Data Inputs for a Machine Learning Platform

Ryan McKinney
Data Strategy and Partnerships
Descartes Labs is a software company building a platform for modeling the world for self-driving enterprises. **Our solution** is enables **custom global scale analytics**

In particular, Descartes Labs offers:

- A single point of access to petabytes of public and private satellite data sources
- A cloud infrastructure to scale local pilots at continental levels on supercomputers
- A set of Python API's and web viewers to explore available geospatial data and present results
- A solutions team that builds highly-customized models for specific client questions

As a spin out from Los Alamos National Labs, with more than $55M in VC funding, and on-pace for $23M in commercial revenue in 2019, the Descartes Labs platform is currently utilized across a range of industries:

- Agriculture
- Oil & Gas
- Mining
- Div. Industrials
- Utilities
- Insurance
The sensor revolution is at an inflection point

- Increased variety
- Greater range & resolution
- Increased frequency
We can now digitize the physical world...
Aerial and other ground truth datasets play a key role when building ML models on a platform.
But the commercial market doesn’t buy raw datasets; it invests in actionable intelligence...
These foundational elements create critical business value

**Data Source**

*We ingest data from numerous sources*

**Data Refinery**

*We prepare that data for fusion with other datasets*

**Digital Twins**

*Factor models are built*

**Insights**

*Data and model results in the cloud supercomputer are as easy to access as a file on your hard drive*
We can refine the daily flood of data...

Raw Sensor Inputs

Refined Input

Analysis Ready Data
Extracting value from data: Crop Classification.
With eyes on every component of the supply chain, we leveraged public and proprietary data to achieve a superior forecast...
Properly leveraged data provides tangible benefits when combined with a ML platform

**STRATEGIC ROI**
Reduce mineral exploration from 6 mo to 6 days
Predict crop yields within 1% error
Forecast the natgas price changes daily
Connect property characteristics to actuarial risk 90% precision
Predict monthly freight forward rates within X%
Forecast fertilizer demand within 3% error
92% accuracy of high value crop acreage

**TACTICAL ROI**
Cut the cost of third party data acquisition by 200%
Significant reduction in data cleansing and calibration preparing data for analysis
all your data in one place; super computing access from your desktop [permission/governance]