

ELVIS NDAH, PhD

Email: elvis.ndah@gmail.com
Tel +32 486 963 187
Code repository <https://gitlab.com/ndah-e>
LinkedIn <https://www.linkedin.com/in/elvisndah/>
Nationality Belgian

WORK EXPERIENCES

- 11/2019 – Present **Data Scientist, Anju Life science Software Antwerp, Belgium**
- Developed custom ETL scripts to prepare analysis-ready data files from open source and proprietary sales data within quality-controlled standards.
 - Work independently, or as part of a team, to design and develop dashboards and analytics solutions to provide insights into medical device sales for our clients.
 - Use the python pymc3 package to develop simulation models for clinical trials patient recruitment to provide insights into study design and patient recruitment process.
 - Research in collaboration with the software engineering team on the development and implementation of a deep learning-based semantic search model in our clinical business intelligence platform.
- 01/2018–09/2019 **Data scientist, Bisnode Brussels, Belgium.**
- Design and develop dashboards to visualize customer segments based on descriptive statistics and unsupervised machine learning algorithms on socio-demographic data of the Belgian population.
 - Developed interpretable supervised machine learning models based on deep-dive statistical analysis and predictive data modelling that were used to identify the most valuable prospects for target advertisement and marketing campaigns.
 - Design and developed a marketing hotspot identification platform that assigned a propensity score for every neighbourhood in Belgium based on a look-alike machine learning model. The platform was deployed in a tableau dashboard.
 - Work closely with the development team to develop and implement credit scoring machine learning models for self-employed and SMEs into the company BI platform.
 - Perform drift analysis using Jason Shannon divergence to inform our clients on the evolution of their customers.
 - Work in collaboration with the development team to train and implement (system exposed as a REST API on AWS) a coupon recommendation.
 - Advise, train, and assist management, sales team and other professionals on working with and understanding the organization's data and BI platforms.

- 05/2013–10/2017 **PhD researcher in bioinformatics, *University of Gent, Belgium.***
- Design custom ETL pipelines for multi-omics (RIBO-seq, RNA-seq and MS proteomics) data processing such as data cleaning and filtering, transformation and integration for use in novel protein region identification and re-annotation of genomes, visualization in a genome browser and/or developing prediction models.
 - Conceive and develop a random forest model to predict protein-coding regions in bacterial genomes using sequencing data from ribosome profiling experiments.
 - Develop statistical models to identify differentially expressed genes from large scale ribosome profiling and RNA-seq sequencing data.
 - Developed a zero inflated statistical model to identify the most likely transcript undergoing translation in an active gene using ribosome profiling experimental data.
 - Communicate research results through internal and international conference presentations, scientific publications, or project reports.
- 07/2012–05/2013 **Research assistant in Bioinformatics, *DERI-University of Galway, Ireland***
- Developed statistical models to identify differentially expressed microRNA biomarkers to identify patients likely to respond to chemotherapy.
 - Investigate biological and functional patterns existing in genes associated with cancer cells using semantic web technologies.
 - Design and develop RDF databases of cancer genes and their biological and functional properties.
 - Provide training in introductory R/Bioconductor to colleagues.
- 02/2007–08/2009 **Webmaster, *Mawum Holdings SA Bamenda, Cameroon***
- Frontend web development using HTML, CSS and Dreamweaver.
 - Graphic design and photo editing with Adobe Photoshop and Illustrator
 - Responsible for all client and inhouse website management and maintenance.
 - Requirement analysis and website design and development.

SKILLS & ABILITIES

Programming	Python, Perl, R, SQL and SAS
Big data and ETL	Spark(PySpark) and Amadea iSoft
Data visualization	Power BI, Tableau, Dash, Seaborn, Matplotlib.
IDE	Jupyter notebook, Visual Studio Code and Pycharm
Cloud/Versioning	Git, AWS and Docker
DL frameworks	Keras and Pytorch
Web framework	Django and Bootstrap (HTML, CSS)

EDUCATION AND CERTIFICATION

01/2019	U2U Training Belgium, <i>Analysing your data with Power BI for BI Professionals.</i>
08/2017	SAS Education, <i>SAS Certified Base Programmer.</i>
10/2017	The University of Ghent, <i>Doctorate in Bioinformatics.</i> Thesis topic: <i>Complementary use of ribosome profiling and proteomics for mapping the translation landscape.</i>
10/2012	Katholieke Universiteit Leuven, <i>Master of Science in Artificial intelligence.</i> Thesis topic: <i>Algorithms for the reconstruction of genealogical trees of ancient texts.</i>
10/2011	Universiteit Hasselt, <i>Master of Science in Statistics.</i> Thesis topic: <i>Large-scale prediction of phenotypic variables using gene expression data.</i>
10/2006	The University of Buea, <i>Bachelor of Science in Mathematics and computer sciences.</i>

LANGUAGES

English	Native proficiency
French	Intermediary.
Dutch	Basic.

ACTIVITIES & HOBBIES

60/2015 – present	President, NGEMBA cultural and development association (NGEMCUDA). The organisation aims at promoting the NGEMBA culture in Belgium, fund developmental projects in Cameroon, and help migrants integrate into the Belgian society.
08/2016 – present	Vice-President, Starlight ex-student organisation (STESA) Europe. STESA is an association of Alumni from Starlight College Nkwen in Cameroon, the main objective is to raise scholarships funds for outstanding students enrolled in the school.
Hobbies	Watching and playing (physical and digital) football.

PUBLICATIONS

- MCP 2019 *Verbruggen & Ndah, et al.* PROTEOFORMER 2.0: further developments in the ribosome profiling-assisted proteogenomic hunt for new proteoforms. DOI:[10.1074/mcp.RA118.001218](https://doi.org/10.1074/mcp.RA118.001218)
- NAR. 2017 *Ndah, et al.* REPARATION: Ribosome Profiling Assisted (Re-)Annotation of Bacterial Genomes. DOI:[10.1093/nar/gkx758](https://doi.org/10.1093/nar/gkx758).
- BMC. 2017 *Giess, Ndah, et al.* A Unique Ribosome Signature Reveals Bacterial Translation Initiation Sites. doi:[10.1186/s12915-017-0416-0](https://doi.org/10.1186/s12915-017-0416-0).
- NAR. 2017 *Fijałkowska, Verbruggen, Ndah, et al.* eIF1 modulates the recognition of suboptimal translation initiation sites and steers gene expression via uORFs. DOI: [10.1093/nar/gkx469](https://doi.org/10.1093/nar/gkx469).
- MCP 2017 *Willems, Ndah et al.* N-terminal proteomics assisted profiling of the unexplored translation initiation landscape in Arabidopsis thaliana. DOI:[10.1074/mcp.M116.066662](https://doi.org/10.1074/mcp.M116.066662).
- MSB. 2016 *Gawron, Ndah, et al.* Positional proteomics reveals differences in N-terminal proteoform stability. DOI: [10.15252/msb.20156662](https://doi.org/10.15252/msb.20156662).
- NAR. 2015 *Ndah, Crappe, et al.* PROTEOFORMER: deep proteome coverage through ribosome profiling and MS integration. DOI: [10.1093/nar/gku1283](https://doi.org/10.1093/nar/gku1283).

CONFERENCES

- HUPO 2017, Dublin. **Poster:** "Identifying translated isoforms using ribosome profiling data".
- BePac2016, Ghent. **Oral presentation:** "Identifying translated isoforms using ribosome profiling: The case of the zero-inflated model".
- ECCB 2016, Hague. **Poster:** "REPARATION: Ribosome Profiling Assisted (Re-)Annotation of Bacterial genomes".
- GCC2015, Norwich. **Poster:** "Deep Proteome Coverage Through Ribosome Profiling and MS Integration".
- BBC2013, Brussels. **Oral presentation:** "Protein identification based on ribosome targeted mRNA fragments".