

20240613 dbt links

dbt for Machine Learning

- [dbt_ml_preprocessing](#) package for light feature engineering
- [Article on ELT-ML pipeline](#) inside dbt and how ML team could benefit
- [Meetup talk video](#) by Sam Swift, VP of Data & AI at Bowery Farming
- [Data Science use cases breakdown](#) directly from the dbt Labs

dbt + Airflow

- [Cosmos package](#) for creating Airflow tasks/jobs from dbt project
- [dbt+Airflow](#) by Astronomer authors of the package

dbt + Dagster

- [Dagster + dbt](#) - official article introducing to dbt support

Write-Audit-Publish

- [WAP: dbt + BigQuery](#) - intro video
- [WAP: dbt + BigQuery](#) - PDF from the above presentation
- [Conf talk referencing WAP strategy](#) by employees video intro by Michelle Ufford, Staff Engineer at Netflix

dbt-core technicalities

- [analytics engineering handbook](#) by Hifly Labs
- [Talk on custom CI setup for dbt](#) with automated reports a'la terraform plan in GitHub Actions
- [Article on User Defined Functions](#) managed inside dbt project and included in lineage
- [dbt guide](#) by GitLab Data team

other

- [DataOps Manifesto](#) that you can sign on
- [awesome-dbt](#) collection on GitHub

Companies deploying and utilizing dbt

Company	Location	Technologies and Tools	Founding Year	Clients	Website
Rittman Analytics	London, UK	dbt, Looker, Snowflake, Google BigQuery,	2017	Rittman Analytics Clients	rittmananalytics.com

Company	Location	Technologies and Tools	Founding Year	Clients	Website
		Segment, Fivetran, Stitch, Supermetrics			
HiflyLabs	Budapest, Hungary	dbt, Tableau, Power BI, Alteryx, AWS, Snowflake	2012	Erste Bank, Magyar Telekom, Sanofi, MOL Group, Philip Morris	hiflylabs.com
DataRoots.io	Leuven, Belgium	dbt, Kubernetes, Docker, Apache Spark, TensorFlow, Kafka, Databricks	2016	KBC, Proximus, Port of Antwerp, Barco, Gimv, Fluxys	dataroots.io
calogica.com	San Francisco, USA	dbt, data warehousing, analytics, Looker, Snowflake	2019	TBA	calogica.com
elementary-data.com	Remote	dbt, Data observability, data quality, analytics	2021	TBA	elementary-data.com
re_data	Remote	dbt, Data observability, monitoring	2020	TBA	getre.io
esentri.com	Karlsruhe, Germany	dbt, API management, cloud services, process automation	2009	TBA	esentri.com
analytics8.com	Chicago, USA	dbt, Data analytics, BI, data management, data visualization	2002	McDonald's, Blue Cross Blue Shield, Randstad, Geico, Northrop Grumman	analytics8.com
altimate.ai	Singapore	dbt, AI, machine	2020	TBA	altimate.ai

Company	Location	Technologies and Tools	Founding Year	Clients	Website
		learning, big data			
DataCoves	USA	dbt, data engineering, analytics	2020	TBA	datacoves.com
Y42	Berlin, Germany	dbt, Data integration, data pipelines, BI	2020	BMW, P&G, HelloFresh, BuzzFeed, Klarna	y42.com
Estuary.dev	New York, USA	dbt, Real-time data integration, data pipelines	2019	TBA	estuary.dev
kuwala.io	Munich, Germany	dbt, Data integration, data enrichment, data connectors	2021	TBA	kuwala.io
Vaultspeed	Leuven, Belgium	dbt, Data vault automation, data warehousing	2018	TBA	vaultspeed.com
Guinsoolab	Remote	dbt, Data observability, data reliability	2019	TBA	guinsoolab.com

Anything missing? Please fill in feedback <https://tally.so/r/mKoVZK> or contact me directly:
pawel@bogumilo.co