# Dat Tran

Born April 30, 1988 in Hanoi, Vietnam

| | |
|---|---|
| Address: | Breite Straße 33 |
| | 13597 Berlin, Germany |
| ✉ E-Mail: | datitran@gmail.com |
| ☎ Phone: | +(49) 151 2296-3660 |
| ☉ Web: | https://dat-tran.com/ |
| in LinkedIn: | https://www.linkedin.com/in/dat-tran-a1602320 |
| ☉ Github: | https://github.com/datitran |
| ☉ Twitter: | @datitran |

## Summary

Experienced technology and business leader with a strong focus on smart (data) products. Passion for building up high performance, cross-functional and self-managed balanced teams (from scratch). Open-source and machine learning evangelist with global reach. Currently I'm heading Axel Springer AI which is the artificial intelligence unit of Axel Springer SE, one of the largest digital publishing house in Europe. My interests are diverse from traditional machine learning, deep learning, AI in general to computer vision. I have a lot of experiences from devising realistic data-driven use cases to the actual implementation into a real product; more than capable of distinguishing hype, buzzwords and wannabes from substance. Big fan of agile practices, lean product development and open-source technologies. I mainly use Python and the PyData stack for anything that I do but I also know a bit about Apache Spark, TensorFlow, R, SQL and many other stuff. I also regularly blog about my work on Medium and also speak at conferences such as PyData.

## Experience

- **Axel Springer Ideas Engineering GmbH**                                   Berlin, GER
  *Head of AI at Axel Springer AI*                                    *06/2019 - present*
  - Establish and lead Axel Springer AI, the artificial intelligence unit of Axel Springer SE, one of the largest digital publishing house in Europe
  - My team's research focus was mainly on Voice (TTS & STT) and NLP topics. We also open-sourced several projects:
    * ForwardTacotron (Generating speech in a single forward pass without any attention!) - https://github.com/as-ideas/ForwardTacotron
    * headliner (Easy training and deployment of seq2seq models.) - https://as-ideas.github.io/headliner/
  - Driving Axel Springer's change into a real "tech" company:
    * Established and managed the Axel Springer techblog
    * Co-organized the Axel Springer TechCon 2019 which was the first installment of a group-wide conference within the Axel Springer group
    * I'm the "face" of the new employer branding for technology within the Axel Springer group
    * Represented Axel Springer internally and externally at several renowned conferences, podcasts, magazines etc.
    * Established a more open-source and research-driven mindset
    * Evangelizer for Tech & Product first

- **Humydo - Pham & Tran GbR**                                               Berlin, GER
  *Co-Founder / Managing Director*                                    *12/2019 - present*

- – Humydo provides high-quality clothing for an affordable price

- **idealo internet GmbH** — Berlin, GER
  *Head of Data Products at idealo* — *10/2017 - 05/2019*
  - – Co-head of the data team (25+ people) with a special focus on building up the data science/data engineering team (built up from scratch - hired 6 people within 9 months)
  - – Led the data science team directly where my team's research focus is on:
    * Interpretability of deep learning models (mainly computer vision problems)
    * Efficient deployment of machine learning models
    * AutoML
    * Data collection process (in-house labeling and external mechanical turk approaches)
  - – Selected projects of my team:
    * We used deep learning to automatically rank millions of hotel images (300+ million) according to their aesthetic and technical quality which helped to improve the listing for small and mid-sized hotels
    * We trained CNNs to automatically generate nice looking product galleries which reduced the workflow time of our content team by 75%
    * We built a tree-based classifer to predict non-purchases of customers which helped our CRM team to target those customers three days ahead before we receive the actual sales data by the shops
  - – Provided technical and scientific guidance to senior team members and also our machine learning trainees
  - – Created the machine learning powered roadmap for idealo for both product and research
  - – Turned the data team from a pure cost center into a profit center e.g. monetarizing our data for CRM campaigns which resulted in EUR 1 million additional revenue
  - – Set up a regular knowledge sharing session internally and externally across the Axel Springer companies including Bonial, Upday etc.
  - – Established the open-source strategy under the idealoTech umbrella and executed it e.g. created and managed the idealoTech blog (`https://medium.com/idealo-tech-blog`), idealo's Github landing page and also the @idealoTech channel on twitter
  - – My team open-sourced several projects with high business impact:
    * image-quality-assessment (Convolutional Neural Networks to predict the aesthetic and technical quality of images.) - `https://github.com/idealo/image-quality-assessment`
    * image-super-resolution (Keras implementation of the Residual Dense Network for super scaling images.) - `https://github.com/idealo/image-super-resolution`

- **Springboard** — Remote, USA
  *Mentor for the AI/ML career track, `https://www.springboard.com/`* — *01/2019 - present*

- **Data Festival** — Munich, GER
  *Member of the program committee, `https://www.datafestival.de/`* — *05/2018 - present*

- **Thesis Advisor** — Berlin, GER
  *Co-supervised thesis* — *2018*
  - – Hong Ngoc Truong. Bachelor Thesis, co-supervisor with Prof. Dr. Manfred-Erich Jäger-Ambrozewicz, HTW Berlin - University of Applied Sciences, Germany. Title: "Machine Learning: Tree-Based Algorithms"

- **Pivotal Inc.** — Berlin, GER
  *Senior Data Scientist at Pivotal Labs* — *10/2015 - 10/2017*

- Leading and managing the Data Science practice for Pivotal Labs Berlin
- Created machine learning models and put them into production for various clients
- Supported pre-sales activities by creating demos/proof of concepts and blog articles
- Scoped data science projects including proposal writing, resource estimations and budget allocation
- Provided data science thought leadership, for example, helped to establish the smart apps idea and XP practices into data science like TDD (`http://tinyurl.com/gme62kr`), pair programming and CI/CD
- Established an API First culture within the company: `http://tinyurl.com/z9kv5ae`
- Gave talks at PyData Berlin, London, San Francisco 2016 and CF Summit Europe 2016
- Data Scientist from 10/2015 - 06/2016

**Accenture GmbH** — Berlin, GER
*Data Scientist at Accenture Digital* — *08/2014 - 09/2015*

- Built up internal capabilities and showcases in the area of network analytics, operations analytics, text analytics, customer analytics and data visualizations
- Worked on setting up an analytics center of excellence for a large German airline
- Led and co-ordinated workshops at customers to demonstrate the value of data science
- Acted as point of contact for internal requests of various advanced analytics topics such as machine learning and optimization
- Supported pre-sales activities by creating proof of concepts for clients in telecommunication, aviation and automotive
- Helped to organize internal and external events such as TEDxRheinMainSalon Datanauts and LMU DataCrunch

**Hotel Diana Berlin** — Schönwalde-Glien, GER
*Family Business, `http://www.hotel-diana-berlin.com/`* — *01/2018 - present*

- Managed online marketing activities (SEO, review portals and social networks)
- Created and managed the homepage

**Sy Restaurant** — Berlin, GER
*Family Business, `http://www.sy-restaurant.de/`* — *12/2006 - present*

- Supported the formation of a new branch in Berlin Spandau
- Developed and managed the service
- Managed online marketing activities (SEO, review portals and social networks)

# Education

**Humboldt University of Berlin** — Berlin, GER
*Master of Science in Business Administration, Grade 1.9* — *10/2011 - 12/2013*

- Field of Specializations: Operations Research and Econometrics
- Research Interests: Graph Theory, Queuing Theory and Dynamic Programming
- Master Thesis: Optimal Design for an Online Dutch Auction Selling Airline Tickets
  Advisor: Prof. Dr. Kurt Helmes, Department of Operations Research, Grade 1.7
  Digital Copy: `http://tinyurl.com/jsq8hwp`

**Berlin School of Economics and Law** — Berlin, GER
*Bachelor of Arts in Business Administration, Grade 1.5* — *04/2008 - 07/2011*

- Field of Specializations: Banking & Finance and Accounting

- – Bachelor Thesis: Best of N Concept - A Comparison of Dynamic Asset Allocation Strategies
    Advisor: Prof. Dr. Thomas Wolke, Department of Finance, Grade 1.7
    Digital Copy: `http://tinyurl.com/l84jskr`

- **Study Abroad at the University of Sydney** <span style="float:right">Sydney, AUS</span>
  *Exchange Program* <span style="float:right">*03/2010 - 08/2010*</span>
    - – Courses taken in Portfolio Management, International Financial Management, International Business Strategy, Calculus I and Linear Algebra I

- **Friedrich Anton von Heinitz Gymnasium** <span style="float:right">Rüdersdorf, GER</span>
  *Abitur, AP English and Chemistry, Grade 2.5* <span style="float:right">*07/2007*</span>

# Selected Online Publications

- **Deep Learning for Classifying Hotel Aesthetics Photos** <span style="float:right">NVIDIA Developer Blog</span>
  *https://tinyurl.com/yadn9wgk - Jointly with Christopher Lennan* <span style="float:right">*10/2018*</span>

- **What is Minimum Viable (Data) Product?** <span style="float:right">Medium - idealo Tech</span>
  *https://tinyurl.com/ybqs55bt* <span style="float:right">*06/2018*</span>

- **My two cents on what makes a good data scientist nowadays?** <span style="float:right">Medium - Towards DS</span>
  *http://tinyurl.com/y9grlbrg* <span style="float:right">*07/2017*</span>

- **8 "Simple" Guidelines For Data Projects** <span style="float:right">Medium - Built to Adapt</span>
  *http://tinyurl.com/m8nhv2o* <span style="float:right">*04/2017*</span>

- **Continuous Integration for Data Science** <span style="float:right">Pivotal Engineering Journal</span>
  *http://tinyurl.com/j6ghfht* <span style="float:right">*02/2017*</span>

- **Delivery in a Land with No Postcodes** <span style="float:right">Medium - Actionable Data</span>
  *http://tinyurl.com/gqat3bo - Jointly with Megha Agarwal* <span style="float:right">*11/2016*</span>

- **How Automakers And The IoT Are Colliding To Save Lives** <span style="float:right">Pivotal Blog</span>
  *http://tinyurl.com/jjcvaxs* <span style="float:right">*10/2016*</span>

- **Test-Driven Development for Data Science** <span style="float:right">Pivotal Engineering Journal</span>
  *http://tinyurl.com/gme62kr - Jointly with Megha Agarwal* <span style="float:right">*09/2016*</span>

- **API First for Data Science** <span style="float:right">Pivotal Engineering Journal</span>
  *http://tinyurl.com/z9kv5ae - Jointly with Alicia Bozyk* <span style="float:right">*07/2016*</span>

# Selected Talks

- **Future of AI Europe 2019** <span style="float:right">Berlin, GER</span>
  *Keynote Talk: The Future of Media is AI* <span style="float:right">*12/2019*</span>

- **WeAreDevelopers Congress Vienna 2019 -** `https://tinyurl.com/uzy7z8c` <span style="float:right">Vienna, AUS</span>
  *Deep Learning for Aesthetics: Training a Machine to See What's Beautiful* <span style="float:right">*11/2019*</span>

- **Digitale Leute Summit 2019** <span style="float:right">Cologne, GER</span>
  *Fireside Chat about Lean AI* <span style="float:right">*11/2019*</span>

- **Infoshare AI & Data Science** <span style="float:right">Warsaw, POL</span>
  *Deep Learning for Aesthetics: Training a Machine to See What's Beautiful* <span style="float:right">*11/2019*</span>

- **EGG Germany 2019** <span style="float:right">Stuttgart, GER</span>
  *Keynote Talk: From ML Research to Production - the Autobahn Way!* <span style="float:right">*11/2019*</span>

- **Future of Trust Summit** <span style="float:right">Frankfurt, GER</span>
  *Artificial Intelligence - From Science-Fiction to Real Use Cases* <span style="float:right">*11/2019*</span>

- **DataHack Summit 2019** <span style="float:right">Bangalore, INDIA</span>
  *Deep Learning for Aesthetics: Training a Machine to See What's Beautiful* <span style="float:right">*11/2019*</span>

- **AI Show Madrid**  
  *Demystifying the Buzz in Machine Learning! (This Time for Real)*  
  Madrid, ESP  
  *10/2019*

- **Bertelsmann Data Exchange Week**  
  *Keynote Talk: ML @ Axel Springer*  
  Bielefeld, GER  
  *10/2019*

- **VSNE Startup Idea Contest - The Grand Final Round**  
  *Keynote Talk and Panelist*  
  Berlin, GER  
  *09/2019*

- **I.E.C.T. Summer School 2019 -** `https://tinyurl.com/sfl58na`  
  *Keynote Talk and Panelist*  
  Innsbruck, AUS  
  *08/2019*

- **ODSC India 2019 -** `https://tinyurl.com/rrtfdo7`  
  *Image ATM - Image Classification for Everyone*  
  Bangalore, INDIA  
  *08/2019*

- **Chief Data Officer Exchange Europe 2019 -** `https://tinyurl.com/y3dnk9xl`  
  *Keynote Talk: Demystifying the Buzz in Machine Learning! (This Time for Real)*  
  Berlin, GER  
  *06/2019*

- **GCAAI 2nd Annual Meeting 2019 -** `https://tinyurl.com/y3oqk7y8`  
  *Axel Springer AI - the past, present and future*  
  Berlin, GER  
  *06/2019*

- **DataHack Radio (Analytics Vidhya) -** `https://tinyurl.com/y4rf5akx`  
  *Exploring Computer Vision and Data Engineering with Dat Tran*  
  Remote, INDIA  
  *04/2019*

- **hub.berlin 2019 / Big-Data.AI Summit 2019**  
  *Automate the Boring Stuff with Machine Learning*  
  Berlin, GER  
  *04/2019*

- **code.talks commerce 2019 -** `https://tinyurl.com/y4fvkjg8`  
  *AI Use Cases @ idealo*  
  Berlin, GER  
  *04/2019*

- **Data Festival 2019 / Panel Discussion**  
  *AI impact on P&L*  
  Munich, GER  
  *03/2019*

- **Data Natives 2018 -** `https://tinyurl.com/y8mgvpsv`  
  *Demystifying the Buzz in Machine Learning! (This Time for Real)*  
  Berlin, GER  
  *11/2018*

- **Axel Springer CTO Workshop 2018**  
  *AI Use Cases @ idealo*  
  Berlin, GER  
  *11/2018*

- **Süddeutscher Verlag UB Fachinformation**  
  *Keynote Talk: Smart Apps - The New Normal*  
  Munich, GER  
  *07/2018*

- **Chief Data Officer Exchange Europe 2018**  
  *Fireside Chat - Creating Profitable Data Opportunities*  
  Berlin, GER  
  *06/2018*

- **Data Festival 2018 -** `https://tinyurl.com/y7jrneej`  
  *What is the big in (Big) data? And how lean can your data be...*  
  Munich, GER  
  *04/2018*

- **Tech Open Air / Panel Discussion -** `http://tinyurl.com/y8vngang`  
  *New Age of Analytics for Digital Products?*  
  Berlin, GER  
  *07/2017*

- **Pivotal Big Data Roadshow -** `http://tinyurl.com/mnsgcce`  
  *Smart data-driven apps*  
  Seoul, KOR  
  *04/2017*

- **Cloud Foundry Europe Summit 2016 -** `https://youtu.be/z7AmakqpgdQ`  
  *Saving human lives with the IoT*  
  Frankfurt, GER  
  *09/2016*

- **PyData SF 2016 -** `https://youtu.be/HyjZFYDubf8`  
  *How you really get your data science models into production - the cool way!*  
  San Francisco, USA  
  *08/2016*

- **PyData Berlin 2016 -** `https://youtu.be/ZojIGRS3HLY`  
  *PySpark in Practice - Joint talk with Ronert Obst*  
  Berlin, GER  
  *05/2016*

- **PyData London 2016 -** `https://youtu.be/SETpipUZ_Lc`  
  *PySpark in Practice - Joint talk with Ronert Obst*  
  London, UK  
  *05/2016*

## Selected Open-Source Projects

- **imagededup** - `https://tinyurl.com/yyaoe587`  ★ 3000+  ⚇ 200+
  *Finding duplicate images made easy!*
- **Object Detector App** - `http://tinyurl.com/yyf7tmu9`  ★ 1100+  ⚇ 600+
  *Real-Time Object Recognition App with Tensorflow and OpenCV*
- **face2face Demo** - `http://tinyurl.com/y3u6u4cs`  ★ 1100+  ⚇ 350+
  *pix2pix demo that learns from facial landmarks and translates this into a face*
- **Raccoon Dataset** - `http://tinyurl.com/y29v4uhk`  ★ 1000+  ⚇ 700+
  *Trained a Raccoon detector with TensorFlow's Object Detection API*
- **jupyter2slides** - `http://tinyurl.com/y4qfenxv`  ★ 700+  ⚇ 150+
  *Cloud Native Presentation Slides with Jupyter Notebook + Reveal.js*

## Skills & Interests

- Computer

  - Very good knowledge in Python including the PyData stack: scikit-learn, pandas, matplotlib, XGBoost etc.
  - Working experiences with deep learning frameworks such as TensorFlow and PyTorch
  - Practical experiences in solving big data problems using Apache Spark (PySpark) and MPP databases
  - Worked under a Test-Driven and CI/CD environment and constantly embraced Pair Programming
  - Experiences with Microservices, PaaS and Cloud Computing (Cloud Foundry, AWS, GCP)
  - Good knowledge in SQL, Git, VIM, LaTeX, UNIX
  - Experiences with HTML, CSS and JS (mainly React)

- Industry

  - Industry experiences in Media, Telecommunications (Postpaid, LTE and VoIP), Aviation (Customer Segmentation), Automotive (Spare parts management, Aquaplaning prediction), Banking, Logistics and E-Commerce

- Languages

  - German (native), English (fluent), Vietnamese (beginner)

- Others

  - Taekwondo, Soccer (11 years Landesliga), Traveling, Literature, Cooking, History, Strategy Games
  - McKinsey Masters 2013 Regional Winner and second in the Nationals
  - e-fellows.net Scholarship

## References

**Dr. Michael Natusch**
Global Head of Artificial Intelligence - Prudential plc
m.natusch@me.com

**Holger Mai**
Managing Director - Accenture Analytics
holger.mai@accenture.com

Berlin, April 10, 2020