2021 embedded VISION summit

VIRTUAL | MAY 25-28

Optimizing ML Systems for Real-World Deployment

Danielle Dean, Ph.D.

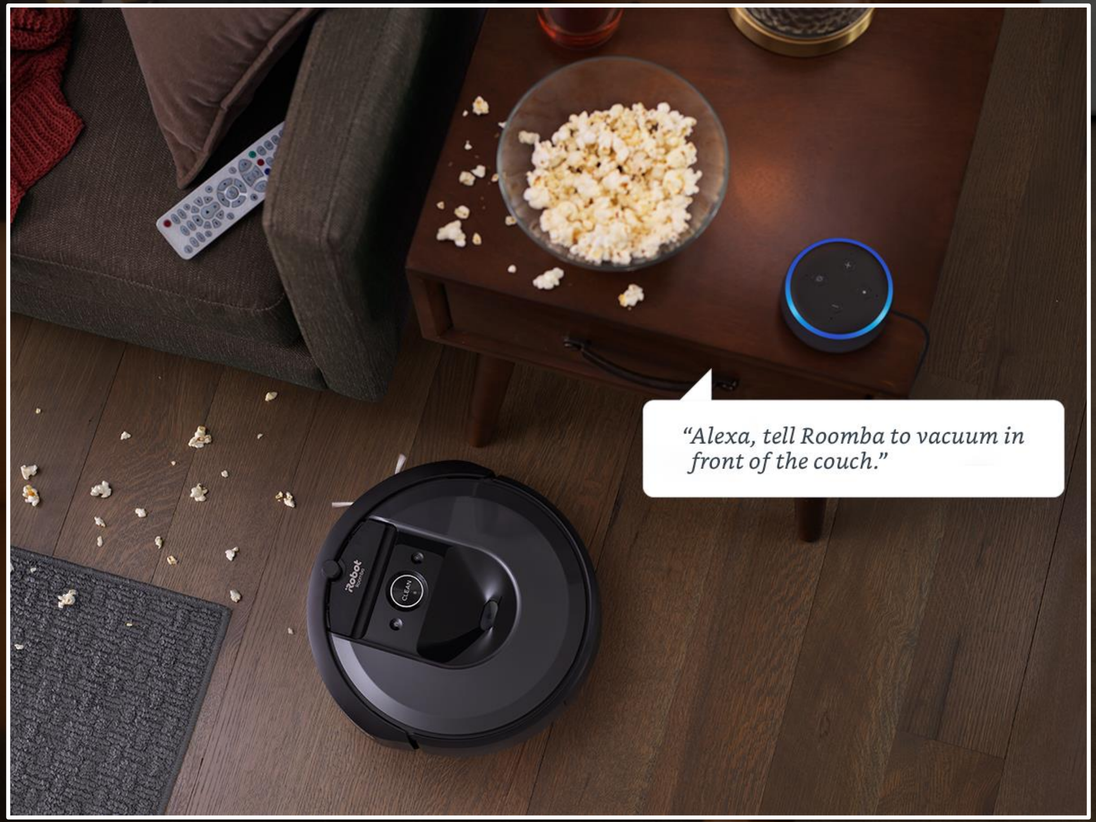
Technical Director of Machine Learning, iRobot

iRobot


Analysis

Agenda

- Brief introduction of ML @ iRobot and problems we're tackling
- Step 1 – Clear System Goals
- Step 2 – Data Collection and Processing
- Step 3 – Testing and Reproducibility

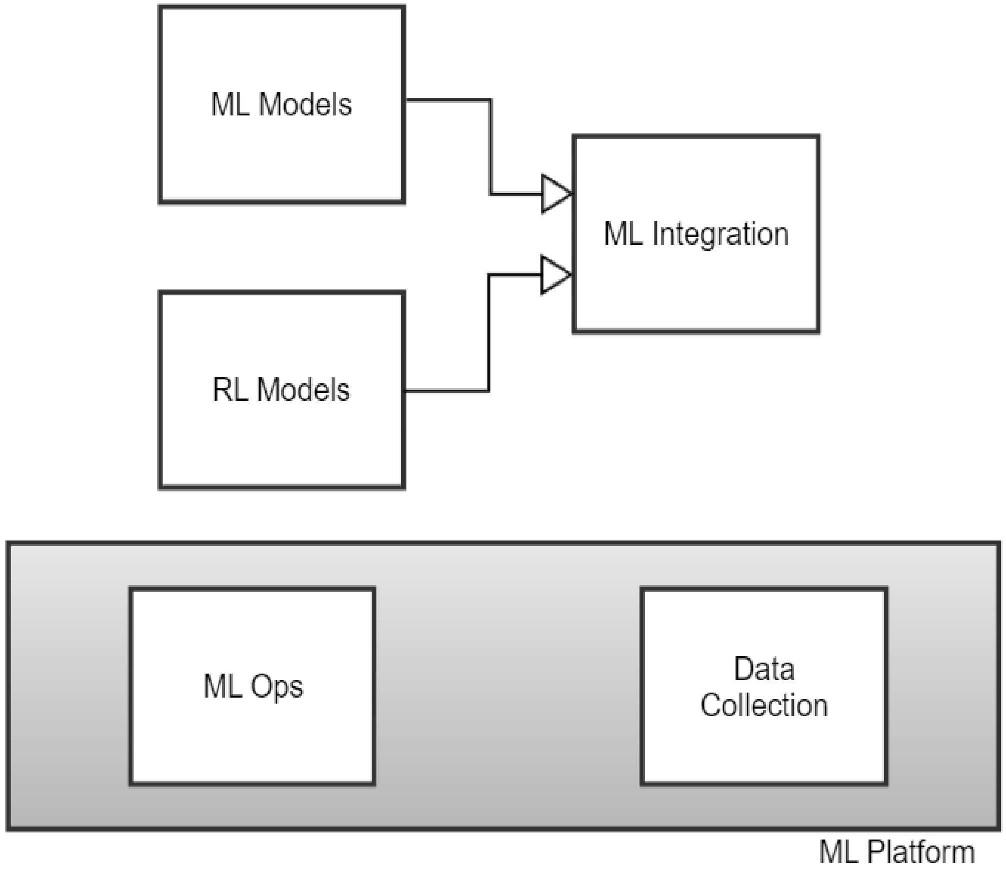


iRobot Genius™

Home Intelligence

ML @ iRobot

- Over ***10 million*** connected robots in production
- Mainly focused on computer vision & reinforcement learning



Step 1: Create Clear System Goals

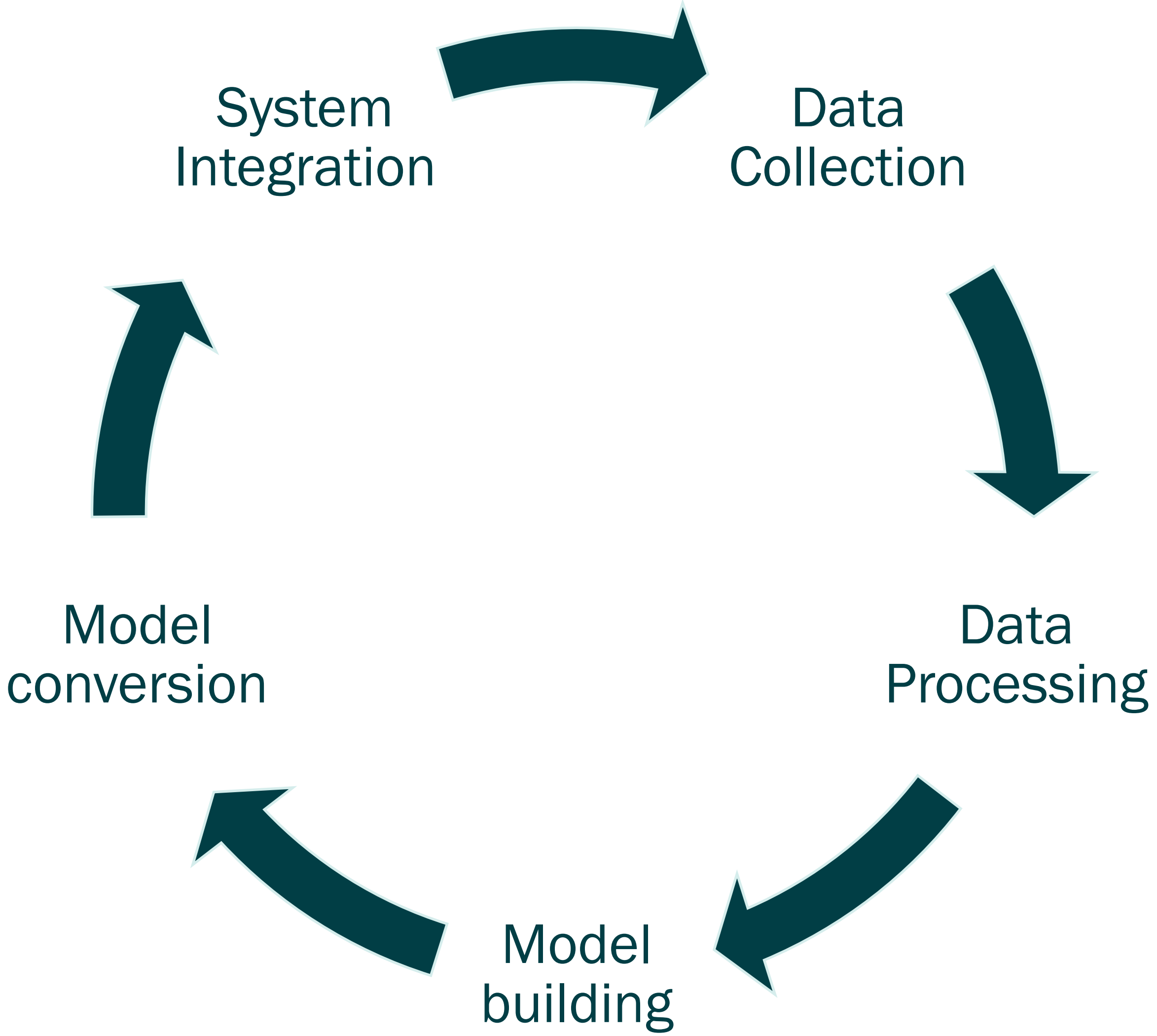
Do you have system KPIs that help drive end to end development?

- Proving that it works is not as simple as it seems
- Measuring success isn't simply comparing against what it is replacing or improving upon

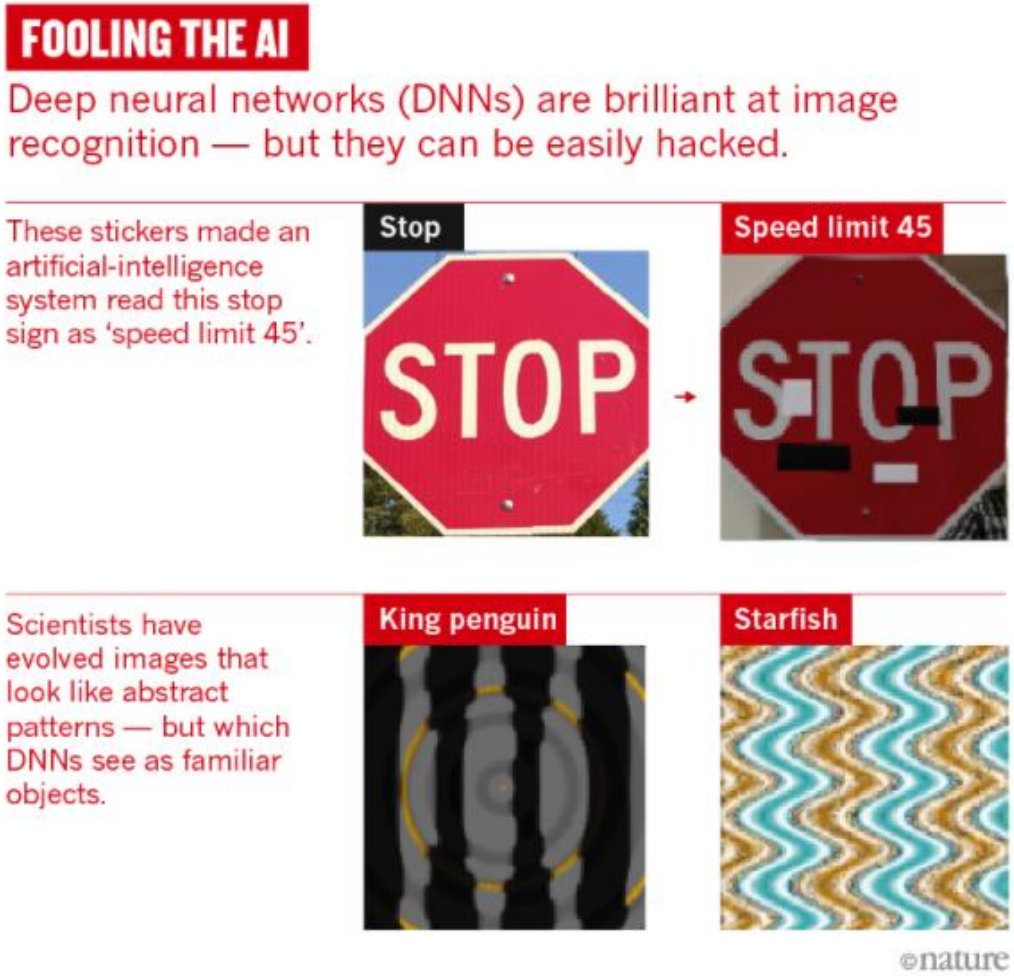
What form should the system even take?

- Does the automation need to take the same form as the thing it is automating?
- What other information can be leveraged to improve the product?

How can you build the system to enable learning where you need to optimize?



How does the system handle ML model failures?



Sources: Stop sign: Ref. 1; Penguin: Ref. 5

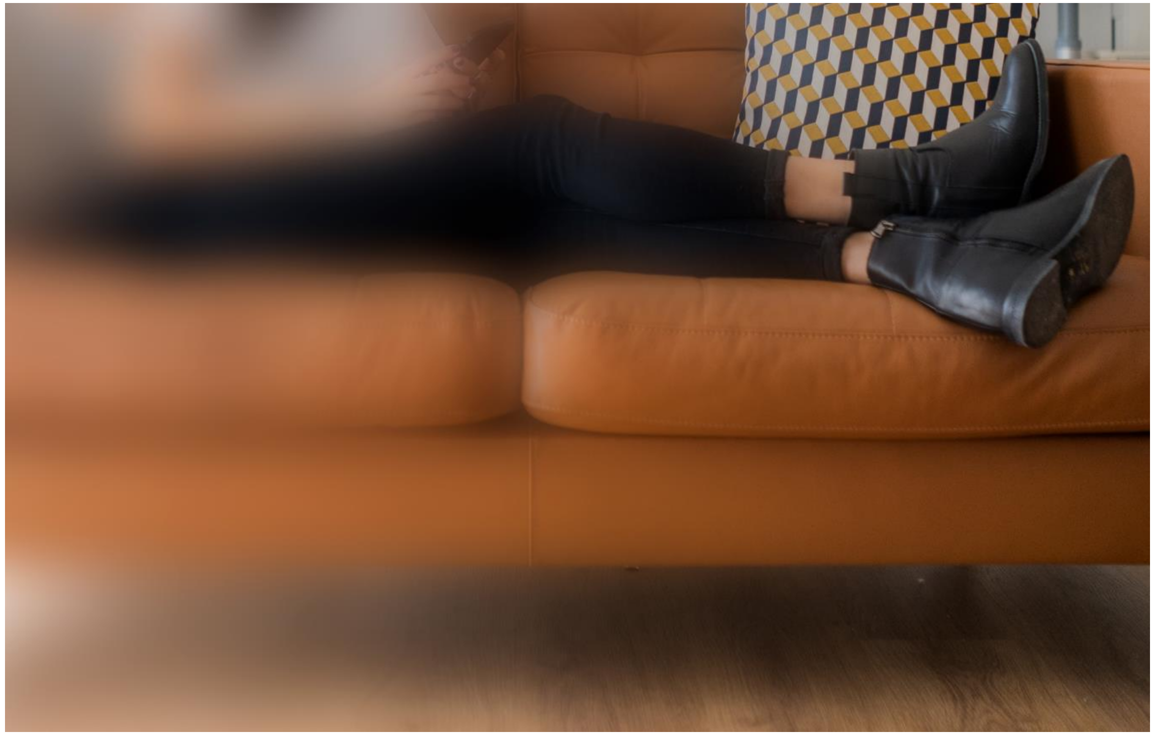
- What other systems can complement?
- How can users be brought into the experience when needed?

Step 2: Build Privacy-conscious Data Collection Processes

Is there data already available you can utilize?

- Consider available data points that can serve as “ground truth” for the machine learning system
- Consider what can be done in simulation or with synthetic data
- Should come without saying to focus on privacy-conscious collection and development processes

How well does your training and validation data capture the real-world deployment environment?



- Be careful not to “over-optimize” data collection procedures which may not represent real-world deployment environment
- Hardware, software, camera parameters, environment…

How does the data drift in production?

- Camera changes
- Hardware wear and tear

arXiv.org > cs > arXiv:3141.59265

Search... All fields Search

Help | Advanced Search

Computer Science > Machine Learning

Surpassing the state of the art on ImageNet by collecting more labels

(Submitted on 2020)

We achieve state-of-the-art 99.5% top-1 accuracy on ImageNet using a ResNet-50. This was achieved by paying people money to clean and grow the training set. First we cleaned up the incorrect labels in the existing training set and validation set. Then we found more unlabeled images similar to the high loss images in the validation set, labeled them, and added them to the training set. We repeated this process until accuracy improved enough. Data for this paper will be made available.

Subjects: **Machine Learning (cs.LG)**; Computer Vision and Pattern Recognition (cs.CV); Machine Learning (stat.ML)

Download:

- PDF
- Other formats

(license)

Current browse context:

cs.LG

< prev | next >

new | recent | 2002

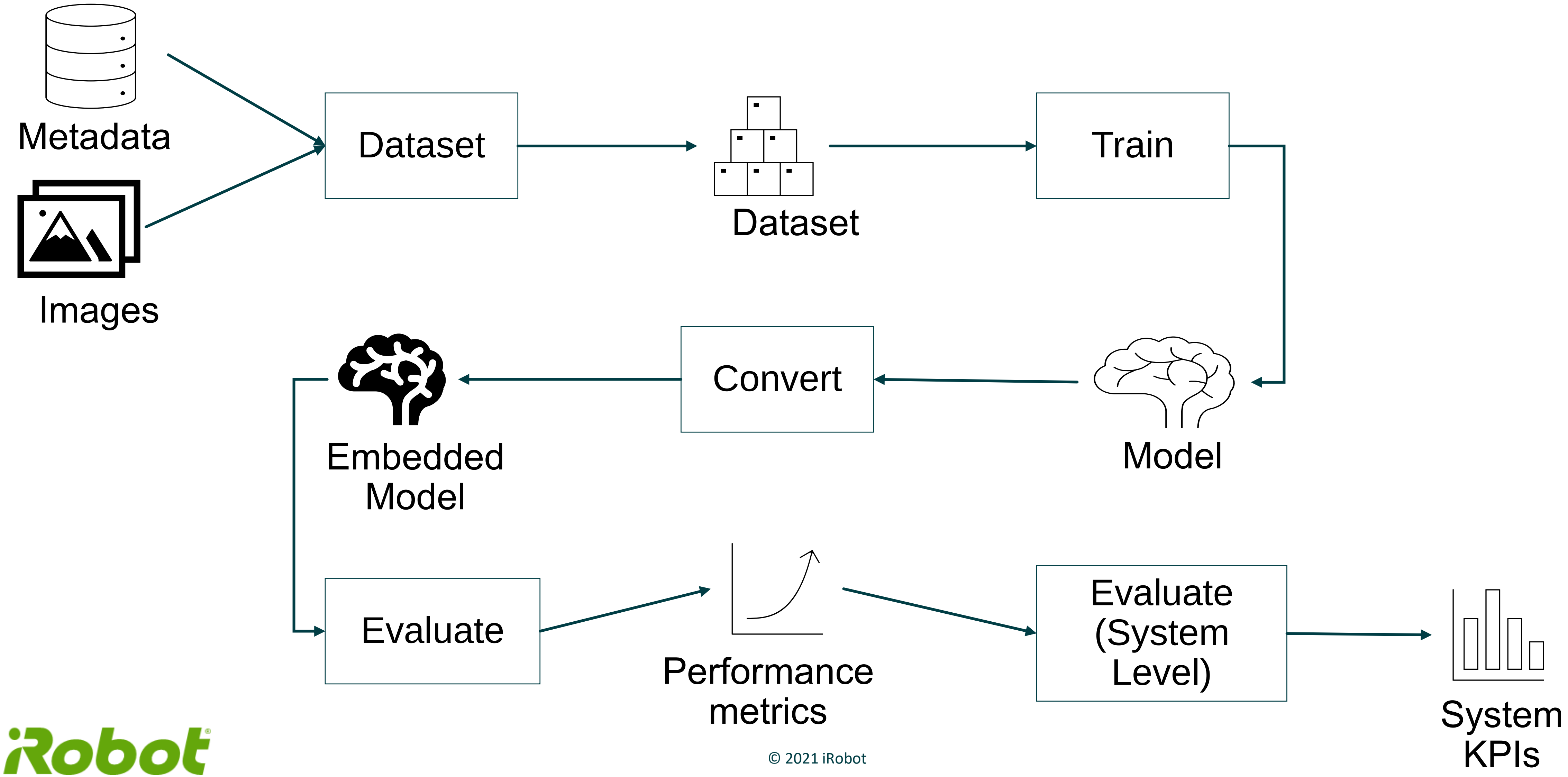
Change to browse by:

cs

https://twitter.com/karpathy/status/1231378194948706306?s=20

Step 3: Testing and Automation

End to End Model Development



Are you producing pipelines rather than models?

- Reproducibility
- Version Control
- Testability
- Integration with rest of system

Questions to consider

- Step 1 – Create clear system goals
  - Do you have system KPIs that help drive end to end development?
  - How can you build the system to enable learning where you need to optimize?
  - How does the system handle ML model failures?
- Step 2 – Data Collection and Processing
  - Is there data already available you can utilize?
  - How well does your training and validation data capture the real-world deployment environment?
  - How does the data drift in production?
- Step 3 – Testing and Reproducibility
  - Are you producing pipelines rather than models?
  - How do you plan to tackle the “long tail” distribution of real-world edge cases?

Resources

More details about ML@ iRobot

iRobot Genius
https://www.irobot.com/About-iRobot/iRobot-Genius.aspx

AWS re:Invent 2020: Running machine learning workflows at enterprise scale using Kubeflow
https://www.youtube.com/watch?v=X3772hxWstI

Join us!
https://www.irobot.com/about-irobot/careers