Data Science Workstations

The Missing Link to Productivity?

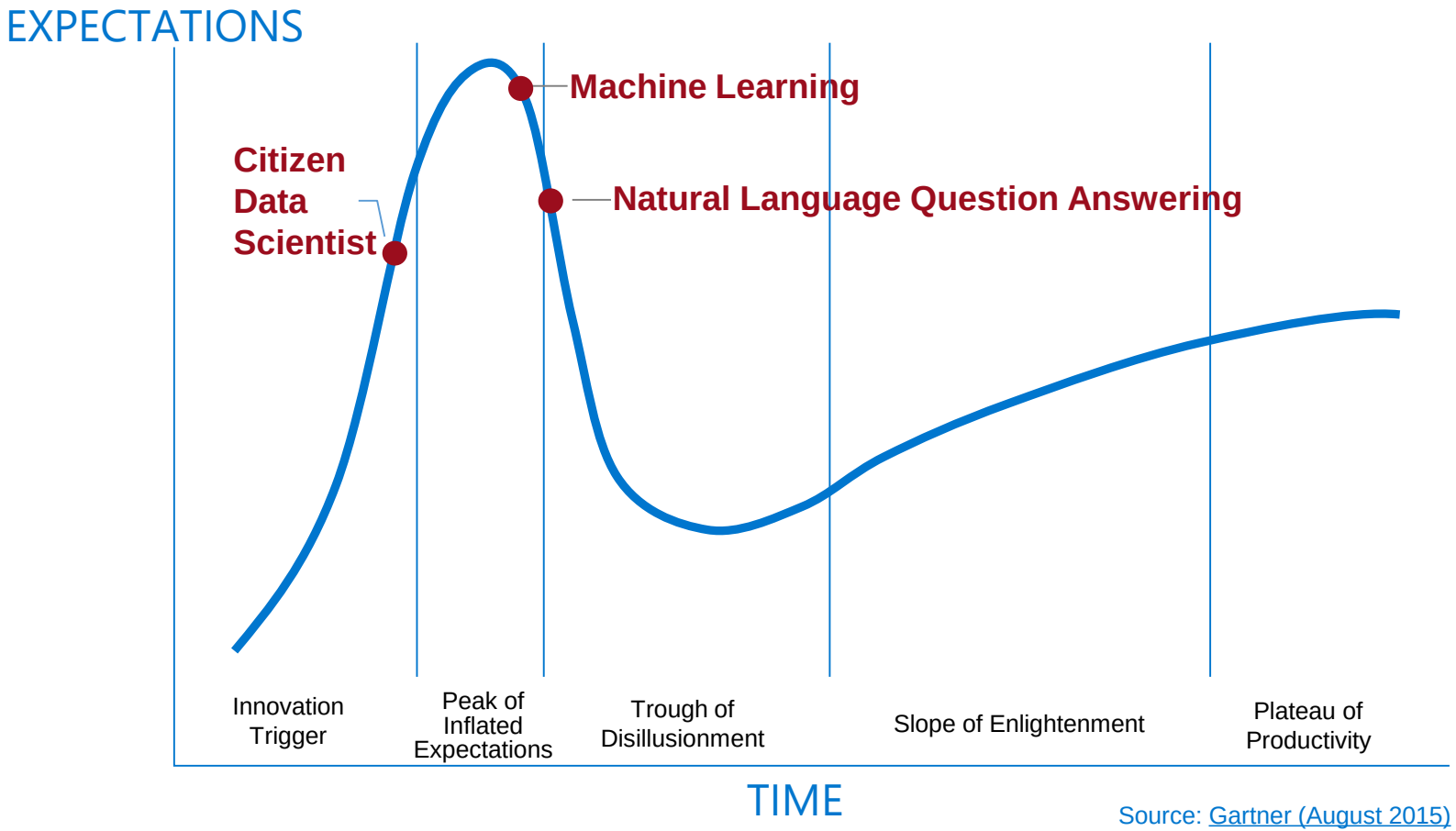
David Patschke, Dell
AI/ML Strategy, Dell Precision Workstations

S9996

Introduction

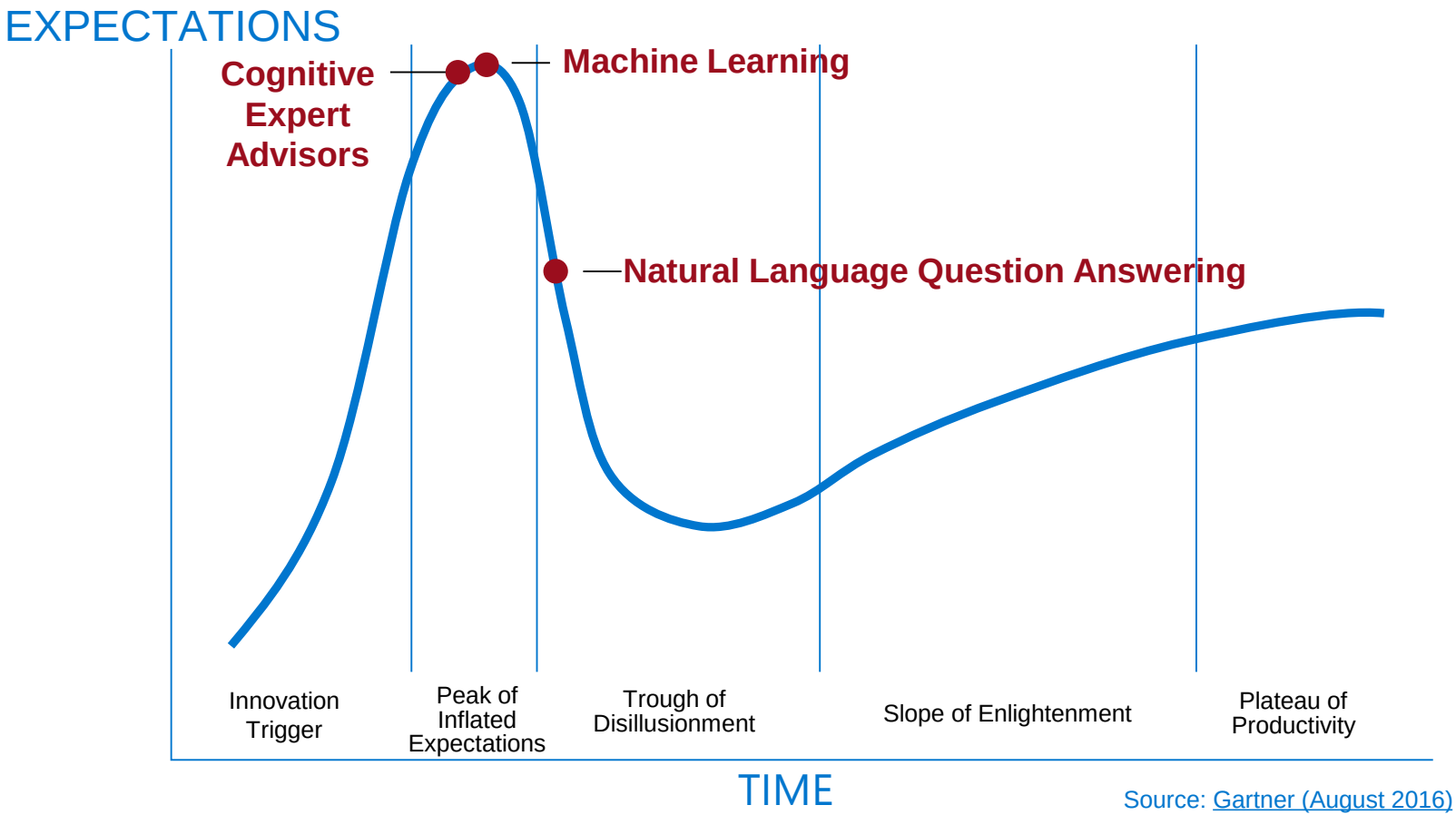
The AI/ML Journey courtesy Gartner Hype Cycle

Gartner Emerging Tech Hype Cycle - **2015**



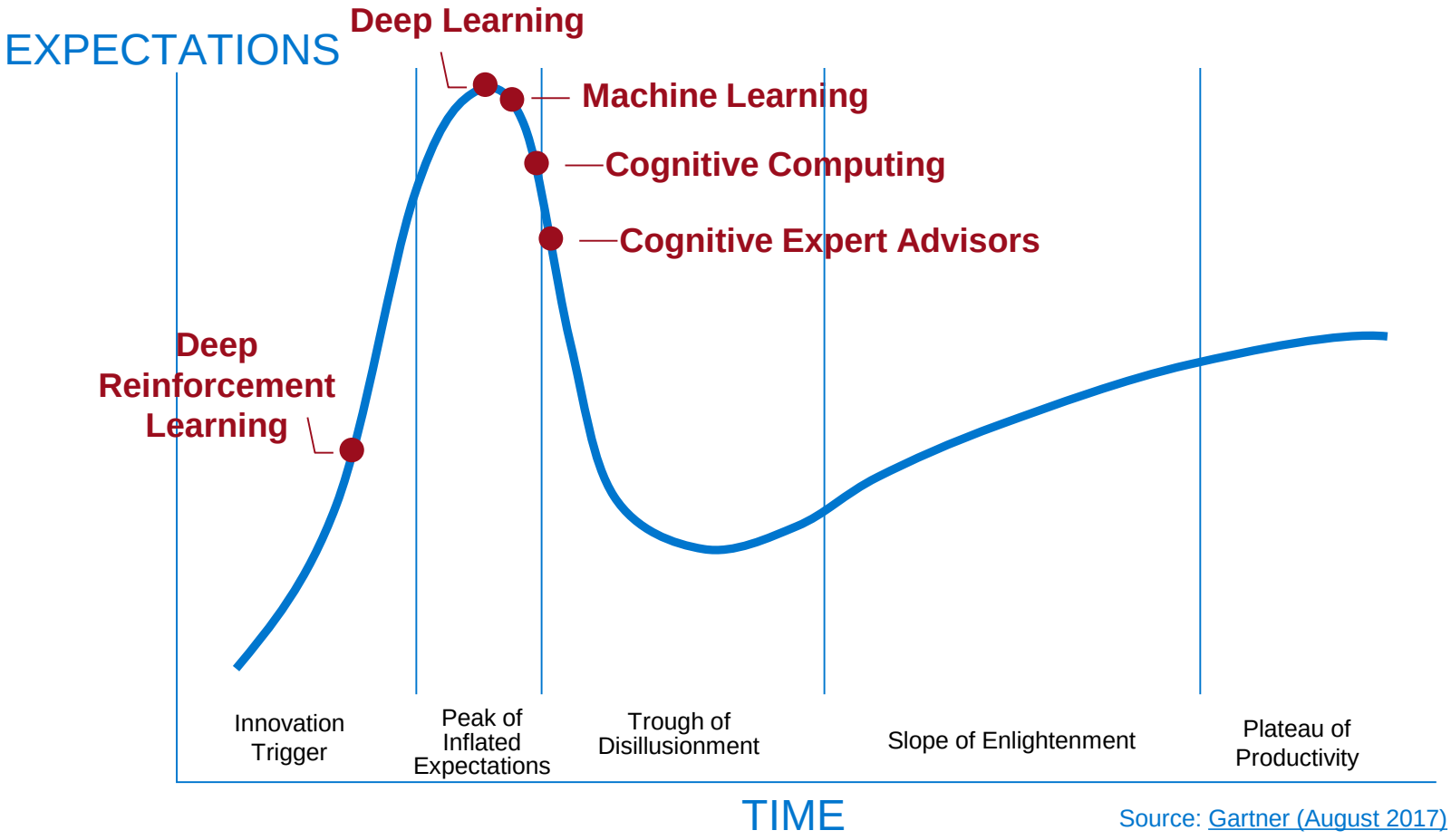
Source: Gartner (August 2015)

Gartner Emerging Tech Hype Cycle - **2016**



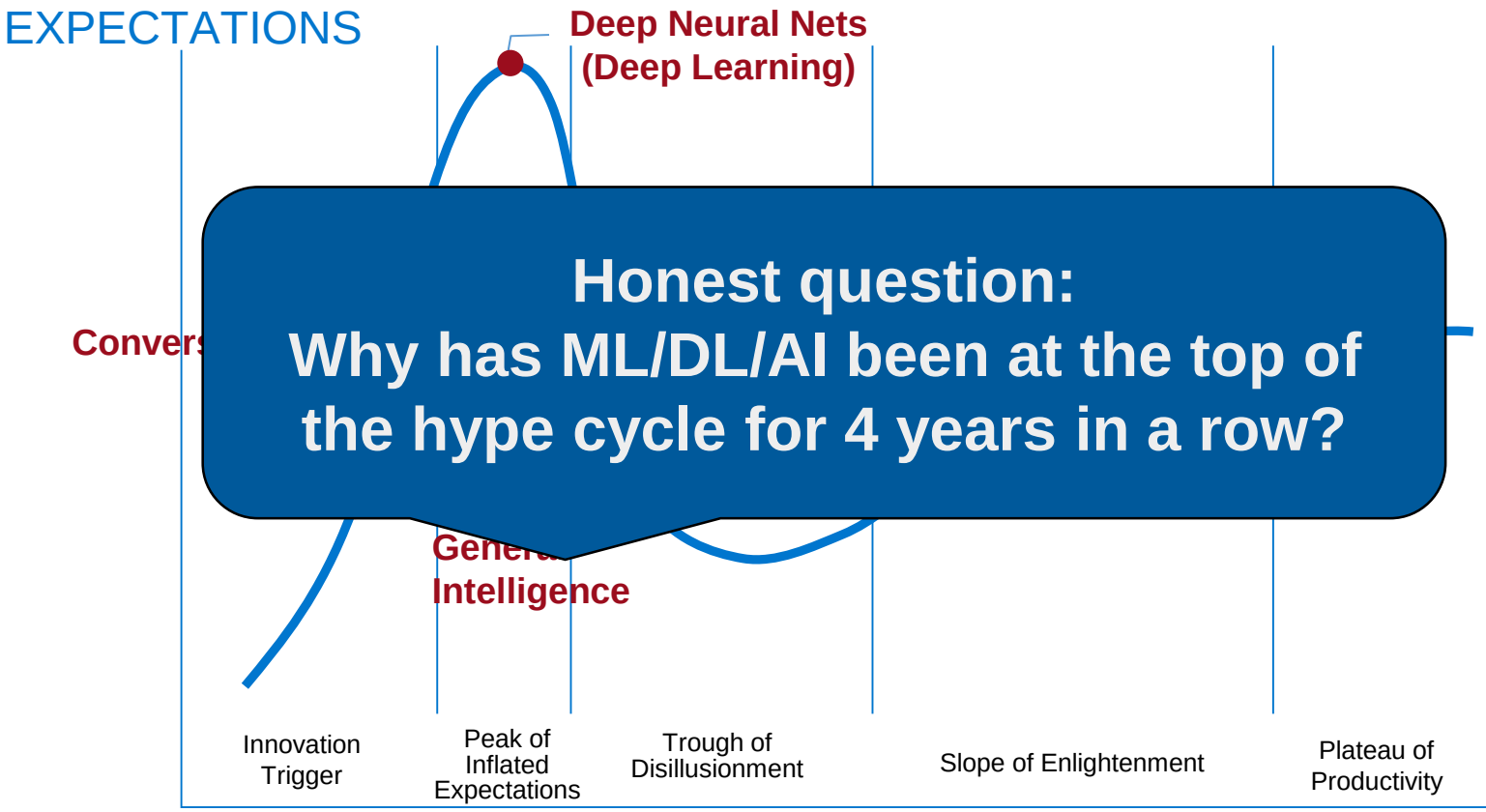
Source: Gartner (August 2016)

Gartner Emerging Tech Hype Cycle - **2017**

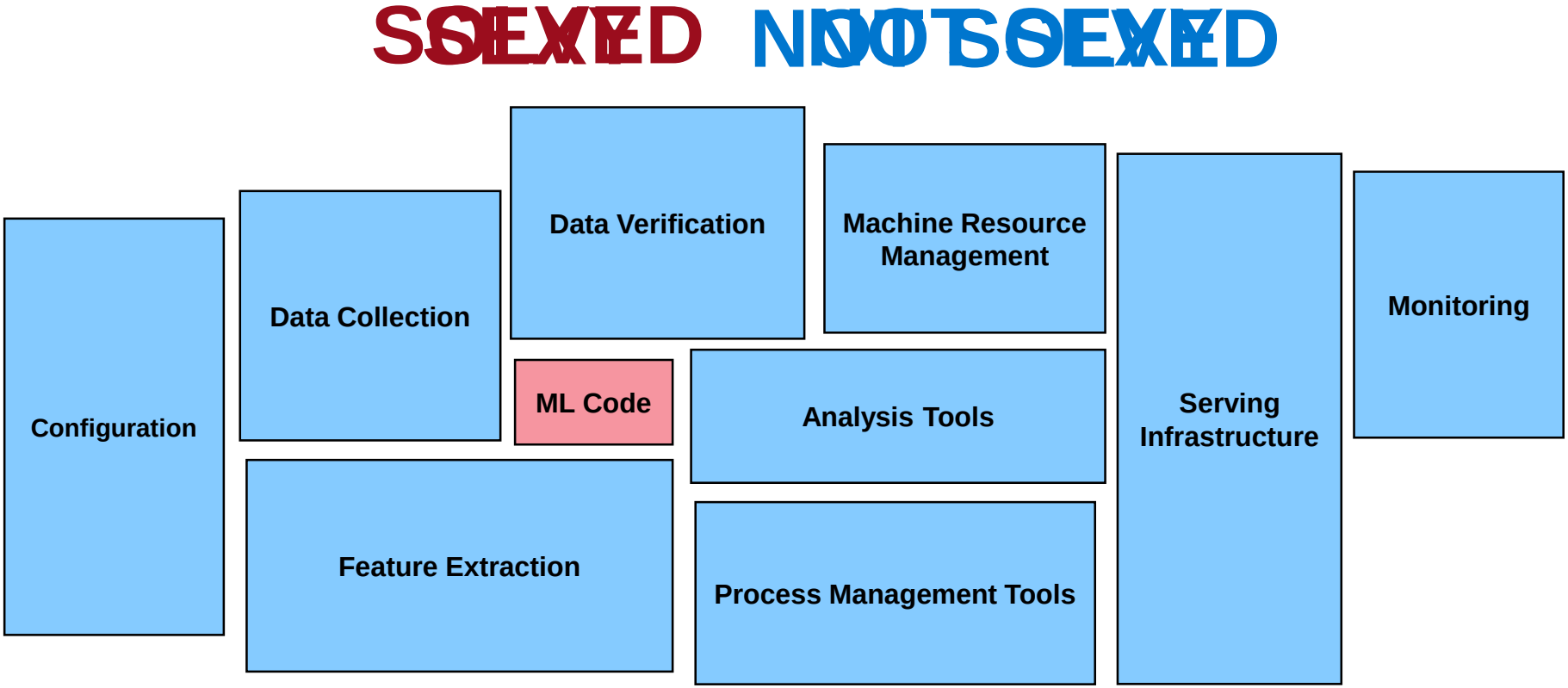


Source: Gartner (August 2017)

Gartner Emerging Tech Hype Cycle - **2018**

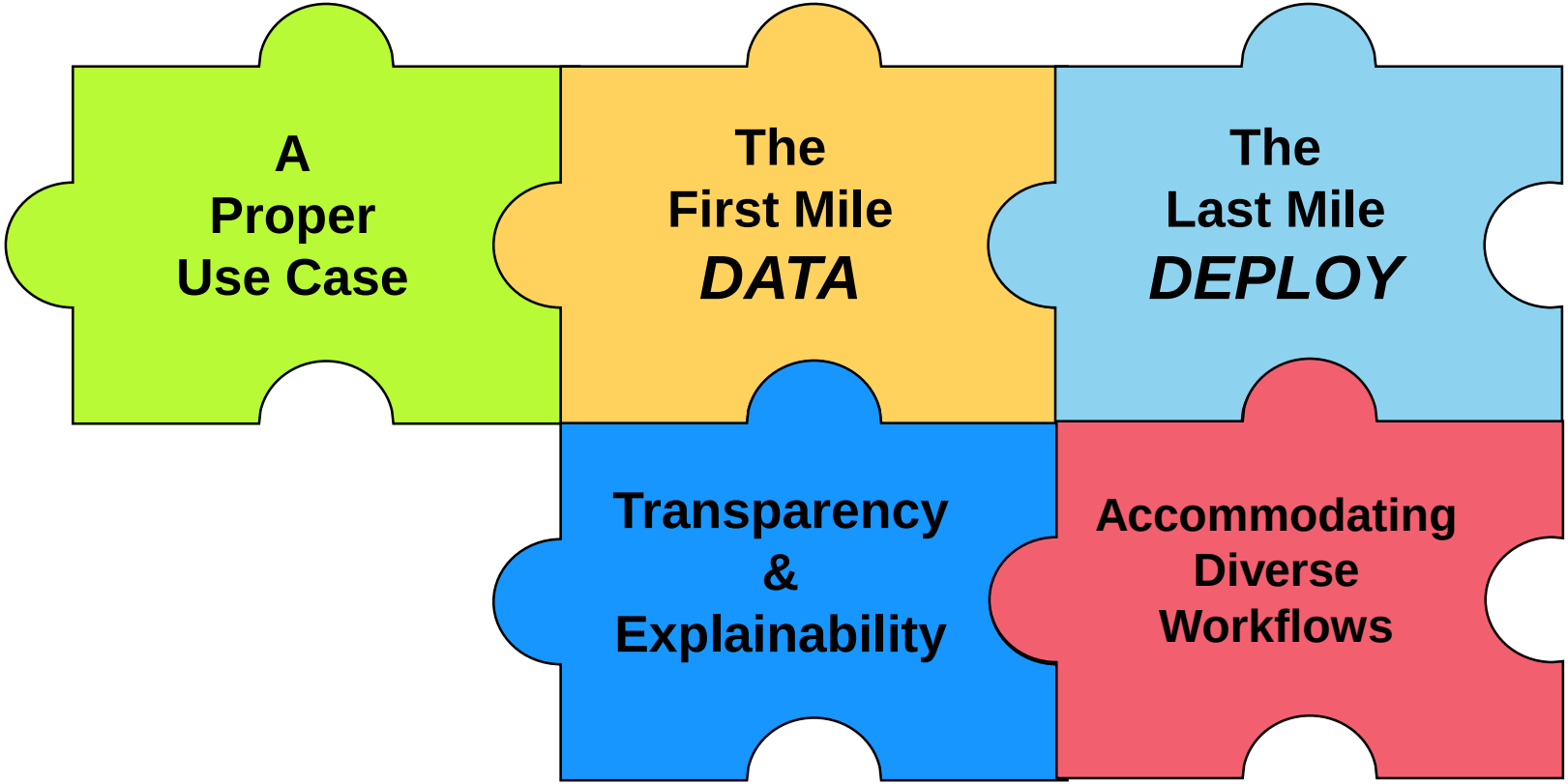


Source: Gartner (August 2018)

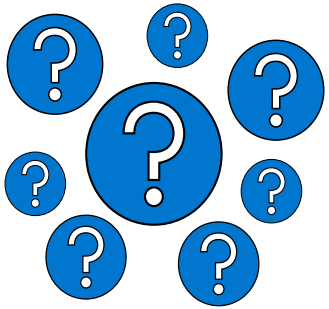


"Hidden Technical Debt in Machine Learning Systems"—Sculley et al, Google

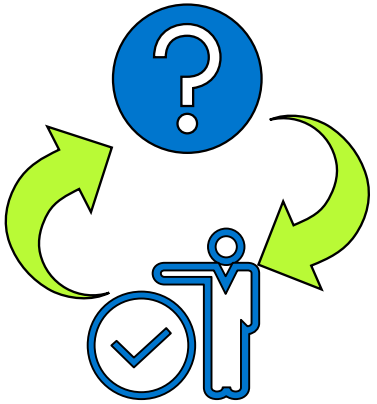
5 Pieces of the Unsolved Puzzle



A Proper Use Case

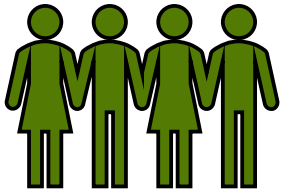


Same Question(s) Asked

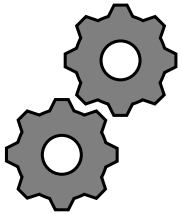


Repeatable Decision in point in time

Measurable Outcome



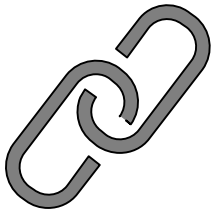
Human Capital Constrained



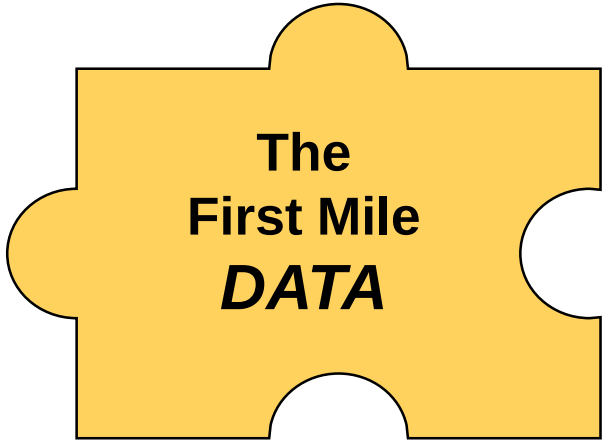
IT-Supportable (Batch, Real-time)

Starting with Wrong Problems

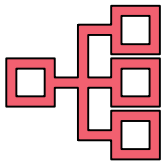
Make Data Come A.L.I.V.E. !



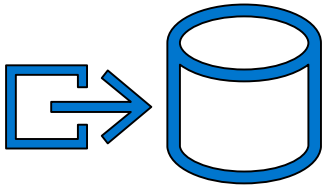
A ggregation



E nhancement



L ineage



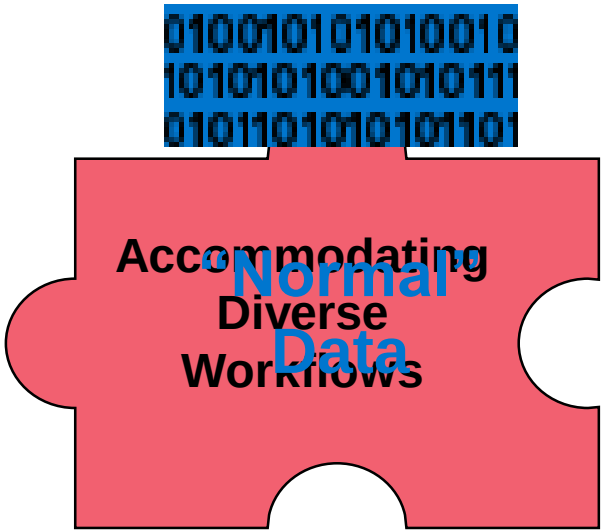
I ngestion

V alidation

What's Your (Data) Problem?

Workflow depends on the data, right?

Big Data



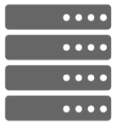
Forrester Study

Respondents site the following reasons for using workstations in AI workloads:

- **Price/Performance**
- **R&D w/ Flexible Timelines**
- **Offload server demand**

Workstations are essential as a development platform:

DEVELOPING AI APPLICATIONS



41% Servers/data centers

33% Workstations

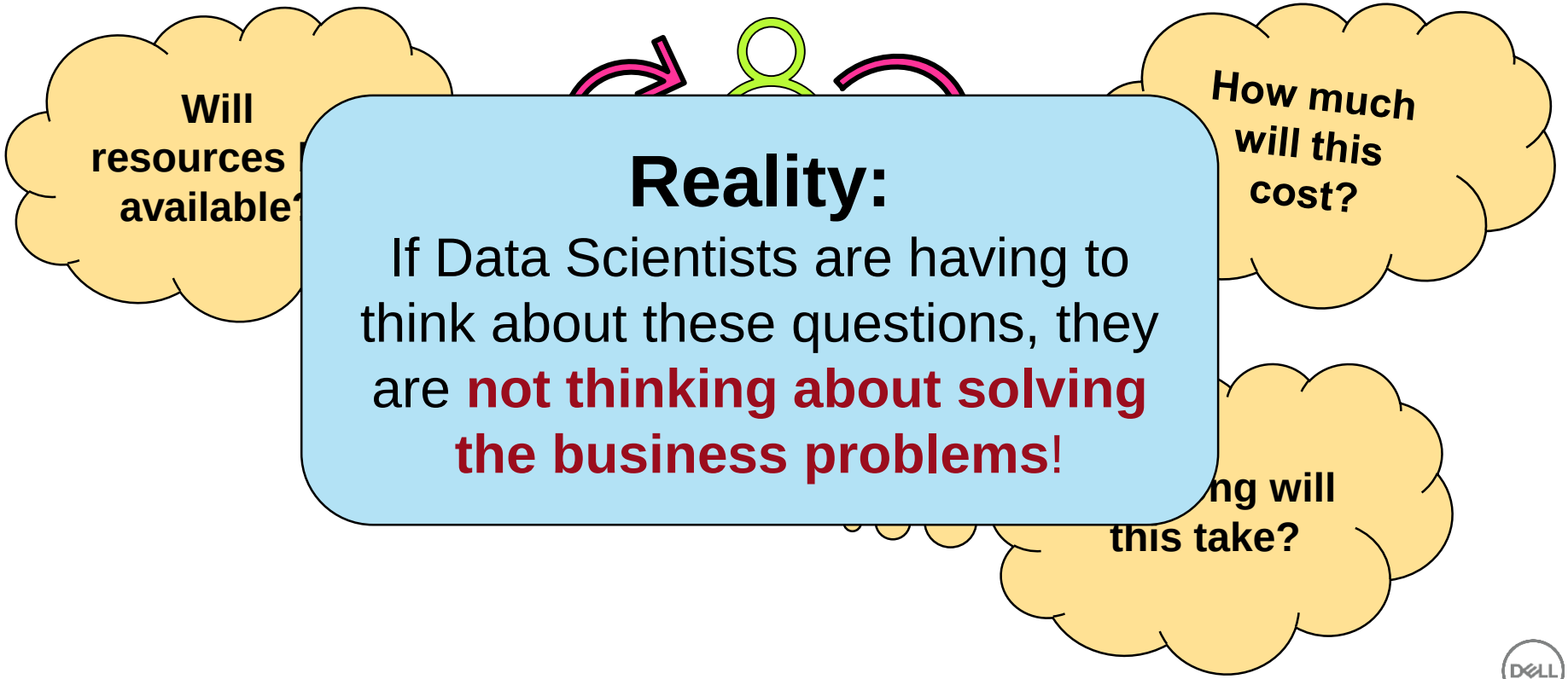
24% Cloud environments

Base: N 210 AI & Workstation decision-makers in North American companies with 100+ employees
Source: A commissioned study conducted by Forrester Consulting on behalf of Dell, April 2018

Data Science Workflow Considerations

- Resources
- Experimentation
- Agility
- Scaling
- Performance

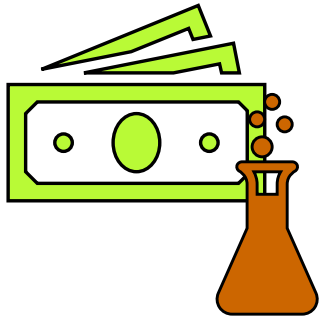
Resource Considerations



Reality:

If Data Scientists are having to think about these questions, they are **not thinking about solving the business problems**!

Experimentation Considerations



Science,
Experimentation,
and
Risk-taking

The
Lower-cost,
“All-Inclusive”
Alternative

Agility Considerations

EDA
&
Baseline Modeling

Before Deep Learning, there were techniques like:

- Chunking
- Subsampling
- Stratification

Surprisingly, they are still rather successful today for:

- Exploratory Data Analysis (EDA)
- Feature Engineering
- Baseline Model Building

DASK

data.table

sparklyr

kinetica BLAZINGDB brytlyt omni·sci

RAPIDS

Scaling Considerations (Containers)

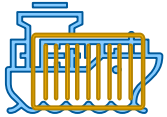
Containers lend themselves

- **Reproducibility**
- **Portability**
- **Streamlined Model Deployment**

Containers possess:

- **Data Science libraries and toolkits with complex dependencies**
- **Ability to simplify DevOps demands**

IT (Container Ship)



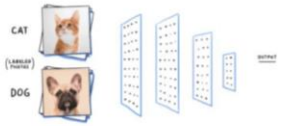
DellEMC Storage

Dell Precision Workstations

DellEMC PowerEdge

Performance Considerations

2-8 GPU server provide adequate performance to train complex state of the art models in reasonable time



Use Cases: Security Cameras, Facial Recognition, Tagging Photos, Autonomous Driving, Customer Service, Retail etc

Train the Resnet50 model from scratch to target accuracy (74.9%)

~10 hours for Resnet50 (2 x GV100)

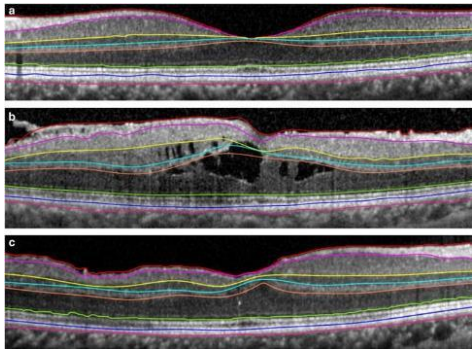
Success Considerations

voxeleron

Revolutionizing Ophthalmic Image Analysis

- Optical Coherence Tomography (OCT) allows for 3-D imaging of the Eye
- Neuro-degenerative diseases can be detected via OCT images (Alzheimer's, Parkinson's, ALS, MS, etc.)
- Using Dell Precision Workstations w/ NVIDIA GV100 graphics cards, Voxeleron has trained Deep Convolutional Neural Networks to detect known neuro-degenerative pathologies and incorporate these models into their InSight and InSight3D software for primary-care ophthalmologists.

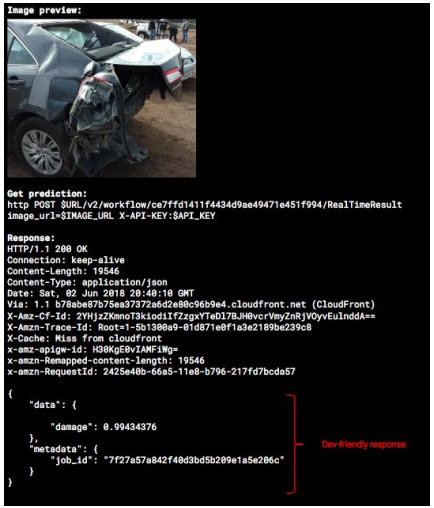
OCT Data Collection



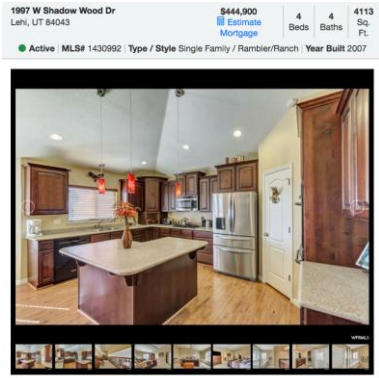
OCT Retinal Layer Thickness

Success Considerations

Transforming Business with AI in weeks not years … with existing teams

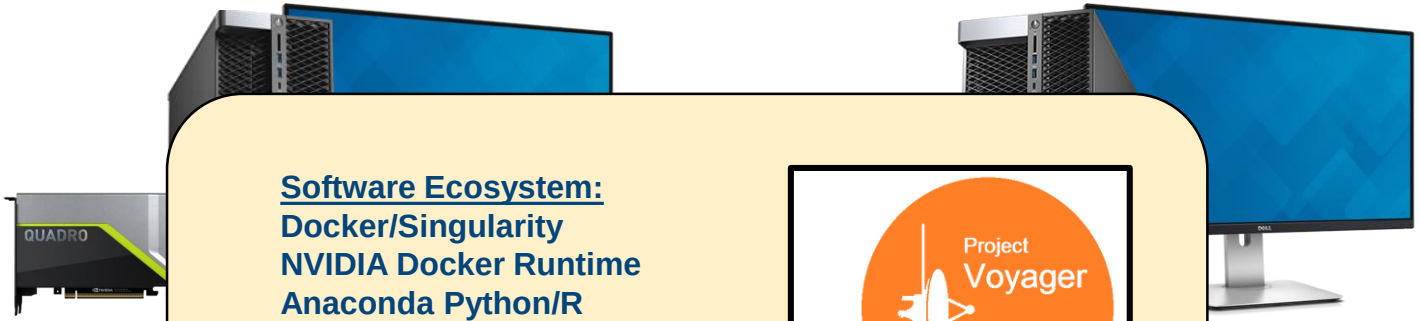


- The World's Only Unstructured Database
  - An AI system in a box.
- No data scientist necessary to build deep learning models.
  - A data engineer retrieves the data.
  - A domain expert ensures that value is being derived
- Models capable of being trained with mixed structure data (relational, image, audio, video, etc.)
  - Car damage – image, structured vehicle information
  - Home values – images, geospatial, structured data
- Using Dell Precision Workstations w/ NVIDIA GV100s in Proof of Concept engagements with customers.
- 20k rows x 6 million features -> trained in a day



Predicted house price
Structured only: r=0.6
ZIFF Holistic: r=0.92

Dell Workstation Offerings



Software Ecosystem:
Docker/Singularity
NVIDIA Docker Runtime
Anaconda Python/R
RAPIDS software
Deep Learning frameworks
NVIDIA GPU Cloud (NGC)
….. and more

Project Voyager
Precision

P **T7920**

Hardware:
GPU: 1x Quadro
CPU: Intel Xeon
Memory: 128GB D
HDD 1: M.2 1TB PCIe NVME Class 40 SSD
HDD 2: 3.5" 4TB 7200rpm Nearline SAS HDD
OS: Ubuntu 16.04

8000 w/ NVLink (48GB/96GB)
18-core (2.3GHz, 3.7GHz turbo)
2666MHz
HDD 1: M.2 1TB PCIe NVME Class 40 SSD
HDD 2: 3.5" 4TB 7200rpm Nearline SAS HDD
OS: Ubuntu 16.04

Dell Mobile Workstation Offerings

Precision 5530

Hardware:
CPU: Intel Xeon E-2176M, 6-core (2.7GHz, 4.4GHz turbo)
Or
CPU: Intel Core i9-8950HK, 6-core (2.9GHz, 4.8GHz turbo)
Memory: 64GB DDR4 2666MHz
HDD : M.2 1TB PCIe NVME Class 40 SSD
OS: Ubuntu 18.04

Precision 7730

Hardware:
GPU: NVIDIA Quadro P5200, 16GB GPU memory
CPU: Intel Xeon E-2186M, 6-core (2.9GHz, 4.8GHz turbo)
Or
CPU: Intel Core i9-8950HK, 6-core (2.9GHz, 4.8GHz turbo)
Memory: 128GB DDR4 2666MHz
HDD : M.2 1TB PCIe NVME Class 40 SSD
OS: Ubuntu 18.04

Come Visit

Booth #1311