Dotscience Emerges from Stealth to Eliminate the Biggest Pain Points of Operationalizing AI in the Enterprise

DevOps for ML platform to simplify, accelerate and control all stages of the AI model lifecycle for industries including fintech, autonomous vehicles, healthcare and consultancies

LONDON and SAN FRANCISCO – July 30, 2019 — [Dotscience](), the pioneer in DevOps for machine learning (ML), today emerged from stealth with its platform for collaborative, end-to-end ML data and model management. By giving teams the unique ability to collaboratively track runs—a record of the data, code and parameters used when training an AI model—-Dotscience empowers ML and data science teams in industries including fintech, autonomous vehicles, healthcare and consultancies to achieve reproducibility, accountability, collaboration and continuous delivery across the AI model lifecycle. The Dotscience platform is now available as SaaS or on-prem, and on the Amazon Web Services (AWS) Marketplace in August.

"The current state of AI development is a lot like software development in the 1990s. Before the movement called DevOps, modern best practices such as version control, continuous integration and continuous delivery were far less common and it was normal that software took six months to ship. Now software ships in minutes," said Luke Marsden, founder and CEO of Dotscience. "At Dotscience, we are applying the same principles of collaboration, control and continuous delivery of DevOps to AI in order to simplify, accelerate and control AI development."

AI Development and Operations Challenges Today
Data science and machine learning teams commonly face a multitude of issues that make ML projects more likely to fail and create financial, reputational or legal risks for the business. These include wasted time, difficulties collaborating, mistakes made when manually tracking data, no reproducibility or provenance, lack of automated testing, manually deploying models, unmonitored models and losing track of what is running and where it came from resulting in "snowflake deployments."

According to Deloitte's "[State of AI in the Enterprise, 2nd Edition]()," the majority of respondents cited "implementation, integration into roles and functions, and measuring and proving the business value of AI solutions as top challenges of AI initiatives." Expanding on this observation, according to findings from Dotscience's "*[The State of Development and Operations of AI Applications 2019](*)*" market research also released today, the top three challenges respondents experienced with AI workloads are duplicating work (33.2%), rewriting a model after a team member leaves (27.8%) and difficulty justifying value (27%). The report evaluates how businesses are deploying AI today and investigates the need for accountability and collaboration when building, deploying and iterating on AI.

"Data scientists and ML engineers may not even be aware of the problem they have yet because they are accustomed to working with broken processes and are not aware of the

solutions available to do ML better," explained Marsden. "Solving these issues promises more productive, effective AI teams and better and safer ML models."

"Reproducibility is fundamentally important if you're putting machine learning applications into production," said James Kobielus, lead analyst for artificial intelligence and DevOps with SiliconANGLE's Wikibon team. "Dotscience's ability to track AI training runs, maintain a complete audit trail, and provide total visibility into a machine-learning app's provenance makes it well suited to this growing enterprise imperative. Just as important, Dotscience's ability to ensure reproducibility across hybrid-cloud platforms ensures reproducibility across the complex DevOps tool chains in today's enterprise AI environments."

The Dotscience Platform Delivers End-to-End ML Data and Model Management

Dotscience provides a tool that manages the complete AI lifecycle by empowering data scientists and ML engineers to work in ways in which they are familiar. Data science and ML teams can take advantage of a platform that is easy to use and provides a single place to collaborate on, develop, test, monitor and deliver their ML projects.

"In practical terms, and unlike other offerings on the market, this means that teams can continue using the same development tools, ML frameworks, languages, data sources and compute instead of being forced into a walled garden which risks vendor lock-in and steep learning curves," said Mark Coleman, VP of Product and Marketing at Dotscience. "Because Dotscience tracks and packages together every run that goes into the data engineering and model creation process, users can replicate each other's work, collaborate easily and track back as needed."

Dotscience offers data science and ML teams the following key benefits:
- **Seamless flexibility and integration all from one platform:** Dotscience users can easily attach any compute to the platform, whether it is their own laptop, cloud-based VMs or on-prem bare metal. After a user then trains a model, Dotscience integrates with continuous integration and monitoring tools so that they can deploy and then monitor the models in production, keeping all relevant information in one place.
- **Optimal team productivity:** By providing an automated ML knowledge base to eliminate silos, Dotscience removes the "key person risk," making it easy for any data scientist or ML engineer to pick up where another left off—an attribute that is especially important in today's competitive hiring landscape. Dotscience allows teams not only to collaborate seamlessly but also to discover previous work and see exactly how it was built by tracking every version of every element in the model development phase.
- **Flexible access to compute, hybrid cloud portability for ML development environments:** Team members can start working on their laptop, then move their AI workload to a bigger cloud machine or a bare metal GPU rig when they need extra power, all seamlessly and without having to create a support request. The entire package of code, data, environment and hyperparameters that are needed to reproduce the development environment is bundled up and packaged together in such a way that moving from one cloud to another or on-prem is seamless.

- **Ability to work with data from any source:** Dotscience works with flat files stored directly in Dotscience, data in remote object storage (i.e., S3 or S3-compatible, Azure or GCS) and data from SQL, NoSQL and Spark data lakes. This flexibility allows data science and ML teams to get started immediately with whichever data sources are already in use. Dotscience doesn't force the ingest of all data; it can track the provenance of data where it already exists, given a compatible object store.
- **Allows AI and data science teams to use the tools they care about, while removing the obstacles that aren't central to productivity**: Using Dotscience's tracked workflows, data scientists and ML engineers can use open source tools for model training with which they are familiar and love, such as PyTorch, Keras and TensorFlow. They can use Jupyter notebooks natively in the application or choose to work on the command line enabling them to use any IDE of their choice.
- **Guarantees compliance with current and future regulation:** ML models are used to make decisions by design, but if decisions that are made are incorrect, it can lead to serious financial, reputational and legal risk. Dotscience both monitors ML models to detect issues early and also makes it possible to forensically reproduce any issues that occur so they can be quickly addressed and fixes confidently deployed.

Dotscience DevOps for ML Platform Now Available as SaaS, On-prem or Through the AWS Marketplace

Dotscience provides end-to-end ML lifecycle management without forcing users to change their working practices and this approach also extends to the installation options.
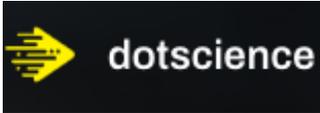
Customers can choose to deploy the hosted SaaS and bring their own compute, or install a fully private version of Dotscience either manually, or through the Dotscience installer in the AWS Marketplace which will be available in August. Installers for Microsoft Azure and Google Cloud Platform will soon be available as well. This flexibility means that a broad userbase can access an integrated ML platform that provides unified version control and collaboration for data scientists.

Dotscience is Trusted by AI Leaders

"The world of ML has a lot to learn from all the best practices developed to handle the Software Engineering lifecycle in the last 10 years. Dotscience has the potential to bring some of those hard-learned lessons to the ML world without forcing data scientists and researchers to completely abandon their tools of choice, like Jupyter Notebooks. It's a bold proposition and has the potential to make a huge impact."

— [Luca Palmieri](), Machine Learning and Data Engineering at TrueLayer

"The processes and tools for collaborating and maintaining ML projects at industrial scale are not yet as mature as for traditional software projects. The ML workflows pose several additional challenges that don't perfectly fit into Software DevOps processes. I am excited to work with

Dotscience to tackle these challenges in our upcoming project, as they are actively focused on making collaboration structured and centralised so that it scales to much larger team and project sizes."

— [Anders Åström](#), Datascience Manager at a global Technology Consulting firm

"The Dotscience product fills a critical gap when it comes to ensuring data provenance for machine learning models. By providing data provenance as a service, Dotscience tracks work without slowing down the Data Science team and gives great visibility into the integrity of the data and the process needed to ensure credibility to key business stakeholders."

— [Terry MacGregor](#), Founder, CTO at LawIQ

Additional Resources
- Visit [dotscience.com](#)
- Follow [@getdotscience](#) on Twitter and [LinkedIn](#)
- Read the Dotscience [blog](#)
- Dotscience's *[State of Development and Operations of AI Applications 2019 Report](#)*
- Dotscience's [DevOps for ML Vision](#)

About Dotscience
Dotscience, the pioneer in DevOps for machine learning (ML), brings DevOps principles followed by high-performing software teams to ML and data science. The Dotscience software platform for collaborative, end-to-end ML lifecycle management empowers ML and data science teams in industries including fintech, autonomous vehicles, healthcare and consultancies to achieve reproducibility, accountability, collaboration and continuous delivery across the AI model lifecycle. Founded in 2017 by container storage veteran Luke Marsden, Dotscience is headquartered in the UK with offices in the US. Its mission is to accelerate and unlock the true value of data and analytics assets through AI.

####

Media Contact:
Jocelynn Stidham
Bhava Communications for Dotscience
(703) 863-1277
dotscience@bhavacom.com