

Data-analyse die er toe doet



bert@hubertnet.nl
<https://berthub.eu/iv-it>

Kiesraad

Toetsingscommissie
Inzet Bevoegdheden



POWERDNS



FOX-IT
FOR A MORE SECURE SOCIETY

POWERDNS

POWERDNS

TU Delft

Opsporing komt in gevaar door disfunctioneren tapkamer politie

Nationale Politie De afdeling Interceptie & Sensing van de politie heeft grote technische en personele problemen. De verhoudingen zijn verziekt.

Dat het toch nog lang goed leek te gaan met het aftappen van criminelen, heeft alles te maken met een ander, nooit in de Tweede Kamer besproken tapsysteem, dat de Nederlandse politie sinds 2010 als aanvullend systeem gebruikte. Dat systeem is juist gericht op het analyseren van complex data-verkeer, en niet op het klassieke tappen van gesprekken. Het is een Nederlands systeem dat werd ontwikkeld door security-bedrijf Fox-IT. Dat systeem, **Replay**, werkte zo goed, dat het onder meer aan andere westerse inlichtingendiensten is verkocht.

Ik mag er zelf niks over zeggen, maar dit stond in de krant:

<https://www.nrc.nl/nieuws/2022/09/28/opsparing-komt-in-gevaar-door-disfunctioneren-tapkamer-politie-a4143537>

< > nature.com

SkewDB, a comprehensive database of GC and 10 other skews for over 30,000 chromosomes and plasmids | Scientific Data

scientific data

View all journals Search Login

Explore content About the journal Publish with us Sign up for alerts RSS feed

nature > scientific data > data descriptors > article

Data Descriptor | [Open Access](#) | [Published: 22 March 2022](#)

SkewDB, a comprehensive database of GC and 10 other skews for over 30,000 chromosomes and plasmids

[Bert Hubert](#)

Scientific Data 9, Article number: 92 (2022) | [Cite this article](#)

687 Accesses | 5 Altmetric | [Metrics](#)

Abstract

GC skew denotes the relative excess of G nucleotides over C nucleotides on the leading versus the lagging replication strand of eubacteria. While the effect is small, typically around 2.5%, it is robust and pervasive. GC skew and the analogous TA skew are a localized deviation from Chargaff's second parity rule, which states that G and C, and T and A occur with (mostly) equal frequency even within a strand. Different bacterial phyla show different kinds of skew, and differing relations between TA and GC skew. This article introduces an open access database (<https://skewdb.org>) of GC and 10 other skews for over 30,000 chromosomes and plasmids. Further details like codon bias, strand bias, strand lengths and taxonomic data are also included. The SkewDB can be used to generate or verify hypotheses. Since the origins of both the second parity rule and GC skew itself are not yet satisfactorily

Download PDF

Sections Figures References

- [Abstract](#)
- [Background & Summary](#)
- [Methods](#)
- [Data Records](#)
- [Technical Validation](#)
- [Usage Notes](#)
- [Code availability](#)
- [References](#)
- [Acknowledgements](#)
- [Author information](#)
- [Ethics declarations](#)
- [Additional information](#)
- [Rights and permissions](#)

<https://www.nature.com/articles/s41597-022-01179-8>

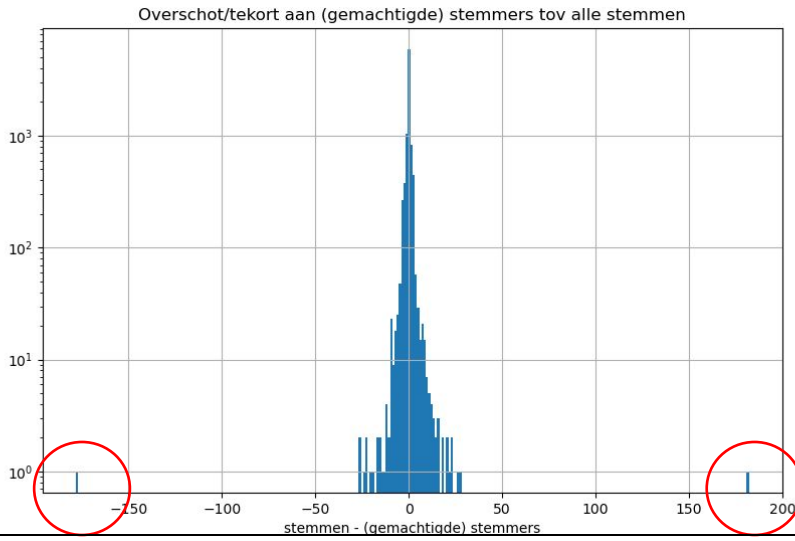


Toetsingscommissie Inzet Bevoegdheden

<https://www.tib-ivd.org/>

```
: plt.figure()  
plt.yscale("log")  
(sbmeta.blanco + sbmeta.ongeldig + sbmeta.stemmen-sbmeta.toegelaten).hist(bins=300)  
plt.title("Overschot/tekort aan (gemachtigde) stemmers tov alle stemmen")  
plt.xlabel("stemmen - (gemachtigde) stemmers")
```

Figure 8



Wordt in oktober open source!



<https://unsplash.com/@jayclark>



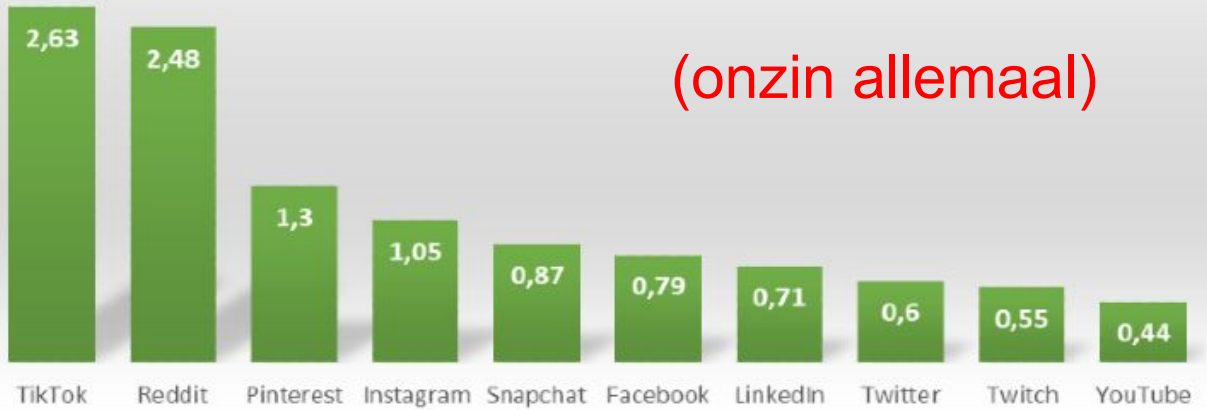
Copernicus heeft echt PETABYTES aan data om mee te werken.

De data maakt hier
echt uit. En dat is nog
een hele strijd.

De cola-aanbieding in een niet optimale week plannen is zo erg niet.

CO2 Emissions Per Social Media Platform (grams per minute)

■ CO2 Emissions Per Social Media Platform (grams per minute)



<https://8billiontrees.com/carbon-offsets-credits/carbon-footprint-of-the-internet/> -
lariekoek!

Minimale, gemiddelde en maximale water verbruiken per maand

Onderstaande tabellen geven het de minimale, gemiddelde en maximale water verbruiken weer.

● Minimale waterverbruik per ... ● Maximaal waterverbr... ● Gemiddeld water...



?????



Provincie Noord-Holland

https://www.noord-holland.nl/Onderwerpen/Water_Bodem/Publicaties/Onderzoek_wat_erverbruik_datacenters

NOS Nieuws • Vrijdag 19 maart 2021, 10:21 •
Aangepast vrijdag 19 maart 2021, 13:49



'Datacenters in Noord-Holland kunnen leiden tot drinkwatertekort'

Datacenters in Noord-Holland gebruiken zoveel water om de apparatuur te koelen, dat er mogelijk een tekort aan drinkwater voor huishoudens kan ontstaan. Dat blijkt volgens [De Telegraaf](#) uit interne stukken van de provincie Noord-Holland. Aan het koelwater worden daarnaast onbekende chemicaliën toegevoegd die in het milieu belanden. Volgens de krant hebben ambtenaren een kritisch onderzoeksrapport hierover bewust niet gepubliceerd. De provincie ontkent dit.

“De Telegraaf maakte uit interne stukken van de provincie Noord-Holland op dat de twee datacenters 525.000 liter drinkwater per uur gebruiken. Op jaarbasis zou dat 4,6 miljard liter drinkwater zijn”

“De krant ging er van uit dat de datacenters 24/7 en het hele jaar door met deze hoeveelheid drinkwater gekoeld moeten worden. Het water wordt echter alleen op de heetste dagen van het jaar gebruikt”

(NRC tuinde er ook twee keer in)

https://www.metropoolregioamsterdam.nl/wp-content/uploads/2023/06/Rapportage-objectivering-datacenters-definitief-juni-2023_.pdf

<https://www.nrc.nl/nieuws/2023/02/28/denk-na-voor-je-op-post-share-of-like-klikt-a4158268>

<https://www.nemokennislink.nl/publicaties/hoe-vervuilend-is-jouw-internetgedrag/>

**Wat IS die data dan
eigenlijk?**

Data-analyse is zoiets
als verven. Boel
plamuren, schuren en
afplakken. En het echte
schilderen **is pas het
staartje.**

Als voorbeeld - je wil weten hoeveel woningen er ergens gebouwd zijn in 2022. Maar die data is er niet. Er is wel het aantal verleende bouw- en sloopvergunningen, en hoeveel kavels er gemaakt zijn en hoeveel panden er gesplitst en samengevoegd zijn. Voor je dat allemaal begrepen en rechtgelegd hebt ben je wel even verder.

Subject: Review reminder for Gene Reports
From: [redacted] <em@editorialmanager.com>
Date: 03-09-2023 07:01
To: "Bert Hubert" <bert@hubertnet.nl>

This is an automated message.

Manuscript Number: GENREP-D-23-00627
Oligonucleotide skews enable comprehensive and insightful characterization of GC- and TA-skew properties observed throughout the human genome with support of unsupervised AI with reference to gene- and Alu-polarity skews

Dear mr Hubert,

You kindly agreed to review the above referenced manuscript on Aug 21, 2023 and your completed review was due by Sep 02, 2023.

Your review is now 1 days late. As you know, timely reviews are of utmost importance to authors, therefore I would be grateful if you would submit your review as soon as possible.

To view the manuscript, please click here: [View Submission](#).



https://en.wikipedia.org/wiki/The_Treachery_of_Images

From em.genrep.15aa7.85c8a4.785e07fb@editorialmanager.com Sun Sep 3 07:01:51 2023
Return-Path: <em.genrep.15aa7.85c8a4.785e07fb@editorialmanager.com>
X-Original-To: bert@hubertnet.nl
Delivered-To: bert@hubertnet.nl
Received: from nssmtp01.editorialmanager.com (nssmtp01.editorialmanager.com [216.205.0.108])
by D964BEAE.static.ziggozakelijk.nl (Postfix) with ESMTPS id 4BB67A80792
for <bert@hubertnet.nl>; Sun, 3 Sep 2023 07:01:51 +0200 (CEST)
Received: from NSWEB19 ([10.200.66.38]) by nssmtp01.editorialmanager.com with Microsoft SMTPSVC(10.0.17763.1);
Sun, 3 Sep 2023 01:01:49 -0400
MIME-Version: 1.0
Sender: em.genrep.15aa7.85c8a4.785e07fb@editorialmanager.com
From: [REDACTED]@em@editorialmanager.com
To: Bert Hubert <bert@hubertnet.nl>
Reply-To: [REDACTED]@cv@gmail.com
Date: 3 Sep 2023 01:01:49 -0400
Subject: Review reminder for Gene Reports
Content-Type: text/html; charset=utf-8
Content-Transfer-Encoding: base64
Message-ID: <NSSMTP01FapNuT1FM1J00113113@nssmtp01.editorialmanager.com>
X-OriginalArrivalTime: 03 Sep 2023 05:01:49.0624 (UTC) FILETIME=[BF251380:01D9DE23]

07:01:50 smtpd[2669294]: connect from nssmtp01.editorialmanager.com[216.205.0.108]
07:01:51 smtpd[2669294]: 4BB67A80792: client=nssmtp01.editorialmanager.com[216.205.0.108]
07:01:51 cleanup[2669302]: 4BB67A80792: message-id=<NSSMTP01FapNuT1FM1J00113113@nssmtp01.editorialmanager.com>
07:01:51 qmgr[1749089]: 4BB67A80792: from=<em.genrep.15aa7.85c8a4.785e07fb@editorialmanager.com>, size=4877, nrcpt=1
07:01:51 local[2669303]: 4BB67A80792: to=<ahu@server1.berthub.eu>, orig_to=<bert@hubertnet.nl>,
relay=local, delay=0.2, delays=0.18/0.01/0/0.01, dsn=2.0.0, status=sent (delivered to command: /usr/bin/maildrop)
07:01:51 qmgr[1749089]: 4BB67A80792: removed
07:01:51 smtpd[2669294]: disconnect from nssmtp01.editorialmanager.com[216.205.0.108] ehlo=2 starttls=1 mail=1
rcpt=1 bdat=1 quit=1 commands=7

En hoe goed snappen
we de data?

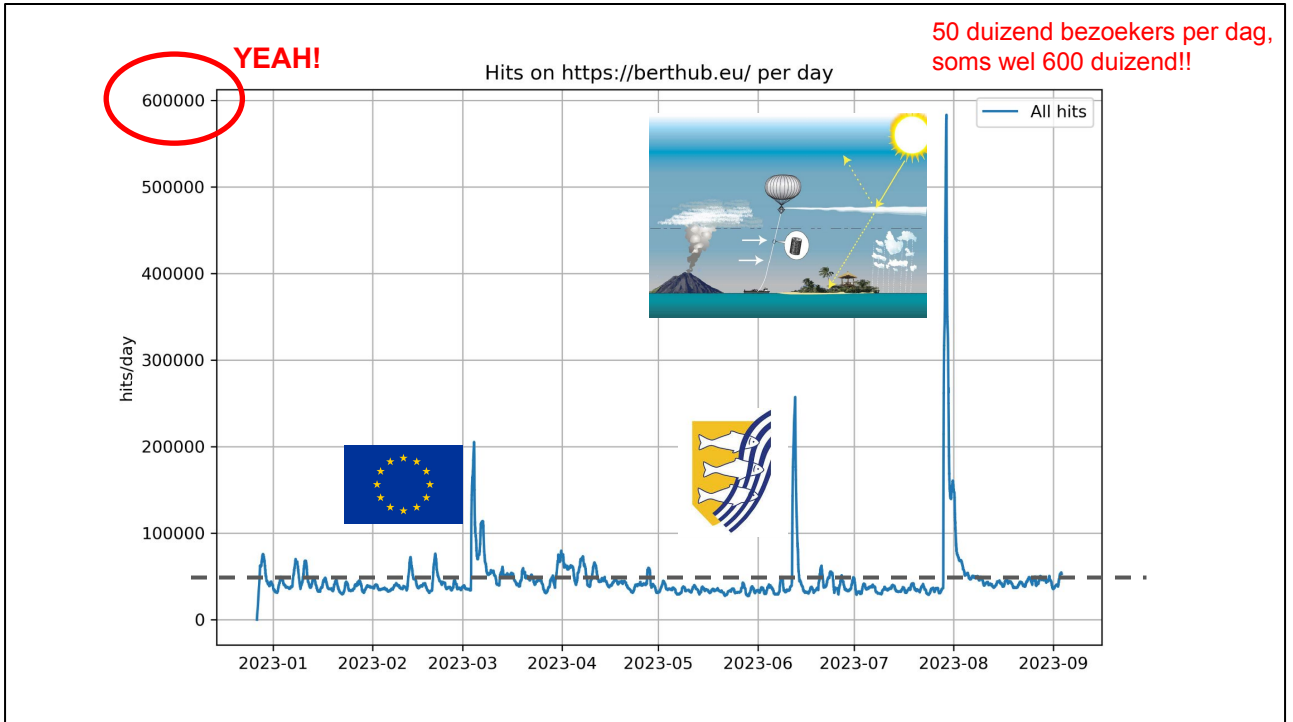
**Ontdek de toekomst
op 4 september bij
onze Tech Talks**

“The first principle is that you must not fool yourself—and you are the easiest person to fool.”

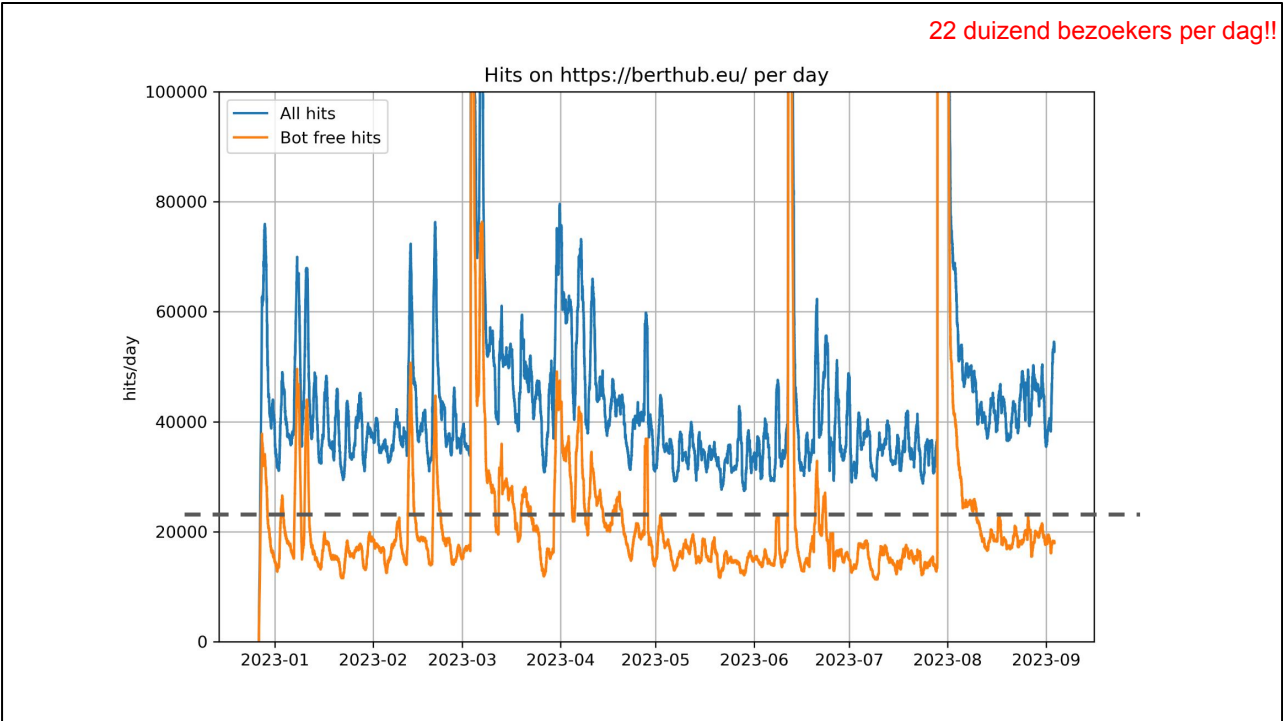
Richard Feynman “Surely You’re Joking, Mr. Feynman!” (1985)

<https://sites.cs.ucsb.edu/~ravenben/cargocult.html> dit moet je echt lezen

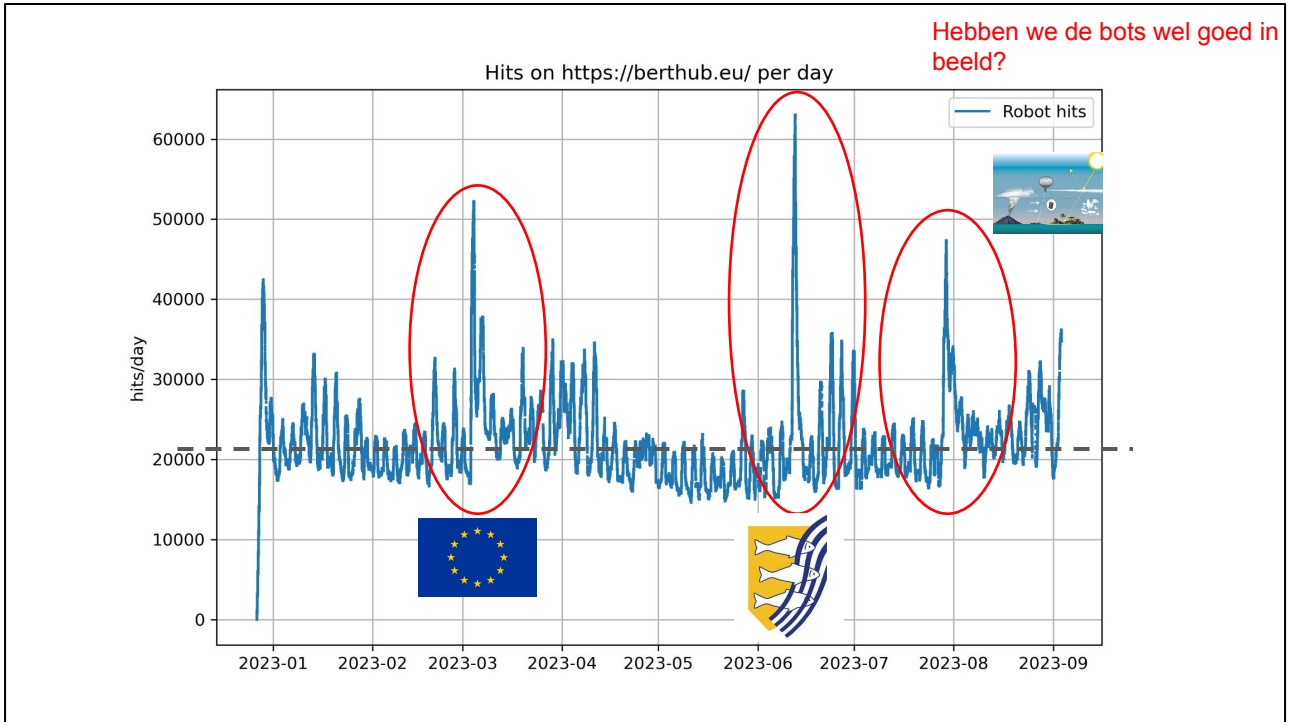
De website heeft
superveel bezoekers
na onze nieuwe
layout!



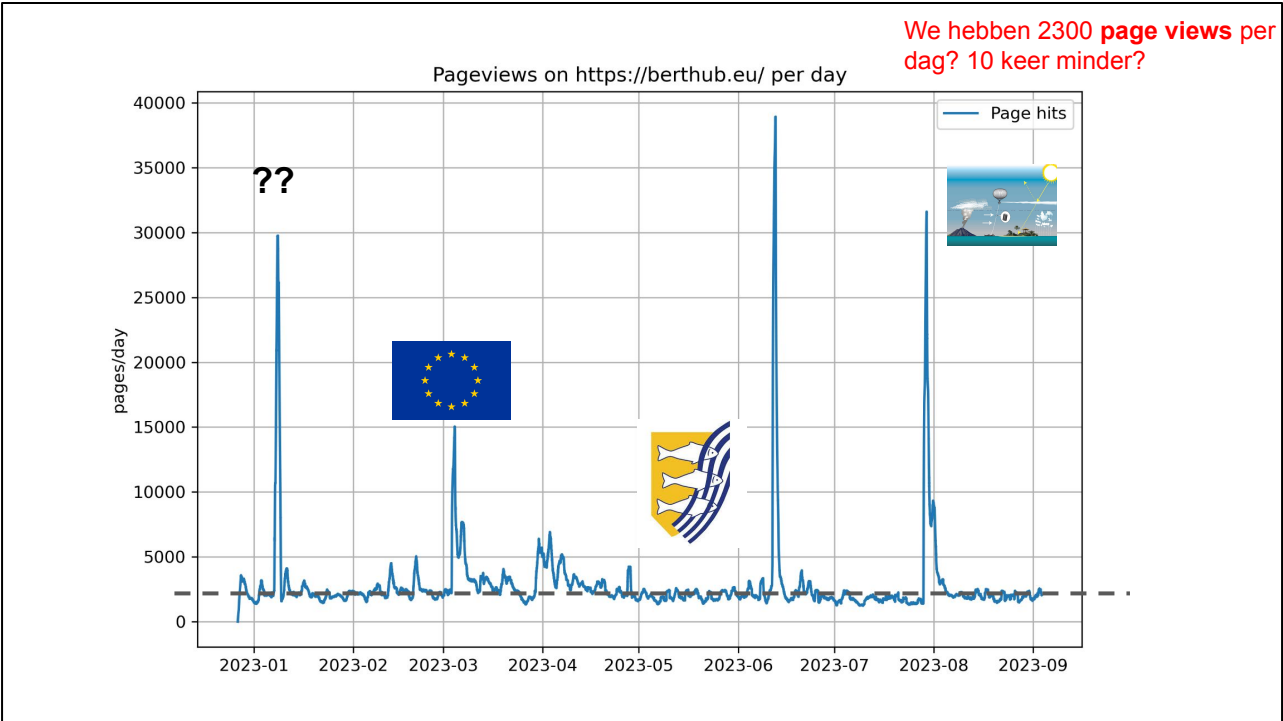
U voelt de bui natuurlijk al hangen - hits zijn geen bezoekers. Maar als je graag wil denken dat je site goed bezocht is is het best aantrekkelijk die vergissing te maken!



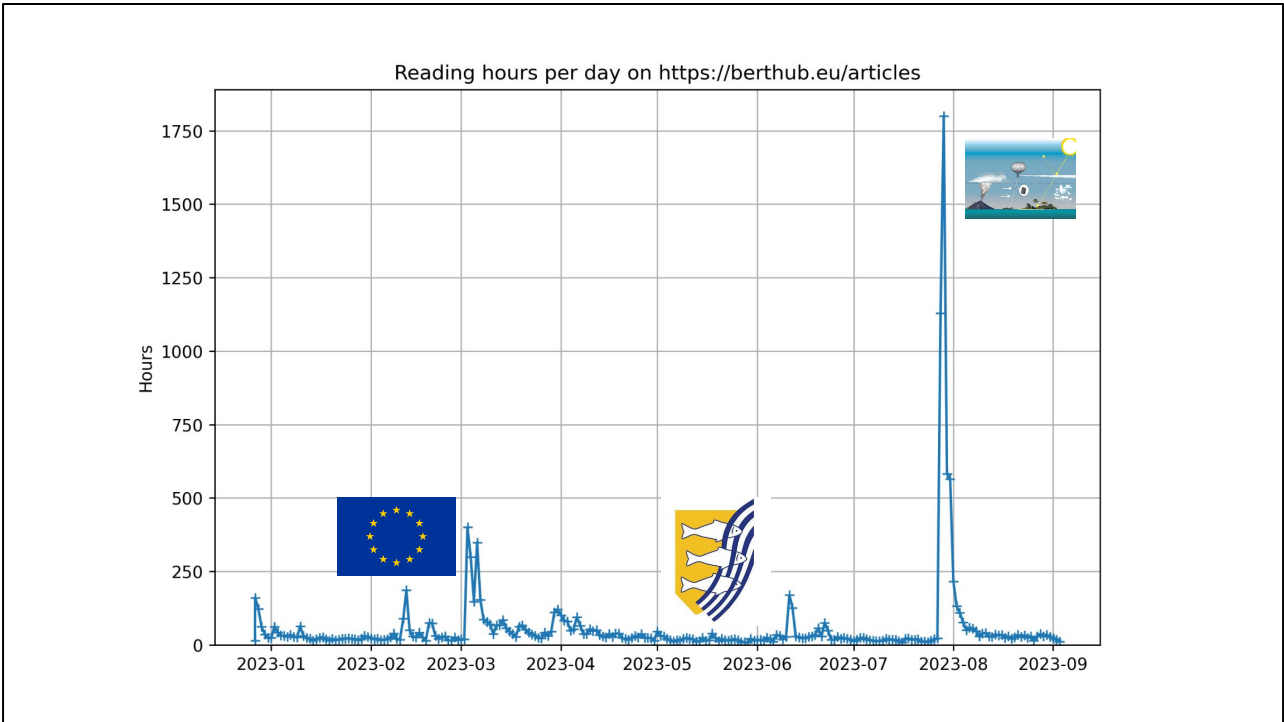
Ho zeg, de helft van ons verkeer zijn bots!



Dit is het verkeer alleen van bots - een beetje sus is dat de aandacht van bots ook toeneemt als onze site drukker bezocht wordt. Zou je uit moeten zoeken. Misschien klopt je botfilter niet.

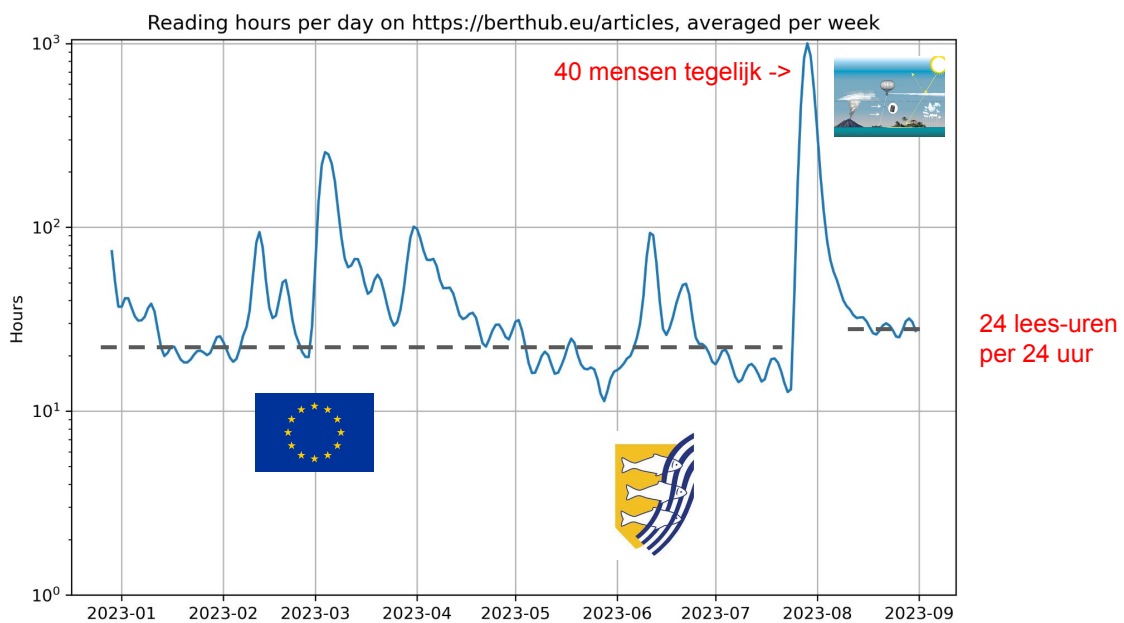


Nu meten we echte HTML page visits, en dat zijn er gelijk een heel stuk minder. En we hebben iets raars begin januari, 30000 bezoekers aan html pagina's. Blijkt iemand te zijn die de hele site heeft gemirrored.



En nu in anatal lees-uren per dag. Soms heel veel, meestal niet.

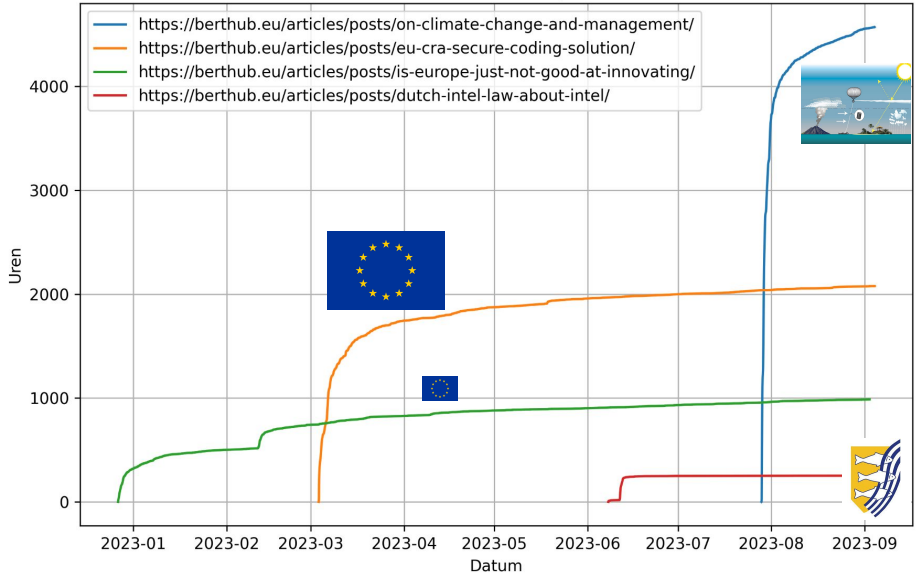
ER ZIT GEMIDDELD 1 IEMAND
ONZE SITE TE LEZEN!



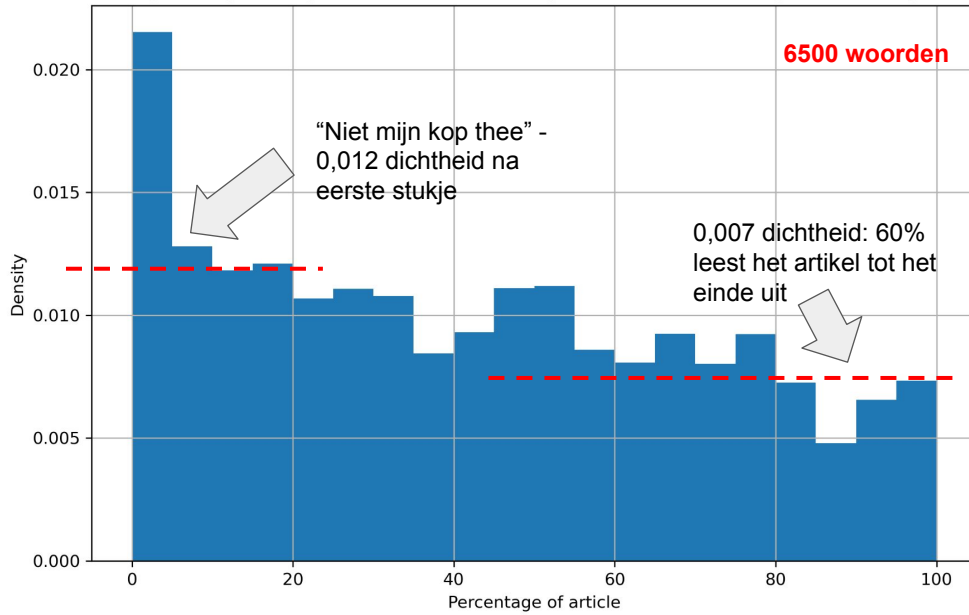
Dat valt tegen. 1 bezoeker.

Heeft onze
voorlichting wel zin
dan?

Totaal aantal lees-uren per artikel



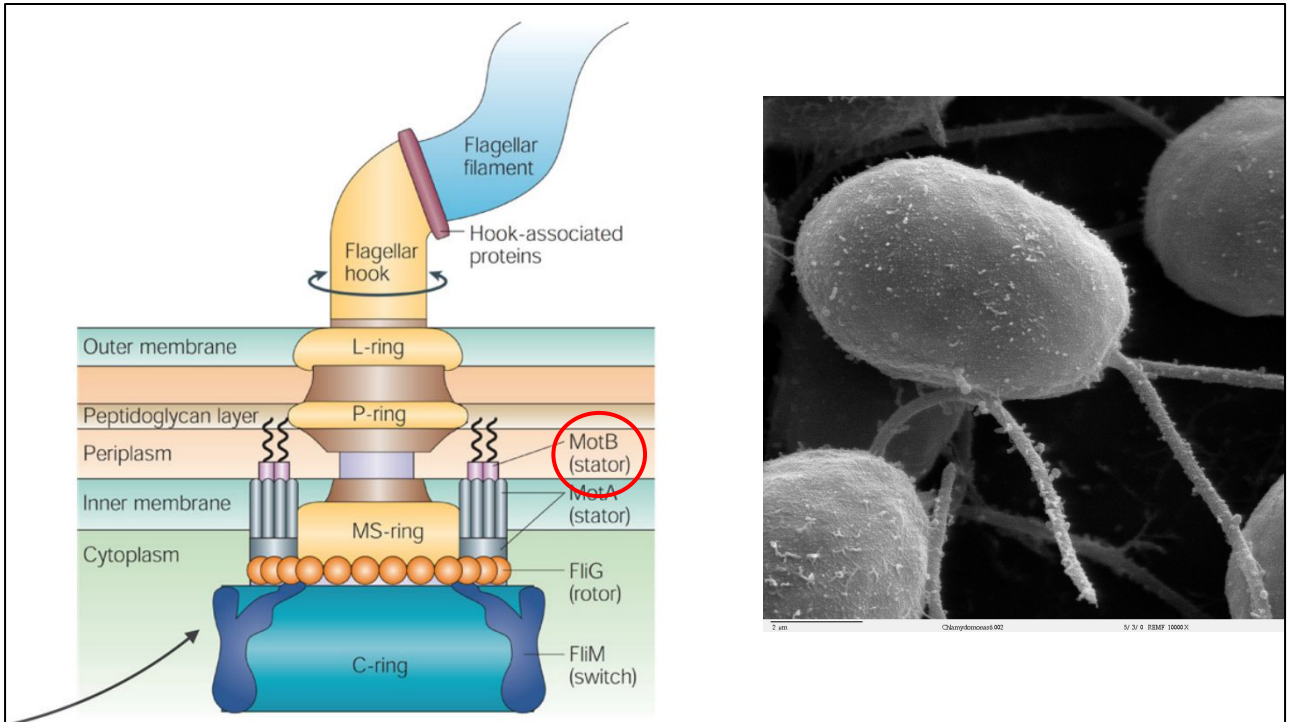
Reading density for <https://berthub.eu/articles/posts/on-climate-change-and-management/>



Dit komt uit <https://berthub.eu/articles/posts/tracking-free-audience-statistics/>

Data-paradox: vaak een **overschot aan data** en dan een dashboard met **precies niet wat je moet weten.**

Heb je ooit een indicatie gekregen van hoeveel lees-uren een artikel heeft opgeleverd? Of hoe ver het gelezen wordt?



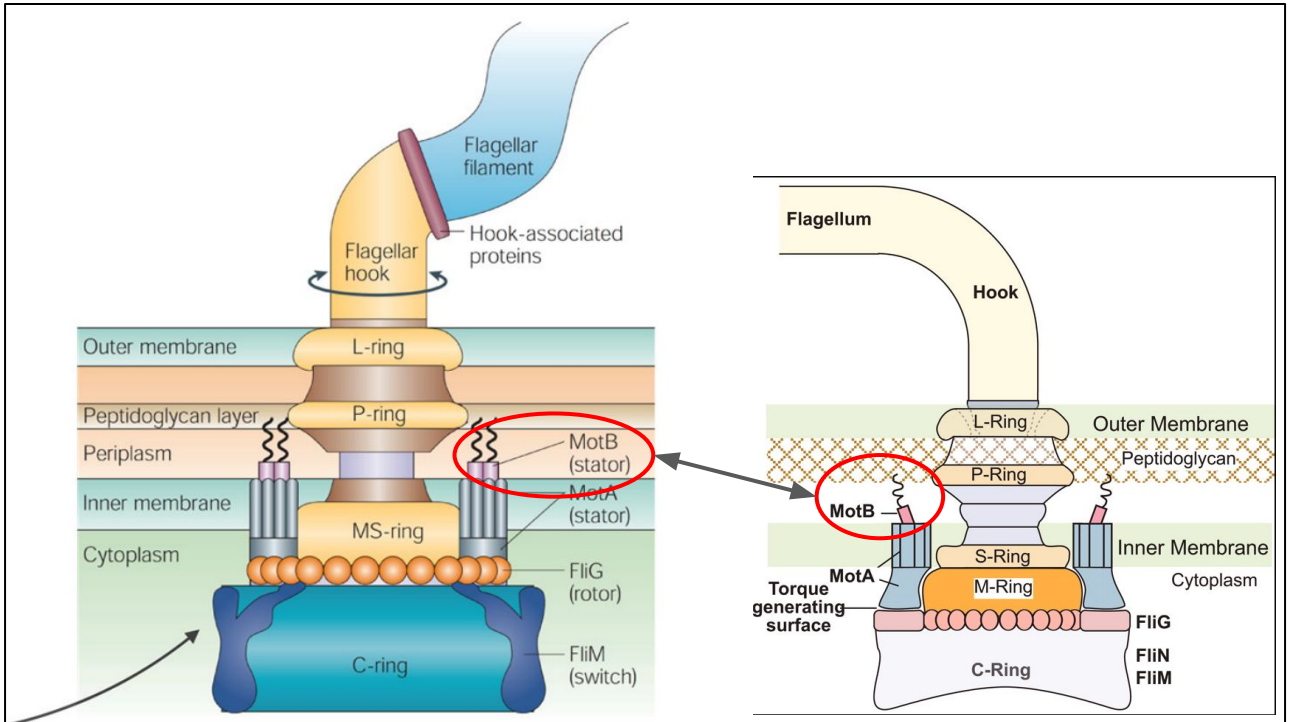
[https://en.wikipedia.org/wiki/Flagellum#/media/File:Chlamydomonas_\(10000x\).jpg](https://en.wikipedia.org/wiki/Flagellum#/media/File:Chlamydomonas_(10000x).jpg)

MotB DNA source code: iedere 3 letters -> 19 atomen

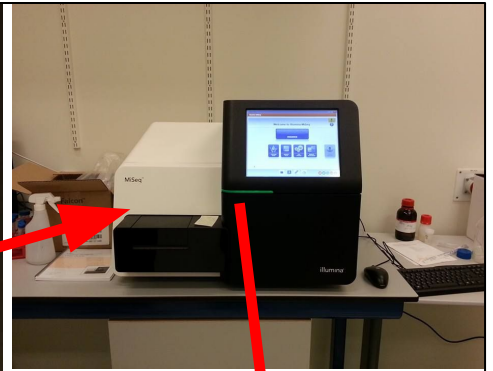
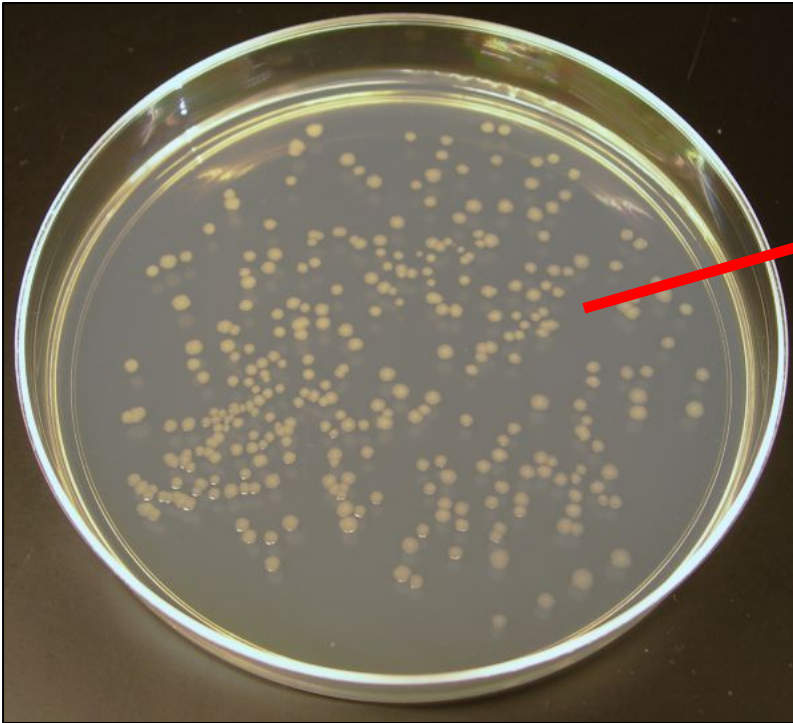
```
1 tcacctcggt tcggetgatg gcattgtggg aacctgacc tgtggtgcaa cctcaggttt
61 ttccagggcg cttactgget cattetgget ttcggcggtt tcatgcaaaa tggcctgttc
121 ggettgtttg ttcagtacca gcaggetgat gcgacggtt acggcatcat caggtcgcg
181 atcgettaag cgcacgttg cgcctatgcc gacgacagt aacctttgc cgetatccaa
241 ccctccgacc atcagttcgc ggcgggatgc attggcccga tcggcagaaa gctcccagtt
301 getatatcet ttctaccgc tggcgtaggg gaaatcatcg gtatgacctg aaaggetaat
361 acggttggga ataccgttca gtacaggcgc aatggcgcgc agaattgtcgc gcataaggg
421 ttcgacatcg gcactgccgg ttetaaacat cgggcgattc tggetatcga tgatctgaat
481 acgtagacct tcctggacca gatcgatttt gagatgggga cgtaacgccc gcagtttcgg
541 atcggactct atcaactggt cgagatcacc ccgcaatttc cgcaatcgac tttgetccat
601 gcgttttttc agctettcga tgttcggetg cttattcaet tccccctggc tttgggtgta
661 atcatcacca ccgccgggaa ttgggettcc actattagaa atgcgatcgc cgcccgtaac
721 gcagtcgcc agtggagtcc ggaagtactc cgcaatctga atcagctett ttgggetgga
781 gatggagatc agccacatca ccagaaaaaa ggccatcatc gcagtcataa agtcggcata
841 agcaatcttc cacgatccat gtgctgcccc gtggettittg gctttgcgtc gtttgacgac
901 aataatcgga tgcgettgat tcctcat
```

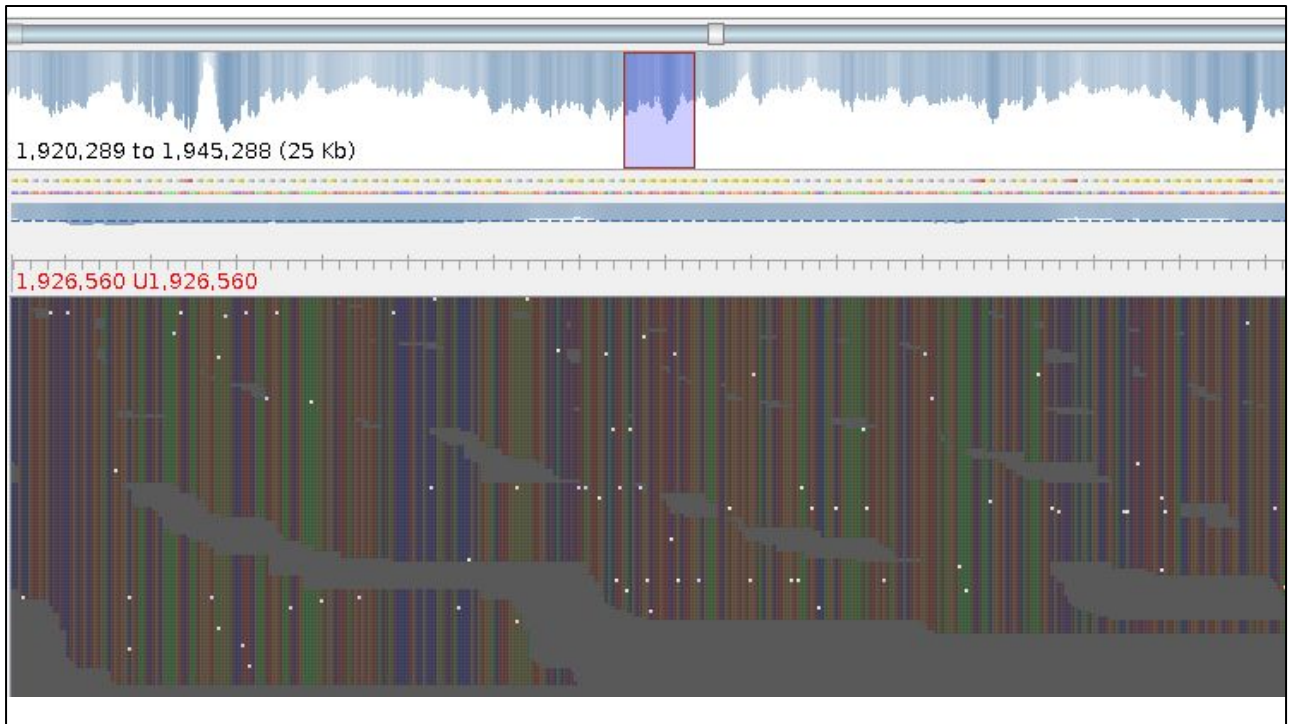
<https://berthub.eu/amazing-dna/>

<https://berthub.eu/amazing-dna/> - https://en.wikipedia.org/wiki/Motility_protein_B



[https://en.wikipedia.org/wiki/Flagellum#/media/File:Chlamydomonas_\(10000x\).jpg](https://en.wikipedia.org/wiki/Flagellum#/media/File:Chlamydomonas_(10000x).jpg) - stop de MotB van de ene bacterie in de andere, en kijk hoe dat evolueert





De witte stipjes zijn mutaties

En toen.. Niks.. alleen ruis. Ik kon er niks van maken. **Te weinig data.**

[nature](#) > [the isme journal](#) > [original article](#) > [article](#)

Published: 03 July 2015

Density-dependent adaptive resistance allows swimming bacteria to colonize an antibiotic gradient

[Felix J H Hol](#) , [Bert Hubert](#), [Cees Dekker](#) & [Juan E Keymer](#) 

The ISME Journal **10**, 30–38 (2016) | [Cite this article](#)

4708 Accesses | 31 Citations | 14 Altmetric | [Metrics](#)

Abstract

During antibiotic treatment, antibiotic concentration gradients develop. Little is known regarding the effects of antibiotic gradients on populations of nonresistant bacteria. Using a microfluidic device, we show that high-density motile *Escherichia coli* populations composed of nonresistant bacteria can, unexpectedly, colonize environments where a lethal concentration of the antibiotic kanamycin is present. Colonizing bacteria establish an adaptively resistant population, which remains viable for over 24 h while exposed to the antibiotic. Quantitative analysis of multiple colonization events shows that collectively

<https://www.nature.com/articles/ismej2015107> - aan de andere kant van de gang hadden ze wel geld.

Wat is die data dan?

Wat vertelt de data me echt?

Leer ik het juiste?

Of misschien helemaal niets?

Data-analyse die er toe doet



bert@hubertnet.nl
<https://berthub.eu/iv-it>