

Daniel Pascal Lamblin		m: dlamblin@gmail.com
Brooklyn NY U.S.A. & Remote		t: +1 646 644 3587
		gh: dlamblin.github.io
		in: linkedin.com/in/daniellamblin
Skills	<i>JVM:</i> Java, Guava, Guice, Dagger, SpringBoot, Lucene, GWT, Scala & Kotlin; <i>Unix:</i> Python, Go, C, Perl & Bash; <i>C#:</i> ASP.Net; <i>Data:</i> Airflow, MySQL, Hive, Presto, Spark, Megastore, Beam/Cloud Dataflow; <i>Web:</i> NodeJS, Vue.js, HTML, CSS & JavaScript; <i>AWS:</i> ALB, ASG, EBS, EC2, EFS, EKS, EMR, Lambda, RDS, CloudTrail, CloudWatch, CloudFormation, S3 ...; Agile & TDD	
Education	Andrew Ng's Stanford ML-Class.org; Online – Became Coursera	Oct 2011–Dec 2011
	Columbia University; New York, NY – Non-Degree Semester	Sep 2007–Dec 2007
	Worcester Polytechnic Institute; Worcester, MA – BS Computer Science	Sep 1995–May 1999
Experience	Coufang , <i>Senior Backend Software Engineer</i> Clickstream Web Logging	Seattle Nov 2020–Present
	<ul style="list-style-type: none"> Developed Customer Experience Analytics Platform tool for self-service funnel, journey and trend analysis of web logging data MVP for CXAP tool: Stood up the web front on Vue.js 2 for customers of Business Analytics team Designed the scheduling of CXAP jobs Delivered Journey and Trend analysis for backend of CXAP with Spark and Scala Mentored intern project to port and extend UI on Vue.js 3 & converted to hire Assisted Retail Delivery team's mission-critical multi-AZ expansion projects 	
	<i>Senior Data Engineer</i> Data Platform Tools and Infra	Seattle Jul 2019–Nov 2020
	<ul style="list-style-type: none"> Migrated enterprise data warehouse's ~4000 Airflow ETL jobs from Airflow 1.8 to 1.10+, with custom semi-automated tooling. Right-size teams' infrastructure resource usage by monitoring for efficient use of EMR and Spark jobs Split Airflow into multiple deployments and onboard users to latest versions On call rotation as in Seoul. 	
	<i>Big Data Engineer</i> Big Data Platform	Seoul Aug 2017–Jul 2019
	<ul style="list-style-type: none"> Onboard teams (edw, growth, pricing, catalog, search, retail etc.) into airflow; configure and maintain EMR clusters for teams' jobs Develop Data Platform Portal tool with Cluster start-stop API, management, and Data Discovery Scale out Airflow to ~6000 dags with multiple deployments On call rotation for Airflow, Presto, Hive, Hue, Zeppelin, Spark, HDFS, Zookeeper etc. 	
	<i>Data Engineer</i> Data Platform	Seattle Jul 2016–Aug 2017
	<ul style="list-style-type: none"> Cloud migration. Lift and Shift on-prem IBM Netezza and HDFS Hive to EMR Hive on S3 Update ETL into Airflow from Oozie and Talend Monitor data readiness with oncall rotations 	
	Insight Data Science , <i>Data Engineering Fellow</i>	New York City Sep 2015–Jan 2016
	<ul style="list-style-type: none"> Realtime and batch processing on NYC MTA's GTFS stream with Kafka, Spark, HDFS, HBase, S3 Generated user data in large scale for testing aim of notifying users of train delays Cluster on AWS EC2; project information and presentation at dlamblin.github.io/mta-delay-monitoring 	
	Google , <i>Software Engineer</i>	New York City Jul 2010–Nov 2014
	<ul style="list-style-type: none"> Ported legacy Studio product, a rich advertisement authoring and QA web-app, to Google's Web Toolkit front-end with a stubby rpc backend and megastore datastore Developed a dashboard to track component usage data of ads by comparing html5 vs. flash authoring, common formats, layouts, and generated impressions. Utilized cross-team apis and internal versions of GFS, Cloud Dataflow and Drill Implemented critical preview features for monitoring ad unit interactions and compliance Migrated user records and assets to support new multi-account users and unified asset library view Reduced reprocessing and conserved storage of assets by fingerprinting uploads, both on individual files and within archives 	
	Sigato Group , <i>Software Engineer</i>	Remote Aug 2009–Jan 2010
	<ul style="list-style-type: none"> Engineered provider search functionality for New West Health with R-tree based search, maps, and directions, as a module for Drupal using PHP, MySQL, and Perl Integrated single sign-on to converge features from New West Health and partners 	
	Travelocity , <i>Senior Web Application Developer</i>	New York City Jul 2006–Jan 2009
	<ul style="list-style-type: none"> Established features for IgoUgo.com using the ASP.NET 2.0 framework, C#, PrototypeJS Introduced a Lucene based index of content to offload db search as a wsdl service in Java and Spring with auto-completed suggestions for key geo-entities Boosted traffic ten fold through optimization of page structure, URLs and image file names 	
	Hairzone Inc , <i>Head Web Developer</i>	Moonachie Oct 2004–Jul 2006
	<ul style="list-style-type: none"> Built and supported the multiple websites of four company brands with ASP 	
	Richmond Research Inc , <i>Programmer</i>	New York City Feb 2003–Oct 2004
	<ul style="list-style-type: none"> Automated reporting and developed web products using Perl, PHP, Visual Basic and ActionScript for clients including Wall Street Journal and Priceline 	
	Mile NY , <i>Website Programmer</i>	Fort Lee Oct 2002–Sep 2003
	<ul style="list-style-type: none"> Established MileNY's e-commerce site for using Perl, CGIs, and MySQL database 	

EMC Corp, *Associate Software Engineer*

Milford, MA Jun 2000–Oct 2001

- ▶ Extended EMC Data Manager Volume & Tape Library Manager's media duplication processes targeting Petabyte capable systems like Sony PetaSite with multiple robots and drives
- ▶ Improved and maintained EDM as a multi-process C based system with Sun RPC, threading, and IPC signals & pipes; Resolved deadlock by refactoring mutex hierarchy