Joshua N.**Grant**
Data Engineer

☻ Contact

12918 Sanderling Lane
Knoxville, TN 37922
United States of America
+1 (865) 803 3495 ☐

jngrant@live.com ✉
notjustadatum.blogspot.com **B**
github.com/sempervent ☺
linkedin.com/in/joshuanagrant **in**
scholar.google.com ✐

◎ Objective

To take theories, ideas, research, and experiments and turn them into production-scalable solutions and to develop and support Python-based applications and data workflows on cloud infrastructure

---

▦ Experience

Senior Software Engineer Data Engineering
Bold Penguin, Columbus, OH (remote) 2022— Present

- Designed a more centralized system for development and deployment using GitHub Actions, AWS, and Python modules
- Supervised a team effort to migrate deployments, flows, and tasks from Prefect 1 to Prefect 2
- Developed Prefect tasks to run on either EC2 or Fargate instances based on conditions such as file size
- Created EC2 and Fargate microservices for assistance in document converstion and data extraction
- Designed and maintained REST APIs to be up-to-date with stakeholder and client demands
- Designed and implemented a scalable message queue (MQ) system to handle high load for a data extraction service, resulting in improved performance and increased efficiency of data processing operations
- Configured GitHub Actions and AWS CodePipeline to automate the build, test, and deployment process of the projects, ensuring consistent and reliable delivery of software updates
- Created and oversaw the deployment of multiple Python and JavaScript projects using AWS CodePipeline, resulting in streamlined, automated, and efficient project delivery

Data Engineer
Oak Ridge National Laboratory, Oak Ridge, TN *2019—2022*

- Served as support staff for a Python-based web scraping project, implementing an ETL workflow from a brittle-based cron system to a more robust Airflow system deployed in a Kubernetes cluster and in the cloud
- Developed and maintained full stack docker images using Python, R, Airflow, and Shiny to automate the acquisition of spatial-temporal data for the WSTAMP project for storage in an Hadoop Data Warehouse
- Led the development team for a COVID-19 tracking project using both R and Python and downstream model creation and analysis
- Developed containerized Airflow ETL workflows, Python microservices, and REST APIs using Swagger for data exchange with PostgreSQL, AWS S3 & RDS, and MongoDB
- Automated web-scraping and social media ingestion via containerized model creation for misinformation classification for implementation in advanced statistical models and network analysis and ingestion into ElasticSearch and visualization in Kibana
- Served as Software Quality Assurance board chair for the Geospatial Science and Human Security Division
- Participated as a member of the ORNL Data Council, which discusses best practices in lab-wide data management and architecture for sharing of lab-generated data

Data Scientist II Subcontractor
Oak Ridge National Laboratory, Oak Ridge, TN *2018 —2019*

- Automated data ingestion of the WSTAMP Web Application via the use of R, Python, Airflow, and Docker
- Authored an R package to unify geography between online data APIs
- Developed a SpatiaLite SQLite database to ease lookup of geographic data
- Developed a zonal statistics application in Python that takes a raster image and shapefile and computes the statistics over the shapes stored in the shapefile

Owner & Contractor-For-Hire
Specrabella, Knoxville, TN *2018 —2018*

- Designed, developed, and deployed an interactive energy informatics leaflet-based web application integrated with Snowflake that provides real-time calculation from machine learning derived models
- Consulted with clients on ground-up web application architecture, database and visualization needs
- Developed data ingestion methods using cURL, PHP, and R to scrape energy data from the web
- Streamlined training and development procedures using ansible, vagrant, PHP, and BASH

Bioinformatician, Project Manager, & Technical Sales Representative
Microbial Insights, Knoxville, TN *2016 —2018*

- Established an automated pipeline using R, MySQL, Python, and Bash for NGS analysis, along with intranet website control written in JavaScript and Shiny
- Automated client-bound statistical calculations such as linear models, ANOVA, SOMs, clustering, PCA, and PCoA
- Developed an ETL customer data visualization tool using R, PHP, JQuery, and MySQL to view qPCR results in the context of other samples' and dynamically selectable parameters
- Optimized data delivery to clients via a custom R package and local shiny applications to quickly generate LaTeX PDF reports
- Constructed, populated, and maintained an intranet wiki based using PostgreSQL and PHP to aid in project management and customer service

Graduate Research Assistant, Dr. Neal Stewart's Plant Biotechnology Lab
University of Tennessee, Knoxville, TN *2014 — 2016*

- Summarized statistical findings of cell suspensions using linear models, ANOVA, and PCA in Python and R
- Evaluated suspension cultures via chemical and spectral processes for lignin formation and statistically analyzed and summarized my findings for inclusion in DOE reports
- Collaborated on a novel single cell suspension and cryopreservation robotic system
- Advanced and executed monocot genetic modifications for *Panicum* spp., *Oryza* spp., *Sorghum* spp., *Saccharum* spp., *Zea mays*

Laboratory Assistant, Dr. Neal Stewart's Plant Biotechnology Lab
University of Tennessee, Knoxville, TN *2012 — 2014*

- Developed automated statistical methodology for screening of lignin content and imaging of cell characteristics both *in vivo* and *in vitro* using Python
- Extracted genomes from NCBI, cleaned and normalized the data, and performed exploratory data analysis using Python
- Complied with USDA-APHIS regulations regarding transgenic plants
- Implemented an *E. coli* bioreactor for production of proteins

Laboratory Assistant, Dr. Paris Lambdin's Biosystematics and Biological Control Lab
University of Tennesse, Knoxville, TN *2002 —* 2012

- Designed instructional modules for undergraduate and graduate level courses using HTML, CSS, JavaScript, and Flash
- Collected samples and data from field locations
- Assisted in community outreach programs such as Bloomsdays, Buggy Buffet, and 4-H Camps
- Performed forest coverage analysis using ArcGIS

🏛 **Education**

Master of Science in Plant Sciences —Plant Molecular Genetics
University of Tennessee — Knoxville, TN — *Spring 2017* — **GPA: 3.72/4.0**

Bachelor of Science in Plant Sciences —Biotechnology
University of Tennessee — Knoxville, TN — *Spring 2014* — **GPA: 3.74/4.0**
Magna Cum Laude

📖 Publications

[1] N. Labbe L.M. Kline M. T. Windham D.G.J Mann C. Fu J.N. Grant C.R. Poovaiah H. Shen W. A. Wuddineh A. Ziebell M. F. Davis F. Chen T. J. Tschaplinski J. R. Mielenz Z-Y Wang R. A. Dixon C. N. Stewart Jr Baxter, H.L. M. Mazarei. Field-based experiments on low-lignin switchgrass as a feedstock for lignocellulosic biofuel production. 2012.

[2] N. Labbe L.M. Kline M. T. Windham D.G.J Mann C. Fu J.N. Grant C.R. Poovaiah H. Shen W. A. Wuddineh A. Ziebell M. F. Davis F. Chen T. J. Tschaplinski J. R. Mielenz Z-Y Wang R. A. Dixon C. N. Stewart Jr Baxter, H.L. M. Mazarei. Field-based experiments on low-lignin switchgrass as a feedstock for lignocellulosic biofuel production. 2012.

[3] Joshua N Grant, Jason N Burris, C Neal Stewart, and Scott C Lenaghan. Improved tissue culture conditions for the emerging c 4 model *Panicum hallii*. *BMC biotechnology*, 17(1):39, 2017.

[4] Joshua N. Grant, Jonathan D. Willis, and C. Neal Stewart. Functional analysis of a putative membrane-bound switchgrass endo-β-1,4-glucanase in *Escherichia coli*. 2012.

[5] Joshua N Grant, Jonathan D Willis, and C Neal Stewart. Comparisons of switchgrass cell wall components containing an overexpression of a putative switchgrass endoglucanase. In *IN VITRO CELLULAR & DEVELOPMENTAL BIOLOGY-PLANT*, volume 49, pages 475–475. SPRINGER 233 SPRING ST, NEW YORK, NY 10013 USA, 2013.

[6] Joshua Nathaniel Grant. Evaluation of hall's panicgrass (*Panicum hallii* vasey) as a model system for genetic modification of recalcitrance in switchgrass (*Panicum virgatum* (l.)). Master's thesis, 2017.

[7] Dasha Herrmannova, Gautam Thakur, Josh Grant, Varisara Tansakul, Bryan Eaton, Olivera Kotevska, Jordan Burdette, Martin Smyth, and Monica Smith. Challenges in automated detection of covid-19 misinformation.

[8] Jinyi Liu, Nana Chen, Joshua N Grant, Zong-Ming Cheng, C Neal Stewart Jr, and Tarek Hewezi. Soybean kinome: functional classification and gene expression patterns. *Journal of experimental botany*, 66(7):1919–1934, 2015.

[9] Robert Stewart, Jesse Piburn, Sarah Walters, Jason Kaufman, Evan Ezell, David Anderson, David Axley, Josh Grant, Bryan Eaton, Alexandre Sorokine, et al. A taxonomic classification approach for global spatio-temporal data. 2021.

[10] Jonathan D Willis, Joshua N Grant, Mitra Mazarei, Lindsey M Kline, Caroline S Rempe, A Grace Collins, Geoffrey B Turner, Stephen R Decker, Robert W Sykes, Mark F Davis, et al. The tceg1 beetle (tribolium castaneum) cellulase produced in transgenic switchgrass is active at alkaline ph and auto-hydrolyzes biomass for increased cellobiose release. *Biotechnology for Biofuels*, 10(1):230, 2017.

👤 References

References available upon request.