· Justin Gosses · Houston, Texas USA · jgosses82@gmail.com · https://justingosses.com ·

Justin Gosses

Summary

Offers experience leading inner source, open source, and open data programs and a successful track record aligning stakeholders and building tooling that improves code sharing effectiveness and efficiency. Seeks opportunity to make large scale impact in the inner source problem space through new policy, tools, and features that generates insights, lowers friction, and nudges developers in ways that increase organizational productivity.

Experience

2018/04 - current **Technical Program Manager & Principal Data Scientist** S.A.I.C. contractor to NASA

Embedded in NASA's Office of Chief Information Officer Transformation & Data Division with mandate to improve NASA's ability to work with code and data. As technical program manager of open-innovation program, responsible for operations and new feature development of several public and internal facing IT systems, including those involved with inner source and open source. As principal data scientist, responsible for analyzing data for executives, delivering software products, contributing to data infrastructure efforts, and advising senior leadership.

- *Created a supportive team environment* where open innovation program & data analytics staff could excel on a remote team. Ensured staff across multiple time zones had psychological safety, technical guidance, development opportunities, and management support to help them achieve career goals.
- *Shepherded dozens of teams* through open source release process and developing in public as administrator of gitub.com/nasa. Wrote documentation, created processes, and ran operations with two junior staff.
- *Negotiated free seats for contributors of inner source* suitable code on the only NASA-wide code platform, leading to 75+ code repositories newly tagged as inner source in the first year.
- *Landed changes in NASA-wide policy documents* through collaboration with Legal and Office of Chief Engineer fellows that removed a burdensome requirement to use a specific IT system.
- *Reduced developer uncertainty about inner source requirements* by writing a single document that summarized many applicable policy documents and creating a README template that ensured all required information, disclaimers, and agreements were present.
- *Built a web application* to collect and share ideas for reusable code as an experiment in connecting developers across organization boundaries before code is written.
- *Collaboratively developed* code to measure probable code reuse between federal agencies using a novel technique to overcome data limitations. Shared with industry peers, other U.S. federal agencies.
- *Persuaded senior executives* of the need for a code platform exclusively for secure, private code collaborations with external partners who can not get inside the NASA firewall. Collaborated with another team to capture user stories, develop requirements, and usher through IT security approvals a solution that will radically reduce inefficiencies for collaborations with universities & other government agencies.
- *Advised senior leaders* & collaborated with committee members to write organization-wide data documents, such as the Agency API Guidance, AI Ethics Principles, and agency metadata standard.
- *Increased value of a natural language processing model* originally trained for 1 customer to predict 7000 NASA concepts trained on 3 million documents by building production APIs & pipelines to automatically tag content improving search & discovery on code.nasa.gov & internal WordPress websites.
- *Maintained 8+ open source repositories* including the most popular NASA API, astronomy photo of the day API. Integrated contributions from developers across the world with a range of skill levels.
- *Led a small team to improve user experience* on websites with >10,000 weekly users (data.nasa.gov, code.nasa.gov, and api.nasa.gov) based on hypotheses confirmed through interviews and modeling.
- *Created software to automate metrics* describing both NASA's public datasets, code, and APIs as well as how the public engages with those assets. Provided insights into what public finds useful.

2016/04-2018/04 **Software Developer** Valador Inc. contractor to NASA

Software Developer reporting to the NASA's data evangelist. Increased adoption of emerging data science technology and modern development practices through consulting services, minimum viable products, blog posts, and reusable code.

- *Led problem framing discussions*, recommended architectures, built proof of concepts, and developed working prototypes with 12+ teams across a range of domains as part of a consulting data science team.
- *Evaluated competing applications & code packages* in AutoML, business intelligence, instant message, and other product categories delivering analysis to managers and sharing via internal blog posts.
- *Created "dependencies are skills" knowledge graph* from code dependencies of U.S. federal agency open source. Demonstrated ability to recognize people doing similar work across organizational boundaries. Wrote python dependency extraction code before GitHub provided their dependency graph.

· Justin Gosses · Houston, Texas USA · jgosses82@gmail.com · https://justingosses.com ·

2006/12-2015/04 **Geoscientist** BP Exploration and Production

Geoscientist working in international cross-functional teams to analyze information, generate hypotheses, design experiments, mitigate uncertainty, make predictions, and communicate results that enable decisions.

- *Impacted 3 software choices* via requirements generation, performance testing, and usability testing.
- *Demonstrated ability to quickly immerse myself in teams* as internal charging consultant to 6 teams.
- *Taught* 100+ staff from multiple disciplines & departments over four years in classroom & field settings.
- *Collaborated with international coworkers* across multiple time zones to develop new techniques combining Python & ArcGIS that enabled reservoir quality risks to be identified early than standard methods.
- *Applied statistics, decision trees, random forests, Monte Carlo simulations, as well as other types of uncertainty and risk management methods* to business decisions up to billion-dollar capital investment decision. 3 largest investment recommendations have all been shown to be profitable.
- *Increased organizational efficiency* by ending 3 years of repetitive wasteful studies. Taught myself neural networks in 2007 to identify what combination of data needed to be available to predict reservoir connectivity pre-drill.

Selected Talks, Open Source, and Community Outreach

- **TALK: Improved code discovery using machine learning** : 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, 2019 Fall Meeting, AGU, San Francisco, CA, 9-13 Dec.
- **TALK: Novel way to frame a problem to make it suitable for machine learning**: 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. Turned into open source project, Predictatops.
- **TALK: Taught technical audience about social aspects of 'doing' data science**: 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.
- **Open Source: Wellio & Wellioviz**: Created JavaScript libraries providing an easy solution for visualization of a scientific data format on the web. Success demonstrated by public contributors, funding for continued open source development, and use by companies in their products.
- **Open Source: Awesome Open Geoscience & visualization of it's community**: Initiated the awesome list creation & integrated efforts of 54 collaborators. Top ranked repository on GitHub.com in geoscience topic. Developing an approach to make visible the implicit community of developers that can be mined from the metadata of the repositories found in the awesome list.
- **Community Outreach: Co-lead Houston Data Visualization Meetup**: Introduced audiences to data visualization methods and led Saturday Data Jams to collaboratively explore, share, & practice new skills.

Computer Skills

- Languages: Python, JavaScript, Ruby, R, Bash, and Java
- Databases : SQL, PostsreSQL, Neo4J, Snowflake, and BigQuery.
- Data Visualization: d3.js, vega.js, Tableau, three.js., mapbox, and ESRI certification (GIS).
- Code Platforms: Internal subject matter expert in GitHub and user of GitLab.
- System Administration: Admin AWS cloud environment responsible for user management, updates, & security. Familiar with AzurePipelines part of Azure DevOps & GCP. Use Docker & Airflow.
- Web Development: Experienced in Flask.py, FastAPI.py, Node.js, JQuery.js, Angular.js, JavaScript, HTML, CSS, PHP, and WordPress across multiple projects. Have used React.js, and Django.
- Machine Learning : Scikit-learn, TensorFlow, Keras, Dask, H2O, CMUSphinx, DeepSpeech, NLTK, Gensim, Word2Vec, and spaCy. Jupyter Notebooks.

Education

2004-2006 Master of Science, Geoscience University of Wisconsin – Madison

2000-2004 Bachelor of Science, Geoscience Franklin and Marshall College, Pennsylvania