Justin Gosses

jgosses82@gmail.com – https://justingosses.com – https://github.com/justingosses

Professional History

Senior Data Scientist **S.A.I.C.** [contractor embedded in NASA OCIO] **2018-present**
Reporting to the Agency Data Scientist, my directives included accelerating adoption of modern data science methods at NASA; promoting open-source and inner-source code sharing; and helping the public take advantage of NASA's publicly released APIs, code, and data as technical program manager for open-innovation program.

- **Manager of 2 software engineers & 2 interns.** Led daily stand-ups and applied agile, lean, and design thinking methods geared to task type: web, data science, or platform/process development.
- **Advised senior management on foundational policy documents impacting data science**, including agency Metadata Standards, agency API Guidance, and AI Ethics Principles.
- **Reduced project initiation time** for teams developing speech-to-text applications through sharing of vendor evaluations, audio preparation code, model evaluation code, and guides mapping technology tradeoffs to use cases. Advised mission engineers building on my earlier work to create speech interfaces for astronauts.
- **Improved information discovery using natural language processing model** to tag text with 7000 NASA concepts trained on 3 million human tagged texts. Deployed to code.nasa.gov and internal WordPress plugin.
- **Replaced a manual business task with A.I. solution**. Led a small team to curate a dataset & train a natural language processing model to predict correct meanings of 1000s of NASA-specific acronyms when multiple meanings exist. Model handed over to a team who will operate it as API and browser add-on service.
- **Championed open-source code as admin of github.com/nasa**. Brought on new members, solved policy & technical blockers to enable use cases, and managed pull requests from the public across 5 repositories.
- **Successfully advocated for users not served by existing capabilities** by convincing senior management to fund a code sharing platform enabling NASA to privately collaborate with users without NASA identities.
- **Sourced feedback during SpaceApps, world's biggest hackathon,** on api.nasa.gov & data.nasa.gov then led a team of 4 to develop 6 new features on public facing websites visited by 10,000s people a day.
- **Led collaborations between** office of chief information officer, office of chief engineer, and software release authority to create an inner-source program that reduced the process burden on code re-users from days to seconds and resulted in 75+ new inner-source tagged projects within a year. Ran program operations.

Software Developer **Valador Inc.** [contractor embedded in NASA OCIO] **2016-2018**
As member of the data analytics lab reporting to the agency data evangelist, spread awareness & promoted adoption of emerging data science technology through consulting services, blog posts, and reusable code.

- **Built minimal viable products, recommended architectures, and led problem framing discussions** with 10+ teams across a range of domains, matching business questions with AI / ML solutions.
- **Taught staff in Human Resources** how to use R and Shiny to automate data processing tasks and build self-service business intelligence dashboards that sped deliver of data to decision makers.
- **Replaced a previously human intensive task with a Machine Learning model** with 97% accuracy that disambiguated authors with same name for group responsible for cataloging all NASA publications.
- **Identified process changes necessary to scale an A.I. solution** in successful proof of concept to detect the IPs most likely to be flagged in near future as malicious in terabytes of network traffic. Used Dask & TensorFlow.
- **Successfully demonstrated the value of new technology** to Office of Chief Financial Officer (OCFO) by building an interactive web visualization that mapped spending to strategic goals and objectives, replacing traditional static documents. Directly influenced OCFO to hire their own data science and data visualization staff.

Geoscientist **BP Exploration and Production** **2006-2015**
Worked in international cross-functional teams to create data acquisition plans, integrate wide range of data types, analyze information, make predictions, and communicate results that drove business decisions.

- **Predicted** favorable attributes in well logs using random forests and support vector machines.
- **Applied statistics, decision trees, and Monte Carlo simulations** in natural science and business domains to minimize risk in business decisions at thousands to multiple million-dollar scale in 3 roles.

- **Increased number of active users** of a research consortium's data products by 20%. Identified time required to determine if results relevant to business problems as major barrier and writing short problem specific summaries.
- **Taught** geoscience to 100+ technical and nontechnical staff over 4 years in classroom & field settings.
- **Taught myself neural networks in 2007** to identify what combination of data needed to be available to predict reservoir connectivity pre-drill. Ended 3 years of repetitive wasteful studies increasing organizational efficiency.

Computer Skills

- **Language + Databases -** Python, JavaScript, Ruby, R, Bash, Java, SQL, Neo4J, BigQuery, & Snowflake.
- **Cloud System Administration -** Admin AWS cloud environment responsible for user management, updates, and security. Developed & deployed machine learning & web applications in GCP. Use Kubernetes, Docker, and Airflow. Familiar with azure but not used in professional setting.
- **Web Development -** Experienced in Flask.py, Node.js, JQuery.js, Angular.js, JavaScript, HTML, CSS, PHP, & WordPress across multiple projects. Have used React.js, Django, Kafka, and ElasticSearch.
- **Data Visualization & B.I. tools -** d3.js, vega.js, SandDance, Tableau, three.js., mapbox, ESRI certification
- **Machine Learning -** Scikit-learn, TensorFlow, Keras, Dask, H2O, CMUSphinx, DeepSpeech, NLTK, Gensim, Word2Vec, & spaCy. Alteryx. Jupyter Notebooks.

Selected Talks

- Gosses, J.C., Buonomo, A.R., Thomas, B.A., Yates, E.T., Yuan, R.W., (2019), Reusing Data and Metadata to Create New Metadata through Machine learning & Other Programmatic Methods, Abstract IN23D-0898 presented at 2019 Fall Meeting, AGU, San Francisco, CA, 9-13 Dec.
- Gosses, J.C., Zhang, L. (2019). A Supervised Machine Learning Approach to Stratigraphic Surface Picking in Well Logs from the Mannville Group of Alberta, Canada, American Association of Petroleum Geologists Annual Conference and Exhibit, San Antonio, 2019.
- Gosses, J.C., Lin, Y.I., (2017). Practical Considerations for Data Science Consulting and Innovation in a Large Organization, Rice Data Science Conference, Houston Texas, 2017.

Community Outreach

- **Co-lead Houston Data Visualization Meetup -** Delivered talks that introduced audiences to different data visualization methods and led Saturday Data Jams to collaboratively explore, practice, and share new skills.
- **Social Media for Gulf Coast Section of SEPM (geology society) –** Invited to create the role of social media chair. Expanded digital outreach, brought in new members, and measured reach via quarterly data reports.
- **Active Member in the online community, SoftwareUndergound –** A non-profit digital-first society at the intersection of computing and subsurface geoscience. Led hackathon challenges, gave short-talks, active Slack participant, and created 'awesome-open-geoscience', the top GitHub repository in "geoscience" topic.
- **Creator of Wellio.js & Wellioviz.js –** two open-source JavaScript libraries to assist in visualizing well logs. Fills a gap and enables interactive display of a niche scientific file type on the web. Inspired contributions from developers through word-of-mouth, demos, GitHub searches, & hackathons. Early work funded by a start-up and currently being used by a large business intelligence company.

Education

Master of Science – Geoscience – 2006 – University of Wisconsin – Madison
Bachelor of Science – Geoscience – 2004 – Franklin & Marshall College, Pennsylvania