

# Justin A. Gould

---

Senior Data Scientist  
The Data Mine  
Purdue University  
West Lafayette, IN

[gould29@purdue.edu](mailto:gould29@purdue.edu)  
<https://www.linkedin.com/in/gouldju1/>  
<https://github.com/gouldju1>  
Phone: +1 (248) 877-0751

## Work Experience

12/2020- Present	Senior Data Scientist The Data Mine Purdue University, West Lafayette
04/2020- 12/2020	Product Owner, NLP Services and Connected Vehicles Ford Motor Company AI Advancement Center - Natural Language Processing Research Smart Mobility Analytics - Connected Vehicles Products and Platforms
11/2019- 09/2020	Machine Learning & Artificial Intelligence Scientist Ford Motor Company AI Advancement Center - Natural Language Processing Research
01/2019- 11/2019	Data Scientist Ford Motor Company Smart Mobility Analytics - Driver Assistance Technology
08/2018- 12/2018	Data Science Consultant (M.S. Experiential Learning) Stryker Medical Devices Orthopaedic Instruments
06/2018- 08/2018	Data Science Intern Ford Motor Company Product and Sustainability Analytics
01/2018- 05/2018	Data Science Consultant (M.S. Experiential Learning) Meijer, Inc. Marketing - Advanced Analytics
06/2016- 12/2018	Data Science Intern DTE Energy Electric Distribution Operations
06/2013- 08/2016	Occupant Safety Prototype and Testing Intern ZF Friedrichshafen AG Occupant Safety Systems and Driver Restraints, North America

## Education

- 2018 M.S. Business Analytics  
Broad Graduate School of Management  
Michigan State University
- 2017 B.A. WITH HONORS in English Literature  
College of Arts and Letters  
Michigan State University

## Service to the Profession

### *Current*

**Board Member**, MSU Broad Graduate School of Management M.S. Business Data Science & Analytics Advisory Board

### *Former*

**Organizer**, Ford Motor Company AI Paper Review Series  
**Planning Committee**, Ford Motor Company R Users Guild  
**Planning Committee**, Ford Motor Company Python Users Guild  
**Intern Recruiting Committee**, Ford Motor Company

## Speaking Engagements

### PRESENTATIONS

- Sep. 2021 **Guest Speaker**, “Dangerous A.I.: A Primer on Algorithmic Bias,” Data Science Topics Seminar Series, Broad Graduate School of Management, Michigan State University, East Lansing, MI.
- Apr. 2021 **Technical Expert**, “Leveraging Named Entity Recognition and Named Entity Linking of Service Text Data for Prognostics and Causal Part Attribution,” Cummins, Inc., Columbus, IN.
- Mar. 2021 **Keynote Speaker**, “Is Data Science Still ‘Sexy’? A Recent History and Outlook of Analytics in Industry,” Data Analytics and Software Event, Krannert School of Management, Purdue University, West Lafayette, IN.
- Feb. 2021 **Presenter**, “Connected Vehicle Data Architecture for Large-scale Analytics,” The Data Mine, Purdue University, West Lafayette, IN.
- Jan. 2021 **Guest Speaker**, “Leveraging Flask for Microservice Development and Deployment,” Cummins, Inc., Columbus, IN.
- Dec. 2020 **Technical Expert**, “Applications of NLP for Manufacturing Service-oriented Companies,” Cummins, Inc., Columbus, IN.

- Aug. 2020 **Presenter**, “Deploying NLP Model as a Service using Kubernetes,” Global Data Insights & Analytics, Ford Motor Company, Dearborn, MI.
- Apr. 2020 **Presenter**, “Addressing Structural Ambiguity in Automated Tabular Data Extraction,” AI Advancement Center, Ford Motor Company, Dearborn, MI.
- Jan. 2020- Apr. 2020 **Organizer and Presenter**, “Automotive Industry and Automobile Major Components Lecture Series,” AI Advancement Center, Ford Motor Company, Dearborn, MI.
- Feb. 2020 **Presenter**, “English is the New SQL: Translating Natural Language Questions to SQL,” AI Advancement Center, Ford Motor Company, Dearborn, MI.  
*Selected as 1 of 10 talks to give at 2020 GDI&A Conference ( > 1,000 attendees); cancelled due to COVID-19 pandemic.*
- Nov. 2019 **Presenter**, "Information Retrieval for Question & Answering," The Data Mine, Purdue University, West Lafayette, IN.
- Sep. 2019 **Presnter**, “Creating Effective Recommendation Engines using R,” Ford R Guild Knowledge Share, Ford Motor Company, Dearborn, MI.

#### POSTERS

- Dec. 2019 “Hidden Markov Models for Autonomous Vehicle Map Matching," 2019 Ford Global Control Conference, Ford Motor Company, Dearborn, MI.
- Jun. 2019 “Connected Blue Zones: Crowdsourced Maps for Assisted Driving,” 2019 Global Data Insights & Analytics Conference, Ford Motor Company, Livonia, MI.

#### PANELS

- May 2020 **Panelist**, “AI/ML Experts and Research Panel,” Ford Product Development and Advanced Engineering, Ford Motor Company, Dearborn, MI.

## Publications

#### PAPERS, JOURNAL ARTICLES, & TECHNICAL REPORTS

- 2021 **Gould, J.** A Framework for Auditing Data Center Energy Usage and Mitigating Environmental Footprint. *arXiv*, Computer Science | Distributed, Parallel, and Cluster Computing. <https://arxiv.org/abs/2102.04446>

#### IN PROGRESS

- 2021 **Gould, J. and Roach, J.** Predicting Short-term Cryptocurrency Volatility using Twitter, Sentiment Analysis, and Web Search Data.
- Nov. 2021

**Presenter**, “English is the New SQL: Translating Natural Language Questions to SQL,” 2021 Academic Data Science Alliance Annual Meeting, Savannah, GA.

## Teaching

- Fall 2021 **Purdue University**  
HONR 39900: Special Topics in Data Science: Foundations of Geospatial Analytics
- Summer 2021 **Purdue University**  
RESEARCH AND EXTENSION EXPERIENTIAL LEARNING FOR UNDERGRADUATES:  
Leveraging Python and SQL for Agricultural Geospatial Analytics

## Research

- 08/2020-  
05/2021 **Purdue University & Ford Motor Company**  
***WHERE Support for Natural Language Interface to Database Models***  
This research project is an alliance between Ford Motor Company and Purdue University’s Data Mine. Current approaches to the Natural Language Interface to Database (NLIDB) problem (an extension of Natural Language to SQL for relational databases) ignores WHERE clause prediction, making them unusable in industry. We aim to close this gap by developing a WHERE clause predictor to implement in models on Yale’s Spider (Semantic Parsing and Text-to-SQL Challenge) task.
- 08/2019-  
05/2021 **Purdue University & Ford Motor Company**  
***In-Vehicle, Multi-Turn, Open-Domain Question-Answer Systems***  
This research project is an alliance between Ford Motor Company and Purdue University’s Data Mine. We created an in-vehicle Question-Answer system utilizing the 2020MY Lincoln Aviator manual. Given a question about a user’s vehicle, which can be answered from the manual, the system searches the manual and returns the most probable answer, via ALBERT for text-retrieval and BERT- based Natural Language to SQL (NL2SQL) model for answers stored in tables.
- 04/2019-  
12/2019 **University of Michigan & Ford Motor Company**  
***Smart Vehicles for a Smart World***  
The Ford initiative with the University of Michigan’s Transportation Research Institute (UMTRI) utilizes high-fidelity data collection technology from vehicles and intersections around Ann Arbor to collect > 2 TB of data every month to find use cases in which Ford can use the vehicle as a sensor in various mobility products, such as finding correlations between driving behaviors, quantifying driving behavior, and classifying drivers into aggression segments for Usage-Based Insurance.

## Skills, Tools, etc.

### LANGUAGES

English (native), German (advanced)

### PROGRAMMING

Python, R, Scala, SQL, CSS, HTML, Java, JavaScript, SAS, L<sup>A</sup>T<sub>E</sub>X

### TOOLS/SOFTWARE

**Big Data:** Hadoop, Spark, PySpark

**Data Visualization:** Tableau, PowerBI, Qlik

**ML Framework:** TensorFlow, PyTorch, Keras, Azure Cognitive Services

**Geospatial:** ArcGIS, Kinetica, GeoSpark, PostGIS, OSM

**Data Management:** Maximo, Alteryx, Postgres, MongoDB, SQLAlchemy, Azure

**Data Workbench:** Domino Data Labs, RAPIDS, Databricks

**Orchestration and Deployment:** Kubernetes, Django, Flask

### CERTIFICATIONS

**Agile Project Management**, Purdue University Global, Issued May 2021

### PROFESSIONAL SOCIETIES

**American Statistical Association (ASA)**, Member

**ASA Text Analytics Interest Group**, Member