

**Bram van Es, PhD**

Amsterdam The Netherlands

☎ +31614385474

✉ contact@b-lab.nl

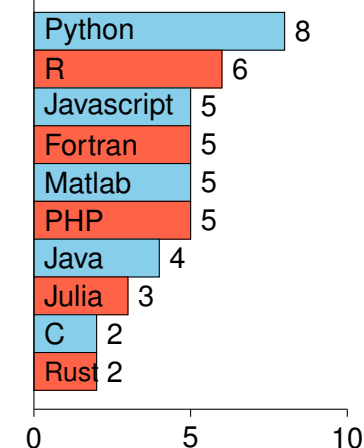
🌐 homepage

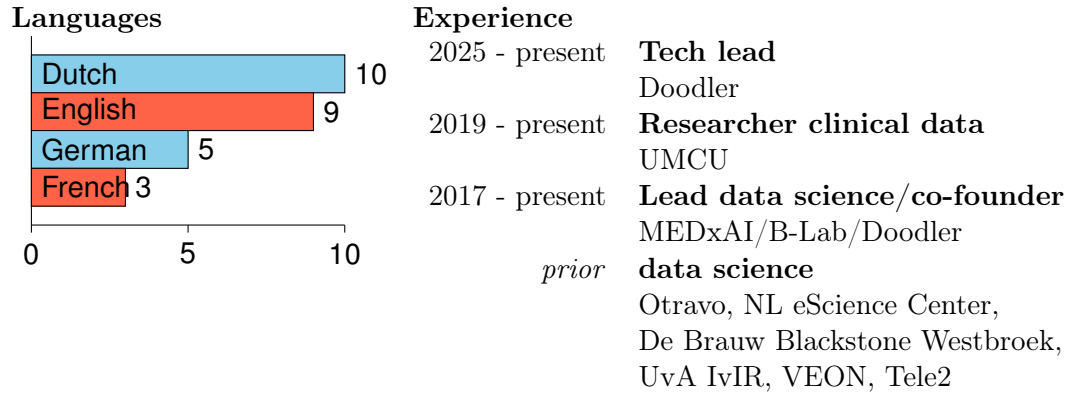
in linkedin

🎓 scholar

📺, 📺, k, 📺 ← *online presence***In short**

Experienced data scientist and machine learner with a background in aerodynamics, applied computational science, data science applied to healthcare, information law and online entrepreneuring. Idea creator with good social skills, can act as a connector and can work independently. Enjoys challenges: (1) in the application and development of novel numerical methods and machine learning algorithms, both in industry and in academics. (2) in the application of artificial intelligence and machine learning to legal practice, medical diagnostics, financial services, predictive maintenance and quality assurance.

Education**PhD applied mathematics****MSc aerospace engineering***pre-master information law***Programming**



Experience, Primary	<u>Senior Data Scientist</u> <i>University Medical Center Utrecht.</i>	1 December 2019 - present
--------------------------------	---	---------------------------

As part of the datalab, I work on novel and practical implementations of data science techniques to assist healthcare professionals in their work. For this purpose I am the lead developer of Argus, an algorithm to predict the presence of cardio-vascular disease. I also collaborate with researchers at the UMCG and the EMC to find bio-markers for lung cancer treatment using multi-omic analysis. I collaborate with IBD & RA researchers to help the early identification of patient sub-groups. I collaborate with NLP/AI researchers to develop/apply clinical language models and clinical NLP models for entity-classification, text embeddings and summarisation. I collaborate with nephrology researchers to develop time-series clustering methods for heterogeneous timeseries. I am personally working on foundational ECG models based on transformer to facilitate the development of downstream ECG-applications.

I am also partially responsible for assuring the quality of data science within UPOD.

keywords: Python, Spark, R, Pytorch, Google TPU computing, NLP, EHR analysis, multi-omic analysis, computer vision

<u>Owner</u> <i>B-Lab.</i>	June 2019 - present
-------------------------------	---------------------

A boutique AI shop: helping companies with specialised, bespoke machine learning solutions, from concept to deployment. As part of B-lab I am a co-founder of MEDxAI, a medtech startup. I am collaborating with legal experts to develop AI-driven legal assistants. I am a co-developer of Doodler, a service that turns patient-doctor conversations into easy-to-digest pictorials, here I am responsible for backend-development (Streamlit→REACT) and secure deployment in Azure. I am an advisor for a MedTech startup where I am focused on the development of ECG information extraction algorithms.

I made a pipeline to create Irish folksongs from regular news using LLMs and

ViTS (LLMs, ViTS). I created an online game to demonstrate the use of LLMs for the Dutch Design Week (REACT, PostgreSQL, Google cloud, LLMs). I also advised a digital artist (Sean Hannan) in developing an exhibition on deep fakes.

keywords: Python, PyTorch, REACT, NEXT.js, TypeScript, LLMs, ViTS, Azure, Google cloud

Tech-lead
Doodler.

January 2026- present

I am responsible for the development and cloud deployment of a multi-modal generative AI solution in the healthcare domain.

keywords: Python, Vite, Vercel, TypeScript, LLMs, Azure, Google cloud, Redis, PostgreSQL

Lead Data Scientist/co-founder
MEDxAI.

1 June 2017 - 1 August 2026

As a spin-out from a hackathon I am part of a small group of people who investigate the use of machine learning to: predict the efficacy of chemotherapy, predict the rejection of donor organs, predict cardiac pathologies, identify cancer(types), predict cancer metastasis and identify biomarkers. I work together with researchers and participate in hackathons (non-profit). We have collaborated with University Medical Center Utrecht, Erasmus Medical Center, the University of Groningen, Maastad hospital in Rotterdam and the Sint Maartenskliniek in Nijmegen. In this capacity I have functioned as the lead in multiple week-long hackathons.

For our first product, we are developing a clinical decision support system to assist cardiologists in diagnosing chest pain patients in second and third line hospitals. We do this in collaboration with the University Medical Center Utrecht and the Expertisecentrum zorgalgoritmen. As part of this project I develop a (1) clinical dashboard in NEXT.js, build pipelines for extracting labels from EHRs and develop pipelines for clinical prediction models. As part of this project we develop a foundational ECG model and Dutch clinical NER models, (2) information extraction system for clinical specialists using LLMs.

keywords: Keras, PyTorch, Python, Spark, Sklearn, MLlib, MMLSpark, interpretable AI

Senior Data Scientist
Otravo.

1 January 2018 - 1 December 2019

I was part of a small but expanding data team that aim to make the online ticket/tour pricing and the tour/travel recommendations data driven. Projects are for instance to develop a predictor for the offer availability, automate dy-

dynamic pricing, implement a recommendation engine or to analyse the factors influencing conversion and price quality.

keywords: Python, Spark, Solr, PostgreSQL, BigQuery, SCALA, timeseries classification, recommendation engine, google cloud

Data Scientist a.i.

May 2018 - December 2018

De Brauw Blackstone Westbroek.

De Brauw Blackstone Westbroek is one of the largest Dutch law firms, and they are exploring the use of data science to improve their service. I helped DBBW set up the infrastructure/codebase/workflow for a more data-driven legal practice through the use of decision matrices.

eScience research engineer
Netherlands eScience center (NLeSc).

1 August 2017 - 1 January 2018

In collaboration with a team of engineers I applied state-of-the-art computational techniques and IT-solutions to help further scientific projects. I was involved in a climate modelling project, aimed at accelerating/improving large scale high-performance computing simulations and I contributed to McFly, an open source code for hyper optimisation of deep neural networks for time series.

keywords: Keras, Fortran, Python, timeseries classification, distributed computing

Data Scientist

19 September 2016 - July 2017

University of Amsterdam, Institute for Information law (IvIR).

IvIR receives several TB's of raw data regarding the online behavior of internet users. This data has to be cleaned, stored and processed using text-mining algorithms. The aggregated results should be presented in a dashboard. The resulting database will be made accessible to research groups all over the world. I was involved in setting up the pipeline, from beginning to end using ElasticSearch, Redis and Hadoop for the backend married together by Python. Also, I assist in developing the plugins for Firefox and Chrome and I co-develop the libraries necessary to extract/analyse the data.

keywords: Realtime data cleaning, storage, filtering and processing, analysis dashboarding, online text-analysis, privacy-by-design, click stream analysis, ElasticSearch, Flume, Docker, Hadoop, Redis, JavaScript.

Data Scientist

01 May 2016 - 01 September 2016

VimpelCom HQ, DMP team.

I was part of a small but expanding team in Amsterdam that is rolling out a data management platform based on open source technologies, across the fourteen operating countries. My focus is now on market segmentation using graph clustering techniques (Markov Clustering, Affinity Propagation) and just-in-time triggers using anomaly detection.

keywords: Graph analysis, market segmentation, PySpark, Zeppelin, Spark, Python, Java, AWS EC2/S3

Quality data analyst

01 November 2014 - April 2016

Tele2, Network Engineering & Operations.

As a quality data analyst at Tele2 I collected and interpreted quality data from multiple sources augmented with external sources. The goal was to measure, analyse and understand the network performance, relate that to customer experience and give recommendations for improvements. My main activities were

- ingesting and analysing quality data using aggregation, regression and machine learning (using Python, R, Matlab),
- automating and visualising quality data analyses (using R, Tableau, Sisense),
- setting up a KPI dashboard, from back-end to front-end (using D3.js, PHP, Sisense, MySQL, Java, R),
- to help design/set up a data architecture and analysis workflow (Java, Spark, Hadoop).

The techniques I used were partially custom-made and novel, especially for causality analysis, time series clustering and HMM parameterisation. Otherwise I used algorithms such as Kmeans, Kmedoids, DBSCAN, OPTICS, SOM, PSD, FFT, Granger's test, Baum-Welch, Decision Trees and Logistic Regression.

keywords: dashboard development, root cause analysis, Java, Python, R, SQL, MongoDB

PhD researcher

2010 - 2014

CWI (Scientific Computing group) / FOM DIFFER.

The goal was the development of new and enhancement of existing numerical methods to approximate the extremely anisotropic diffusion present in nuclear fusion plasma simulations. In short, research entailed:

- development of new numerical schemes or concepts for enhancement of existing schemes; I focussed on lower order finite-volume and finite-difference schemes,
- implementation of algorithms; for rapid code development Matlab was chosen,
- devising and running test cases; test cases were created with fusion plasma simulation in mind, with extremely anisotropic diffusion tensors and with the direction of diffusion fully aligned with magnetic field lines,
- iteration on concepts; concepts were often heuristic in nature,
- interpretation and presentation of results, internally and externally.

Graduate trainee
ECN (Wind Energy)

01 September 2008 - 30 September 2009

As part of the rotorflow project at ECN Wind Energy a new unsteady coupled solver was being developed to compute the unsteady outer flow and the unsteady boundary-layer flow around wind-turbine rotors. The coupled solver has three distinct parts, the earlier mentioned outer flow and boundary layer flow and the coupling between these two parts. It was my task to investigate the integral boundary layer method as a means to solve the boundary layer flow. Both the models and the numerical implementation of finite-difference, finite-volume and discontinuous Galerkin schemes have been studied. Coding was done in Fortran and Matlab.

Graduate trainee
TNO TASS.

01 January - 28 April 2008

At TNO Automotive Safety Solutions I worked on generating and running test cases for the gas flow module of the Madymo solver. The gas flow module uses a compressible multi-species flow model to simulate the gas behaviour inside a deploying airbag. For the purpose of testing I looked at basic gas dynamics cases for which analytic solutions were available, most notably the moving piston cases. In parallel I documented the requirements to implement an existing Riemann solver (Osher scheme). After documentation I implemented the Riemann solver in the gas flow module using the Fortran programming language.

Graduate trainee
Dutch Aerospace Laboratory (NLR)

16-30 July 2007

Short installment at the NLR-facility in Amsterdam. It was my task to investigate the use (and added value) of microsoft Sharepoint for the Flight Test Systems and Applications group.

Owner/developer
www.idntools.com

2005- 2012

During my studies and part of my PhD research I created and maintained a website (www.idntools.com) with paying customers. The site relied on automatically mining/parsing and interpreting online data using self-written code in PHP/MySQL. For some time during my PhD research I was updating/maintaining the algorithms for the new site owner. The information would be retrieved daily through a cronjob of PHP-scripts pulling in compressed .txt files with domain names. PHP-scripts would extract/sanitise the data and extract names of the form 'xn-ABC123.EXT' with the accompanying expiration date. The domain name of the form 'xn-ABC123.EXT' in this extracted list would then be transformed to unicode. Using the unicode representation, e.g. google search/adwords API's, Bing API and Yahoo BOSS would be used using JSON and XML to extract indicators regarding the popularity of the keywords in the domain name and the cost-per-click of those keywords. The resulting data set was placed in a MySQL db and presented to the registered users in an interactive (asynchronous) list that could be ordered by date of expiration, language and by the values of the indicators.

Technical employee
SES Newskies

May 2006- March 2007

SES Newskies was preparing for the launch of the NSS-8 communication satellite. One of the preparations involved testing the telemetry commands issued by/to the satellite. It was my task to apply the pseudo-code, which was delivered by Boeing, to general test cases and to compare the results with the actual implementation in the satellite simulator. To this end I wrote several scripts that rewrote the pseudo-code into working Java code which could be applied to automatically generated test cases. The initiative to write these scripts was mine, it was originally gmail that I would manually transform the code into VB script and generate the test cases manually through an excel interface.

Other experience

traineeship Royal Dutch Shell Amsterdam: Literature study on pulsed eddy currents

traineeship KLM Engineering & Maintenance: Two month internship, assisting in overhaul and repair, reporting activities, presentation of work.

Assistant-manager of movie-theatre: operational management and assembling the films.

Miscellaneous: Cutting and packing flowers, packing flower bulbs, paper delivery, postal services.

Education

<u>PhD in applied mathematics</u> Centrum Wiskunde & Informatica, The Netherlands Thesis title: Numerical Methods for Anisotropic Diffusion Supervisors: Prof. dr. ir. B. Koren and dr. H.J. de Blank	2010 - 2014
<u>Master of Science, Aerospace Engineering - aerodynamics & CFD</u> Delft University of Technology Note: followed extra courses (about 25ECTS). Thesis title: Comparison and application of unsteady integral boundary layer equations. Supervisors: dr. H. Özdemir and dr. ir. B. van Oudheusden.	2006 - 2009
<u>pre-master, Information Law</u> University of Amsterdam	2007 - 2009
<u>Bachelor of Science, Aerospace Engineering</u> Delft University of Technology	2002 - 2006
<u>Bachelor of Aeronautical Engineering</u> Hogeschool van Haarlem	2000 - 2002

Computer experience

coding: C, CSS, **Fortran**, HTML, **Java**, JavaScript (jQuery, D3.js), Julia, Maple, **Matlab**, MPI, openMP, PHP, **Python**, **R**, SCALA, STOL, Spark, Bash, Rust, TypeScript, LaTeX

database: JDBC, MySQL, Oracle SQL, MS SQL, PostgreSQL, MongoDB, MonetDB, ElasticSearch, Solr, BigQuery, Redis, SQLAlchemy

software: Catia, Derive, Eclipse, Fluent/Gambit, KNIME Labview, Madymo, Photoshop, Pro Engineer, PyCharm, RapidMiner, Rstudio, Sisense, Solidworks, Tableau, Tecplot, VScode, Xfoil, Zeppelin, Jupyter, Jira, Confluence, Gephi, Kibana, Grafana

operating system: Linux, Unix, Windows, OSx

editors: Context, Inkscape, Kile, Kate, Notepad++, OpenOffice, Sublime, VIM, Pycharm, VSCode, ZED

Extra-curricular

- I like to tinker with code, e.g. I have made my own Telegram AI-chatbot that I use regularly and I am working on my own search engine using LLM APIs and knowledge-bases.
- I participated in a summer course on European and US privacy law and policy in 2017. .
- I initiated the set up of an online question and answer ticket system for CWI. The system was basically an online forum like stackoverflow.com but with the added feature that people can add tags for which they will receive alerts. The purpose was to connect the expert knowledge which is available within the institute and to provide a low-threshold means to ask questions. The ticket system was based on the open source software Askbot which is written in Python/Django.
- I gave tutoring lessons in physics and mathematics.
- Legal and technological volunteer for Bits of Freedom, an independent organisation that raises awareness with regard to privacy, the freedom of speech and net neutrality, I wrote part of infoleaflet about the internet, meant for the European Parliament.
- To keep myself busy with machine learning I set up a Kaggle team with old colleagues and study friends (CWI, Brown university, IBM Watson, Brain Corp, Sandia labs, etc.) and was involved in CWI's Kaggle team.
- I have worked on several git repos; RexR (machine learning for predicting the best chemotherapy for acute leukemia patients), Robout (for extracting data from ElasticSearch/Hadoop), Plugin/backend code in Python/Javascript) and several hobby repos; PubScience for corpora processing, TimEx for timeseries feature extraction.
- I am an amateur publicist, writing short and long essays on politics, economics, science and society.