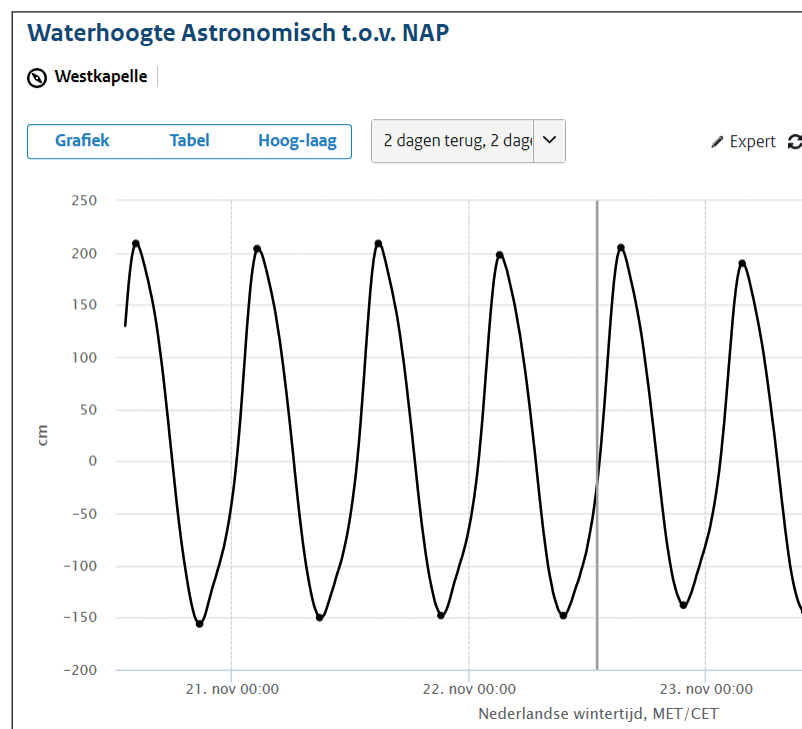
De site Westkappel.info toont in pure HTML het resultaat van verzamelde getijdedata van die plaats. Een scrape-programma kan eenvoudig die code en de inhoud ervan lezen en de getijdedata toepassen. Dit keer gaan we het echter moeilijk doen!

Tegenwoordig zijn de meeste websites veel dynamischer dan alleen HTML-code. Ze maken veelvuldig gebruik van bijvoorbeeld Java of andere scripttalen. Het wordt dan lastiger om met deze dynamica *real-time* te scrapen. Gelukkig zijn ook hiervoor speciale softwarebibliotheken ontwikkeld zoals [Selenium](#), die goed om weten te gaan met dynamische en real-time gegenereerde data. In ons voorbeeld gebruiken we deze bibliotheek.

DE JUISTE DATA

Getijdegegevens worden nauwkeurig bijgehouden op deze webpagina van Rijkswaterstaat: waterinfo.rws.nl/?#!/kaart/astronomische-getij/info. Ons uitgangspunt voor deze workshop is zoals gezegd de locatie Westkapelle op Walcheren, het meest westelijke punt van Nederland. In Westkapelle is de vloed ongeveer 2,2 meter boven NAP en is het eb bij ongeveer -1,7 meter beneden NAP. De data worden op de website in een grafiek (of tabel) als functie van de tijd (interval

10 minuten) weergegeven voor meerdere dagen. De gebruiker kan ook kiezen voor de getijdedata van andere plaatsten in Nederland. Hieronder een weergave van de getijdedata voor Westkapelle:

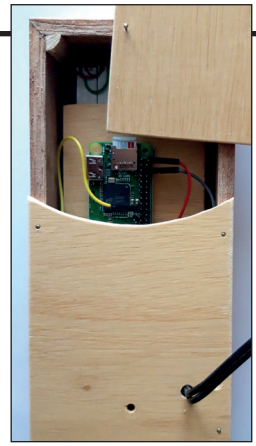


Samen met een microcomputer (microcontroller kan ook) die met internet is verbonden en de softwarebibliotheken BeautifulSoup en Selenium haalt de getijdethermometer zijn informatie op en presenteert deze.

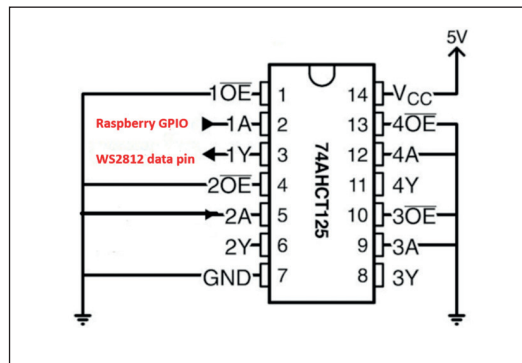
HARDWARE

De Raspberry Pi Zero microcomputer (circa € 10,-) ondersteunt wifi en heeft daardoor internettoegang. Een volledige (gratis) Python 3-omgeving is standaard aanwezig op de Linux-distributie Raspbian, die op een sd-kaartje draait. Om een scrape-programma te schrijven, kun je kiezen uit meerdere *editors*, zoals Geanny, Idle of Thonny. Ook het aansturen van de hardware zoals meer-kleurenleds ([WS2812 op Adafruit.com](#)) wordt eenvoudig, omdat op de Raspberry Pi veel GPIO-pinnen toegankelijk zijn voor het aansluiten van hardware en die makkelijk te benaderen zijn via Python.

De Raspberry Pi Zero werkt met een spanning van 3.3 V. Ook de GPIO-pinnen werken met deze ►

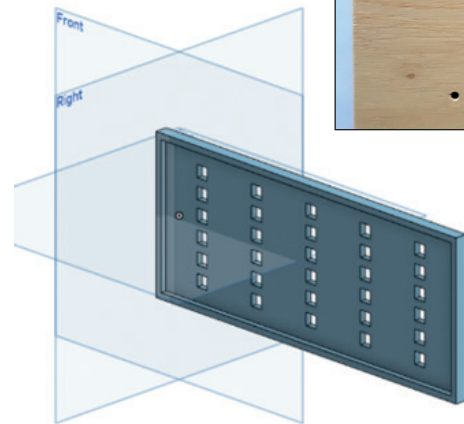


spanning. Interfacen van deze poorten betekent dat ook de aan te sluiten hardware met dit spanningsniveau overweg moet kunnen. Mijn programma GetijdePi gebruikt WS2812 meerkleuren-leds. Deze werken met een spanning van 5 V. De [website van Adafruit](#) doet meerdere suggesties met betrekking tot het aanpassen van de signaalniveaus tussen de GPIO-pinnen van de Raspberry Pi en de hardware. Wij kiezen voor de signaalaanpassing met een 74HCT125, zie de afbeelding hieronder.



Voor de thermometer heb ik een display gebouwd met de leds in een matrix van 5x6. Op deze wijze kan de thermometer nog in andere vormen worden gerealiseerd, bijvoorbeeld vijf leds breed in plaats van nu één led breed. Ook kunnen de overgebleven leds nog worden gebruikt om andere informatie weer te geven, zoals bijvoorbeeld springvloed.

De dertig leds heb ik in een 3d-geprint sjabloon geplaatst. Hieromheen heb ik een houten lijstje met houten achterkant gemaakt, zie de afbeeldingen rechts bovenaan deze pagina.



Met een rode respectievelijk groene led in de rechter bovenhoek van het display geeft de thermometer de status van de vloed en het eb weer. Een rode led in het midden van het display geeft aan dat de Raspberry Pi aan het opstarten is.

SOFTWARE

De software heb ik vanwege de snelheid en vanwege het ontwikkel- en *debug*-gemak eerst ontwikkeld op een Windows 11-pc, met Anaconda Spyder als *editor* en Python 3.8 als *interpreter*. Vervolgens heb ik het programma overgebracht naar de Raspberry Pi Zero en de hardware en aansturing van de GPIO geïmplementeerd. Het programma `getijdePi.pi` wordt ter illustratie hieronder deels in fragmenten weergegeven. Het volledige programma vind je als downloadbaar tekstbestand onderaan de workshop op pactive.nl

Allereerst worden met een 'import' enkele bibliotheken geladen, waaronder de webdriver

```

7
8 import board
9 import neopixel
10
11 from bs4 import BeautifulSoup
12 from selenium import webdriver
13 from datetime import timedelta
14 import datetime, time
15
16 pixel_pin = board.D10
17 num_pixels = 5 * 6
18 ORDER = neopixel.GRB
19 pixels = neopixel.NeoPixel(pixel_pin, num_pixels, brightness=0.1, auto_write=False, pixel_order=ORDER)
20
21
22 url = 'https://waterinfo.rws.nl/?#!//details/publiek/astronomische-getij/Westkapelle(WESTKPLE)/Waterhoogte_
23     _20berekend_20oppervlaktewater_20t.o.v._20Normaal_20Amsterdams_20Peil_20in_20cm'
24

```

(Selenium) en BeautifulSoup. Het installeren van Selenium en BeautifulSoup is even een klusje, maar de uitleg is te vinden op de [Getting Started](#)-pagina van Selenium en op deze [Medium-site](#). De bibliotheek 'board' en 'neopixel' zijn nodig voor het aansturen van de ws2812-ledstrip. Het stroomverbruik van de neopixels kan behoorlijk oplopen als er veel worden gebruikt en als de helderheid van de leds hoog wordt ingesteld. De 5V-voeding kan niet worden geleverd via de GPIO-pinnen van de Raspberry Pi. Een extra voeding is dus nodig (5V, 2A of meer), zeker wanneer de `brightness`-parameter (nu ingesteld op `brightness=0.1`) dichterbij 1 wordt gekozen. Doe dan ook een zonnebril op!

De volledige link van de website waterinfo.rws.nl die moet worden gescraped, wordt weergegeven in de `url`-parameter.

In de `Readtable()`-procedure wordt het daadwerkelijke scrapen uitgevoerd. Samen met de geïnstalleerde chromedriver-API voor Linux (onder Windows kan het ook) en de Selenium-webdriver worden de data opgehaald (`driver.get(url)`). Met behulp van BeautifulSoup worden deze data afgegeven aan de variabele lijsten `datum`, `tijd` en `hoogte`. Eenmaal opgehaald, worden de data ook opgeslagen op de sd-kaart met de procedure `saveData()`. Indien het niet mogelijk zou zijn om de data op te halen (als je bijvoorbeeld geen internetverbinding hebt), dan worden de eerder opgeslagen data gebruikt voor weergave op de thermometer. Na het scrapen zijn de gegevens van de website bekend in het programma en kan de thermometer worden opgebouwd.

In de procedure `Analiseer()` wordt via de `vandaag[3]`-procedure een index `[p]` berekend die staat voor het momentane tijdstip van de dag en die verwijst naar de bijbehorende waterstand (`hoogte[p]`) in de tabel. Deze waarde (in centimeters) gaat vervolgens naar de `thermometer()`-procedure. Bij het aanroepen van de procedure `thermometer()` wordt ook de maximale en minimale te verwachten waterhoogte van de locatie Westkapelle meegegeven.

De `main loop` van het programma (`while True`) kijkt elke 600 seconden in de tabel. Indien het de volgende dag wordt (`old_vandaag != vandaag[0]`), dan wordt de website opnieuw gescraped en worden

nieuwe data in de tabel geplaatst. Als laatste stap in het programma worden de pixels van de thermometer ververst met de procedure `pixels.show()`.

OPSTARTEN PROGRAMMA GETIJDE.PI.PY

Een Raspberry Pi met bijvoorbeeld Rasbian als besturingssysteem erop geïnstalleerd, maakt het eenvoudig om een programma automatisch te laten starten wanneer de stekker in het stopcontact gaat. Het bestandje `/etc/rc.local` moet hiervoor worden aangepast met een editor. Voeg de regel toe (voor exit 0) die aangeeft dat in de directory `/home/pi` het bestandje `startGetijdePi.sh` staat. In het bestandje `startGetijdePi.sh` staat vervolgens de regel `python3 /home/pi/getijdePi.py` die ervoor zorgt dat python opstart.

OPMERKINGEN

Webscraping is een interessante techniek die het mogelijk maakt om de (on?)eindige hoeveelheid aanwezige kennis en data op internet op een automatische manier op te halen, te verwerken en toe te passen zonder tussenkomst van een persoon. Dit biedt weer extra nieuwe mogelijkheden bij het gebruik van het Internet of Things in ons dagelijks leven.

Ons thermometerdisplay is uit te breiden met meer kleuren en meer leds. We gebruiken er hier tenslotte maar zes, terwijl er dertig kunnen worden aangestuurd. Die uitdaging is voor de nabouwer van dit prototype.

In dit display kun je uiteraard ook nog een balkje van zes leds gebruiken om bijvoorbeeld een temperatuurmeting of energiemeting te doen of andere data te presenteren.

Een goedkope Raspberry Pi Zero-microcomputer voldoet zoals gezegd prima voor de taak van webscraping. Het opstarten van `GetijdePi.py` duurt weliswaar ongeveer tien minuten, maar vervolgens werkt het programma moeiteloos. Ik raad aan om voor het ontwikkelen van de (python-) software een wat sneller systeem te gebruiken. Dat mag ook een Windows-systeem zijn. ■

GetijdePi in werking: er kan veilig gevist worden. De groene led geeft de eb-status aan, het water staat laag

