

David S. Hayden

🌐 www.dshayden.com — ✉ [david at dshayden dot com](mailto:david@dshayden.com) — [🌐 LinkedIn](#) — [🎓 Google Scholar](#) — 📍 Palo Alto, CA

Summary I work on robustness, lead world modeling and Perception AI Research at Cruise, and publish in top venues.

Research Diffusion, world models, foundation models, robustness, synthetic data, uncertainty quantification, longtails. MIT GenAI PhD focused on uncertainty, Bayesian nonparametrics, tracking, distributions on manifolds, and behavior analysis.

Experience

Cruise, Staff Research Scientist

Nov 2021 – Current

- Shipped impact to 7 on-road models (5 behaviors, 2 perception) running on 400+ driverless cars in urban environments.
- Led Cruise’s world modeling projects and Perception AI Research (3 ICs).
- Built Cruise’s first world model. Trained on 1B+ internal data. Enabled multi-camera video generation with diverse conditioning (HD maps, 3D trajectories, fine-grained object properties, camera extrinsics, text, reference frames). Outperformed Cosmos. Secured buy-in from Production for synthetic training data and from Sim for closed-loop eval.
- Pioneered diffusion at Cruise by shipping millions of longtail synthetic training data to on-road models, yielding double-digit generalization improvements, resolving paretos, and enabling maneuvers like temporary traffic redirections.
- Developed novel diffusion guidance techniques for generating realistic synthetic data that are longtail from the perspective of existing production models. Enabled a new longtail discovery paradigm of expending offline compute to generatively mine longtail synthetic data and textually describe+summarize it relative to existing real data with VLMs.
- Developed unsupervised diffusion alignment techniques that enabled diffusion models to generate content observed only from external/Internet data, but distributionally-aligned to internal data.
- Architected the decoder for DriveGPT (Cruise’s flagship foundation driving model). Improved sampling diversity that scaled with compute, enabling the first commercial deployment of an autoregressive transformer as an AV planner.
- Developed and deployed real-time on-road ‘known unknown’ signals, enabling increased safety in longtail scenarios.
- Proposed, secured buy-in for, and managed a \$6m multi-tenant infrastructure overhaul that was responsible for \$10m+/year increased, company-wide developer velocity. Unblocked numerous foundation model workstreams, powered more than six publications, and resulted in promotions for 3 ICs.
- Published three papers in ICML ‘25, CVPR ‘25 on robustness, diffusion guidance and synthetic data [1] (first author), foundation driving models [2] (supporting), and causal reasoning in diffusion for closed-loop simulation [3] (advisory).

Augmodo, Technical Advisor

Jun 2024 - Current

- Led foundation model strategy and development for Spatial AI in retail based on passive wearable camera data, enabling real-time, fine-grained visual product identification, in-store localization, and out-of-stock lookups for catalogs of 100k-1m products without barcode reading, environment sensors, or inventory system interop.
- Designed a complete computer vision stack rewrite away from SLAM and towards embeddings, enabling a rapid 12x operations scale-up, 10x reduced human-in-the-loop costs, and \$38m Series-A.
- Built novel ‘jigsaw puzzle transformers’ that localize out-of-stock products with 99% coverage.
- Designed data scale-up to 10m+, directly secured seed and Series-A investments, and set up multi-node training.

Teton Ridge, Companion (Consultant)

2021 - 2023

- Teton Ridge: Developed and deployed from scratch real-time, camera-based, multi-object tracking + performance metrics of championship barrel racers in unseen environments. The system ran live at the American Rodeo with zero failures, was visualized on the jumbotron to an audience of more than 40,000, and aired on television (CBS Sports).
- Companion: Developed pose tracking for canines, enabling a \$1m grant from Elanco. Joined Board of Advisors.

Essistive, Founder

Jan 2012 - Jun 2015

- Commercialized my undergraduate Note-Taker work, bringing it to hundreds of customers before licensing it to Perkins.

Internships (Google, NASA JPL)

Q3 2008 - 2011, 2015

- Google ‘15 (Novel View Synthesis) with Steve Seitz: Adapted Google Street View deep nets for VR-focused stereo panorama generation. Trained from a 16-camera stereo rig and outperformed classical stereo vision pipelines.
- Google ‘11 (Coupling Visual Search and Language Models for Image Retrieval) with Hartmut Neven, Hartwig Adam: Integrated ad-based language models with visual search for improved image similarity metrics in Google Goggles.
- NASA JPL ‘08-‘11 (Metric Learning for Improved Science Return) with Steve Chien: Learned distance metrics for selective data return in space that optimally satisfied multiple, competing mission objectives by distinct science teams. Published in [9, 11, 19, 17].

Education

PhD Computer Science, Massachusetts Institute of Technology (MIT)

Spring 2021

Thesis: Uncertainty Quantification and Structure Discovery for Scalable Behavior Science

Adviser: John Fisher; Committee: Robert Desimone, Justin Solomon

- Conducted long-term, custom video and depth tracking (hundreds of hours) of genetically modified primates, resulting in the first evidence for primate animal models in autism research. Enabled findings by observing statistical differences in behavior between pairs of primates with and without the SHANK-3 gene. Published in Nature [8].
- Developed unsupervised learning of an articulated parts model by observing short video, depth, point cloud, or mesh sequences of objects in motion, assuming that body and part dynamics were described by distributions on the Lie group of rigid-body transformations. Designed multi-object tracking with explicit representation of data association uncertainty and exact posterior sampling that reasoned over latent permutations. Enabled automatic identification of ambiguous outcomes, supporting sparse, algorithmic, or human-in-the-loop corrections for rapid improvement of track quality. Published in CVPR and NeurIPS [7, 5, 6].

MS Computer Science, Massachusetts Institute of Technology (MIT)

Spring 2014

Thesis: Wearable-Assisted Social Interaction as Assistive Technology for the Blind

Adviser: Seth Teller

- Designed a wearable jacket platform with discreet camera, audio, and motion sensors that recognized proximate acquaintances, enabling social interaction analysis as well as serendipitous social interactions for the blind. Conveyed information to the user through private Morse-coded vibrations or audio. Supported by and demoed to Andrea Bocelli.
- Published in CVPR and CHI workshops [15, 16].

BS Computer Science, BS Mathematics, Arizona State University (ASU)

Spring 2011

Thesis: Note-Taker: Enabling the Legally Blind to Take Notes in Class

Adviser: Sethuraman Panchanathan

- Developed a custom pan/tilt/zoom camera and cross-platform (Windows, OSX, Linux) note-taking software with visual servoing, image enhancement, and synchronized video/handwritten notes. Dramatically reduced board-note-board delay for individuals who were legally blind, helping them keep pace in business presentations and classroom lectures.
- Won international competitions and received publicity (NPR, Wired, documentary, white house presentation). Secured \$0.5m NSF funding, co-led a team of CS, EE, MechE, and design engineers, and published in [10, 12, 13, 14, 18].

Publications

13 conference/journal, 5 workshop, 1 book chapter. Venues include Nature, Neurips, ICML, CVPR, ASSETS and CHI.

Conference/Journal

1. ICML 2025 Generative Data Mining with Longtail-Guided Diffusion.

David S. Hayden, Mao Ye, Timur Garipov, Gregory P. Meyer, Carl Vondrick, Zhao Chen, Yuning Chai, Eric Wolff, Siddhartha S. Srinivasa.

2. ICML 2025 DriveGPT: Scaling Autoregressive Behavior Models for Driving.

Xin Huang, Eric M. Wolff, Paul Vernaza, Tung Phan-Minh, Hongge Chen, David S. Hayden, et al..

3. CVPR 2025 Causal Composition Diffusion Model for Closed-loop Traffic Generation.

Haohong Lin, Xin Huang, Tung Phan-Minh, David S. Hayden et al..

4. ChemRxiv 2025 DeepDeg: Forecasting and explaining degradation in novel photovoltaics.

David S. Hayden, Mao Ye, Timur Garipov, Gregory P. Meyer, Carl Vondrick, Zhao Chen, Yuning Chai, Eric Wolff, Siddhartha S. Srinivasa.

5. Neurips 2020 Sequential Bayesian Experimental Design with Variable Cost Structure.

Sue Zheng, David S. Hayden, Jason Pacheco, John W. Fisher III.

6. arXiv 2020 Efficient Data Association and Uncertainty Quantification for Multi-Object Tracking.

David S. Hayden, Sue Zheng, John W. Fisher III.

7. CVPR 2020 Nonparametric Object and Parts Modeling with Lie Group Dynamics .

David S. Hayden, Jason Pacheco, John W. Fisher III.

8. Nature June, 2019 Atypical Behaviour and Connectivity in SHANK3-Mutant Macaques.

Yang Zhou, Jitendra Sharma, Qiong Ke, Rogier Landman, Jingli Yuan, Hong Chen, David S. Hayden et al..

9. ACM TIST 2012 Using Clustering and Metric Learning to Improve Science Return of Remote Sensed Imagery.

David S. Hayden, Steve Chien, David R. Thompson, Rebecca Castaño.

10. ACM ASSETS 2011 Note-Taker 3.0: an Assistive Technology Enabling Students who are Legally Blind to Take Notes in Class.

David S. Hayden, Michael Astrauskas, Qian Yan, Liqing Zhou, John A. Black Jr..

11. IEEE Intelligent Systems 2012 Using Onboard Clustering to Summarize Remotely Sensed Imagery.

David S. Hayden, Steve Chien, David R. Thompson, Rebecca Castaño.

12. ACM ASSETS 2010 Note-Taker 2.0: the Next Step Toward Enabling Students who are Legally Blind to Take Notes in Class .

David S. Hayden, Liqing Zhou, Michael Astrauskas, John A. Black Jr..

13. ACM ASSETS 2008 Note-Taker: Enabling Students who are Legally Blind to Take Notes in Class .

David S. Hayden, Dirk Colbry, John A. Black Jr., Sethuraman Panchanathan.

Book Chapters

14. Impact of Tablet PCs and Pen-based Technology on Education, ISBN: 1557535744 2010 The Note-Taker: a Tablet PC Based Device that Helps Students Take and Review Classroom Notes.

David S. Hayden, Liqing Zhou, John A. Black Jr..

Peer-Reviewed Workshop Publications

15. CHI Workshop on Assistive Augmentation 2014 Unobtrusive, Wearable Social Interaction Detection and Assistance.

David S. Hayden, Robert C. Miller, Seth Teller.

16. CVPR Workshop on Egocentric Vision 2012 The Accuracy-Obtrusiveness Tradeoff for Wearable Vision Platforms.

David S. Hayden, Carl Vondrick, Stella Jia, Yafim Landa, Robert C. Miller, Antonio Torralba, Seth Teller.

17. 10th International Symposium on Artificial Intelligence, Robotics, and Automation in Space 2010 Onboard clustering of aerial data for selective data return.

David S. Hayden, Steve Chien, David R. Thompson, Rebecca Castano.

18. CVPR Workshop: Computer Vision Applications for the Visual Impaired 2010 The Note-Taker: An assistive technology that allows students who are legally blind to take notes in the classroom.

John A. Black Jr., David S. Hayden.

19. IJCAI Workshop on AI in Space 2009 Onboard clustering of aerial data for improved science return.

David S. Hayden, Steve Chien, David R. Thompson, Rebecca Castano.

Academic Awards

Contributed to or directly earned over \$1.2 million in grants, prizes and awards.

2019 McGovern Institute Neurotechnology (MINT) Grant* (\$50k)

2018 NIH Grant 1R01MH111916-01A1* (\$500k)

2018 McGovern Institute Neurotechnology (MINT) Grant* (\$50k)

2016 NSF Graduate Research Fellowship (\$141k)

2014 Google Glass Accessibility Award (\$20k)

2011 IDSA Design Award (Silver)

2011 Microsoft Imagine Cup Software Design, 2nd Place Worldwide (\$8k)

2011 Microsoft Imagine Cup Software Design, 1st Place Nationally

2010 Microsoft Imagine Cup Accessibility, 1st Place Worldwide (\$8k)

2009 NSF Grant 0931278 (2009–2013)** (\$500k)

2009 Additional (Google Lime, NASA MUST) (\$15k)

* Emerged as part of a collaboration with Rogier Landman.

** Emerged from a project I started in Sethuraman Panchanathan's lab.

Talks

Gave 22 academic talks at conferences, universities, and competitions.

- **Generative Data Mining, Longtail Guidance and Discovery** (2025): University of British Columbia
- **Tracking, Parts Modeling, Behavior Analysis** (2020): Harvard, Broad Institute, Massachusetts Institute of Technology, Georgia Tech, Max Plank Institute for Intelligent Systems, Johns Hopkins University, National University of Singapore, Stanford University, CVPR Oral
- **Wearable Social Interaction Assistance and Analysis** (2012 - 2014): CHI Workshop (2014), Andrea Bocelli Foundation Workshop (2013), Andrea Bocelli Foundation Workshop (2012), CVPR Workshop (2012)
- **Note-Taker** (2008-2011): Microsoft Imagine Cup World Finals, 2nd Place, Software Design (2011), Microsoft Imagine Cup National Finals, 1st Place, Software Design (2011), ASSETS Oral (2011), Microsoft Imagine Cup World Finals, 1st Place, Accessibility (2010), ASSETS Oral (2010), ASSETS Oral (2008)
- **Artificial Intelligence for Improved Science Return in Space** (2009 - 2010): i-SAIRAS (2010), IJCAI Workshop (2009)

Mentorship and Teaching

Advised/led 16 students (1 PhD, 7 MS, 8 undergrad). Cathy Wu is MIT faculty. COACH+ received VC funding. Timur is at OpenAI.

1. Timur Garipov *Attention-Based Uncertainty Gating for Transformer MoE (2023)*
2. Zion Hadley *COACH+ Personalized Sports Analytics from Pose Estimation (2020)*
3. Habeeb Salau, *COACH+ Personalized Sports Analytics from Pose Estimation (2020)*
4. Terryn D. Brunelle, *Object Detection for a Robotic Science Laboratory (2020)*
5. Fernando Herrera, *Vision-Driven Interventions in a Robotic Science Laboratory (2020)*
6. Yafim Landa, *Low-Latency Wearable Systems Integration (2013)*
7. Cathy Wu, *User Interaction with a Wearable Vision Assistant (2013)*
8. Sophie L. Diehl, *Sewing Patterns for a Wearable Vision Assistant (2013)*
9. Mike Rush, *Hardware Design of a Custom Pan/Tilt/Zoom Vision Assistance Camera (2011)*
10. Michael Fruchtman, *Aligning Video with Pen Strokes (2011)*
11. Qian Yan, *3D-Printed Design for a Fourth-Generation Note-Taker (2011)*
12. Michael Astrauskas, *Visual Servoing in a Vision Assistance Camera (2009 - 2011)*
13. Liqing Zhou, *3D-Printed Design for a Third-Generation Note-Taker (2010)*
14. Parth Pandya, *Notes Archiving for the Note-Taker (2010)*
15. Andrew Kelley, *Software Design for Aligned Video and Pen Notes (2009)*
16. Shashank Srinivas, *Image Enhancement for the Note-Taker (2009)*

Selected Service

- 2024 **RSS Autonomous Vehicles Across Scales Workshop Organizer**
- 2020 **Assistive Technology (ATHack) Hackathon Organizer**
- 2013 **Andrea Bocelli Foundation Workshop Organizer**
- 2013 **MIT Diversity Summit Panelist**
- 2012 **White House Science Fair with President Obama (Presenter)**

Selected Press

- 2015 Slate
- 2012 Blind Ambition (Documentary)
- 2011 Boston Magazine
- 2011 NPR Science Friday
- 2011 Wired
- 2011 GeekWire
- 2011 Independent

References

Available upon request.