Breakthrough Business Opportunities in the Oil Industry for YPs and "Small Team" Geoscientists*

Susan Nash[1]

Search and Discovery Article #70325 (2018)**
Posted March 5, 2018

*Adapted from oral presentation given at AAPG 2017 Annual Convention and Exhibition, Houston, Texas, April 2-5, 2017

[1]American Association of Petroleum Geologists, Tulsa, Oklahoma, United States (snash@aapg.org)

Abstract

Many of the opportunities for young professionals and geoscientists who are in small, interdisciplinary teams, often as a part of a funded venture to develop a new play or prove a new production—enhancing technology, involve Big Data, Deep Learning, and other forms of analytics.

There are many existing and emerging situations where Deep Learning, that is, artificial intelligence with the ability to be used in "smart systems" can be useful to the geoscientist.

Geophysicists have long used Deep Learning in conjunction with seismic processing and interpretation, particularly where the data sets are constantly changing and are massive, and where the types of data are often dissimilar and have to significantly manipulated in order for them to be incorporated into algorithms that ultimately yield meaningful patterns and relationships. The results are valuable in making decisions, such as assessing the value of a property or in predicting where to find sweet spots.

The new opportunities are in many phases of the exploration and development process. Here are a few:

1. Find ways to manage, store, and retrieve data of all kinds. Once you have identified the data you need, it is important to convert the unstructured into structured data so that you can work with it. Dell EMC and Amazon Web Services are probably the most widely recognized, while Akamai is probably the oldest.

2. Become an expert in the software used for working with Big Data. One example would be the person who decides to become an expert in working with Cloudera or Hortonworks's Hadoop distributed file system and the various solutions for an enterprise data warehouse (EDW).

3. Develop smart systems and deep learning by creating algorithms that can be trained to recognize the patterns that are valuable to your enterprise. Educate yourself in the area of Supervised, Unsupervised, and Reinforcement Learning.

4. Develop new applications that utilize the power of Big Data and Deep Learning to integrate data sets, find patterns, and enable you to create better maps and make better recommendations. A few include developing maps of migration pathways, identifying sweet spots, automatic acreage valuations, “smart” proppants, and smart oilfield chemicals.

Breakthrough Business Opportunities in the Oil Industry for YPs and "Small Team" Geoscientists

Susan Nash, Ph.D.
AAPG

Big Data and Deep Learning Today

- Storing, accessing, and processing data across wide spectrum of possibilities (cloud-based, distributed)
- Instantaneous access at low cost
- Modeling for monitoring and decision-making
- Incorporating wide array of data
- Tackling new kinds of business problems

Examples: Big Data

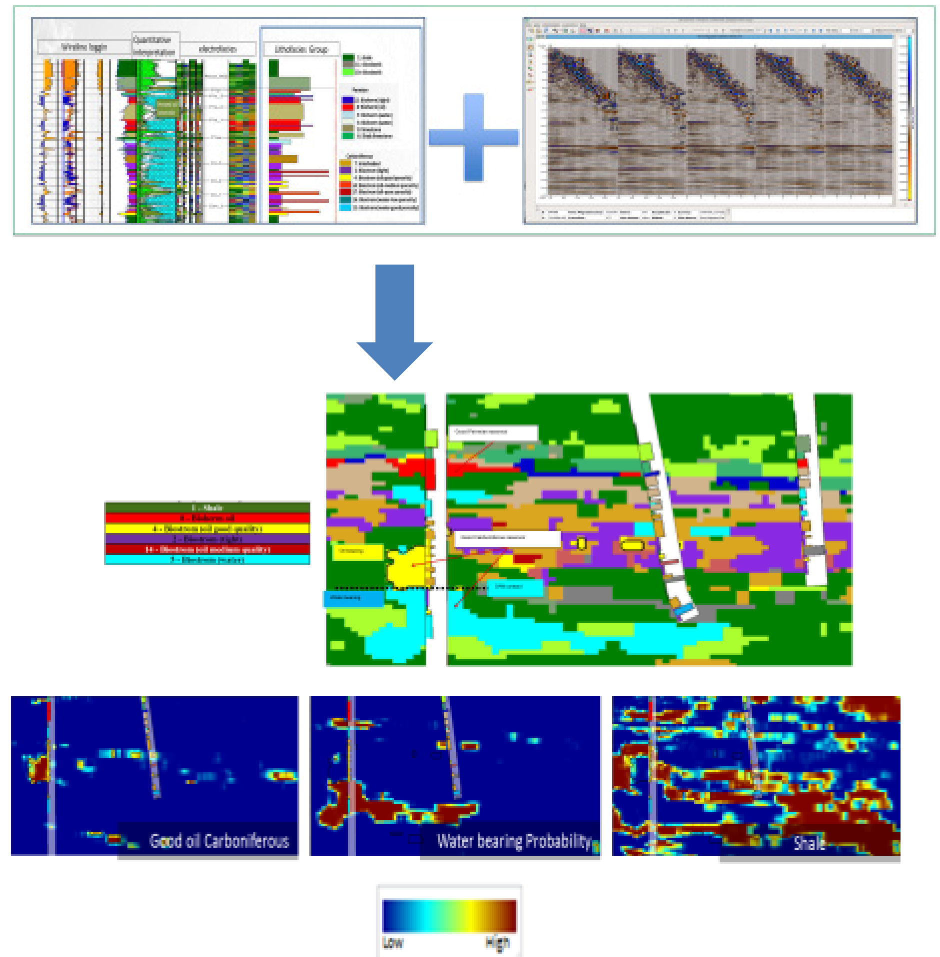
- Evaluating acreage and ranking prospects
- More accurate processing of seismic
- Identify sweet spots
- Reservoir characterization using unique blends of data for deeper understanding
- Identifying opportunities ("botched" shale plays or "botched" EOR) by being able to analyze massive data sets
- Many more examples

Examples: Deep Learning

- Direct Hydrocarbon Indicators of much more complexity, and which integrate surface as well as subsurface data
- “Smart Systems” – self-driving drillbit
- The “responsive” pump (”learns” how best to optimize production from numerous sensors that tie to performance history / history matching

Probabilistic Rock Type Classification (Paradigm)

- Generate probabilistic facies volumes calibrated to facies logs
- Integrated QC steps at wells to validate Neural Network training
- Enables high resolution facies results and good reservoir heterogeneity matching



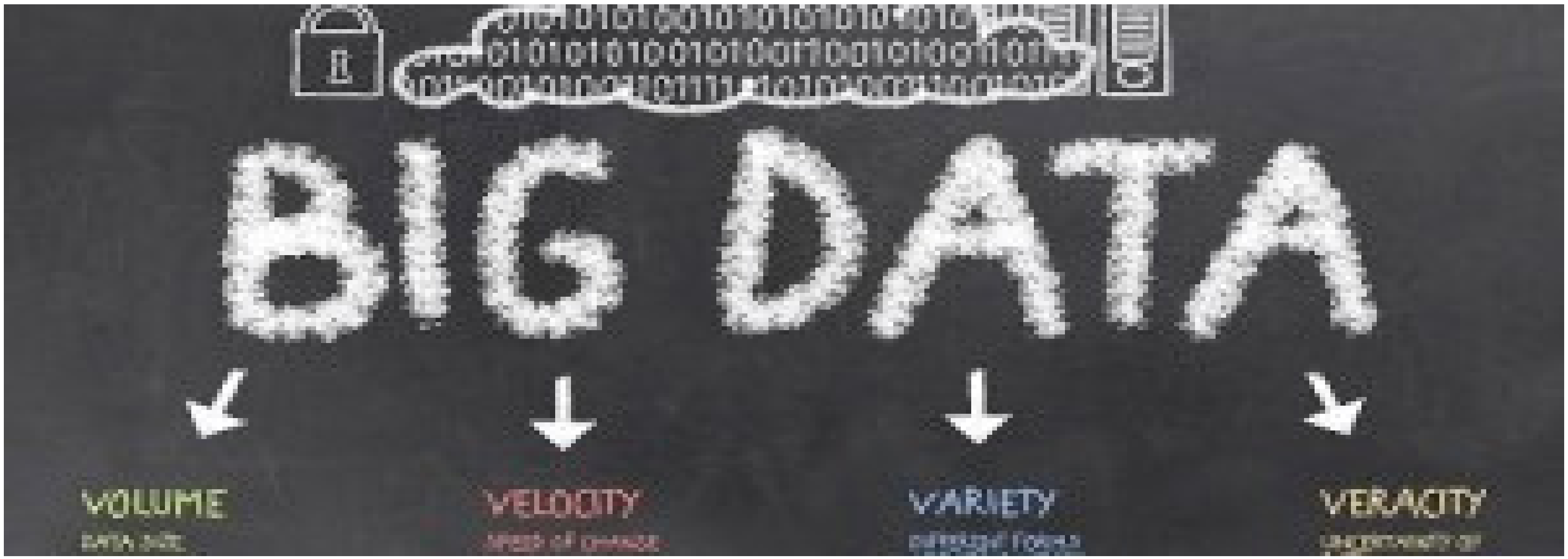
Step 1: Find and Store Data - How?

- Apache Hadoop (highly recommended / wide adoption, flexible)
- Companies that help you with storage
  - Dell EMC
  - Intel's Nervana

Big Data in Upstream Exploration

- Making good decisions quickly (prospectivity indices, sweet spot identification, risk assessment)
- Visualization – understanding the options and relationships in new ways

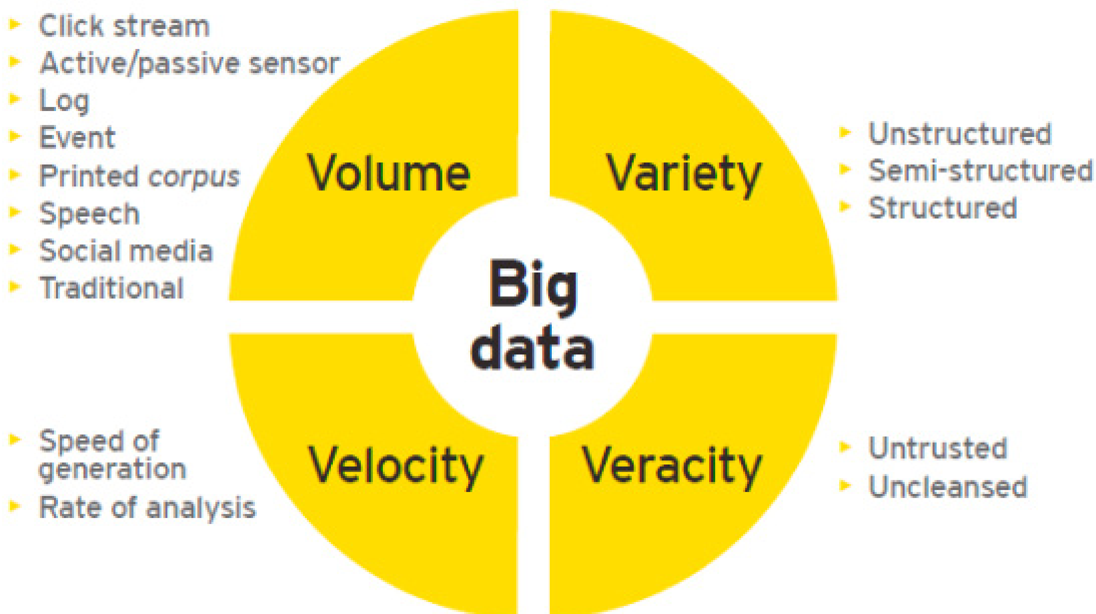
What is Big Data? (Doug Laney)



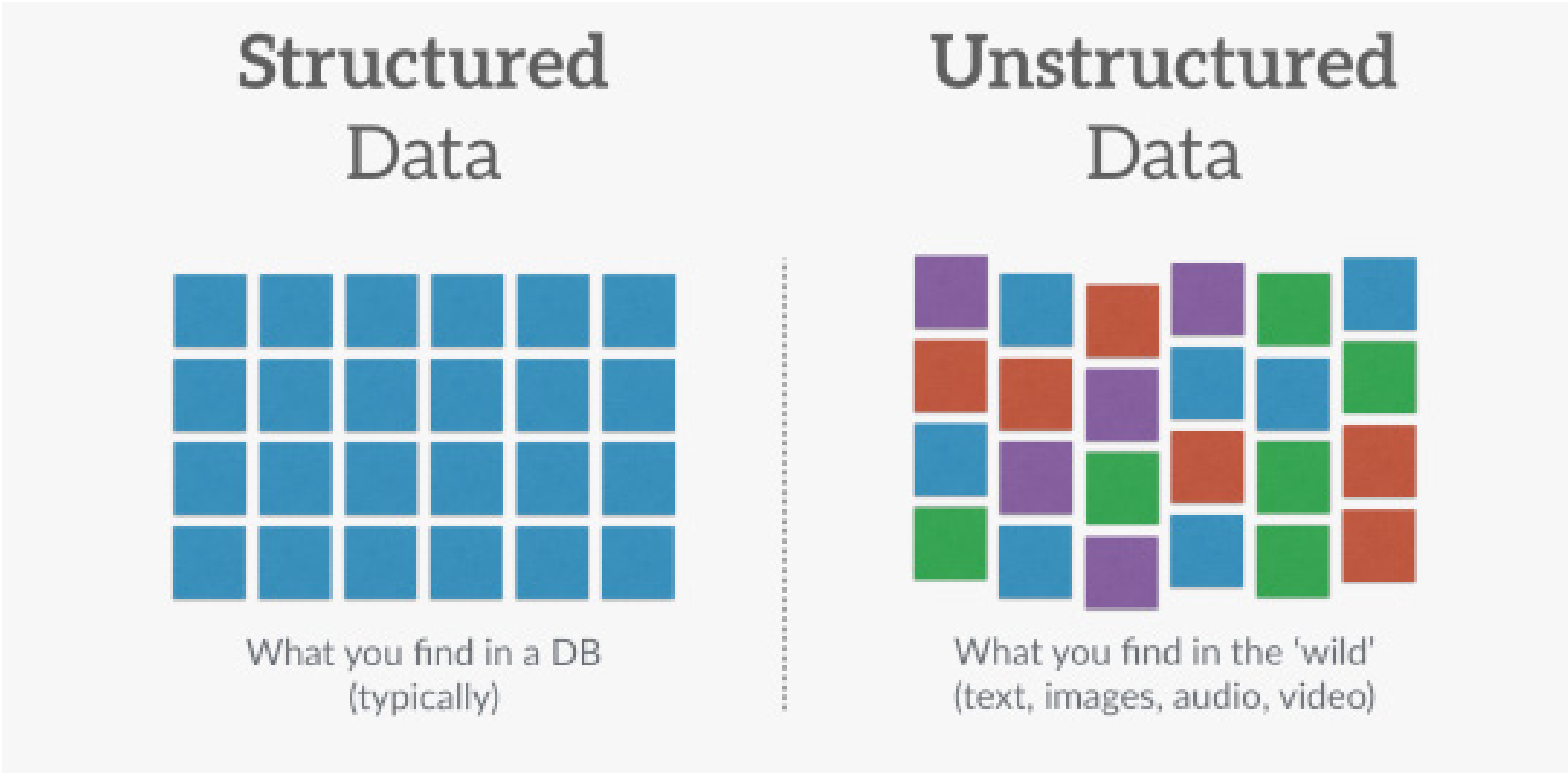
Volume: collecting large volumes of data from sensors or machine-to-machine sources
Storage: new technologies (such as Hadoop) allow distributed cloud storage
Velocity: Data streaming at rapid pace / torrents of data – must deal in near-real time
RFID tags (Radio Frequency Identification -- "chip 'em"
Sensors
Smart metering
Variety: Wide variety of formats from structured to unstructured
Structured Data: traditional databases
Unstructured Data: text documents, email, video, audio, financial transactions, sensors
Other considerations (from SAS)
Variability:
The velocity of data flow can be extremely inconsistent
The types of information are constantly changing (social media, etc.)
Complexity:
Multiple sources of information in different forms
Must find commonalities; a common “language” for the data to converse
Hierarchies
Relationships
Multiple data linkages

Now 7 V's added to the 4 V's

- Volume
- Variety
- Velocity
- Veracity
- Value
- Virtual
- Variable



Unstructured vs Structured Data



Unstructured in G&G

- Unstructured data sources: Data that's not in structured, relational databases. So, they could be structured or semi-structured. Examples: email, social data, XML, videos, audio, photos, GPS, satellite images, sensor data, spreadsheets, petrophysics, well logs, production history, pressure gauges, logging while drilling information, RFID tags, PDF documents.

Step 2: Make the data usable – How?

Making the unstructured data usable

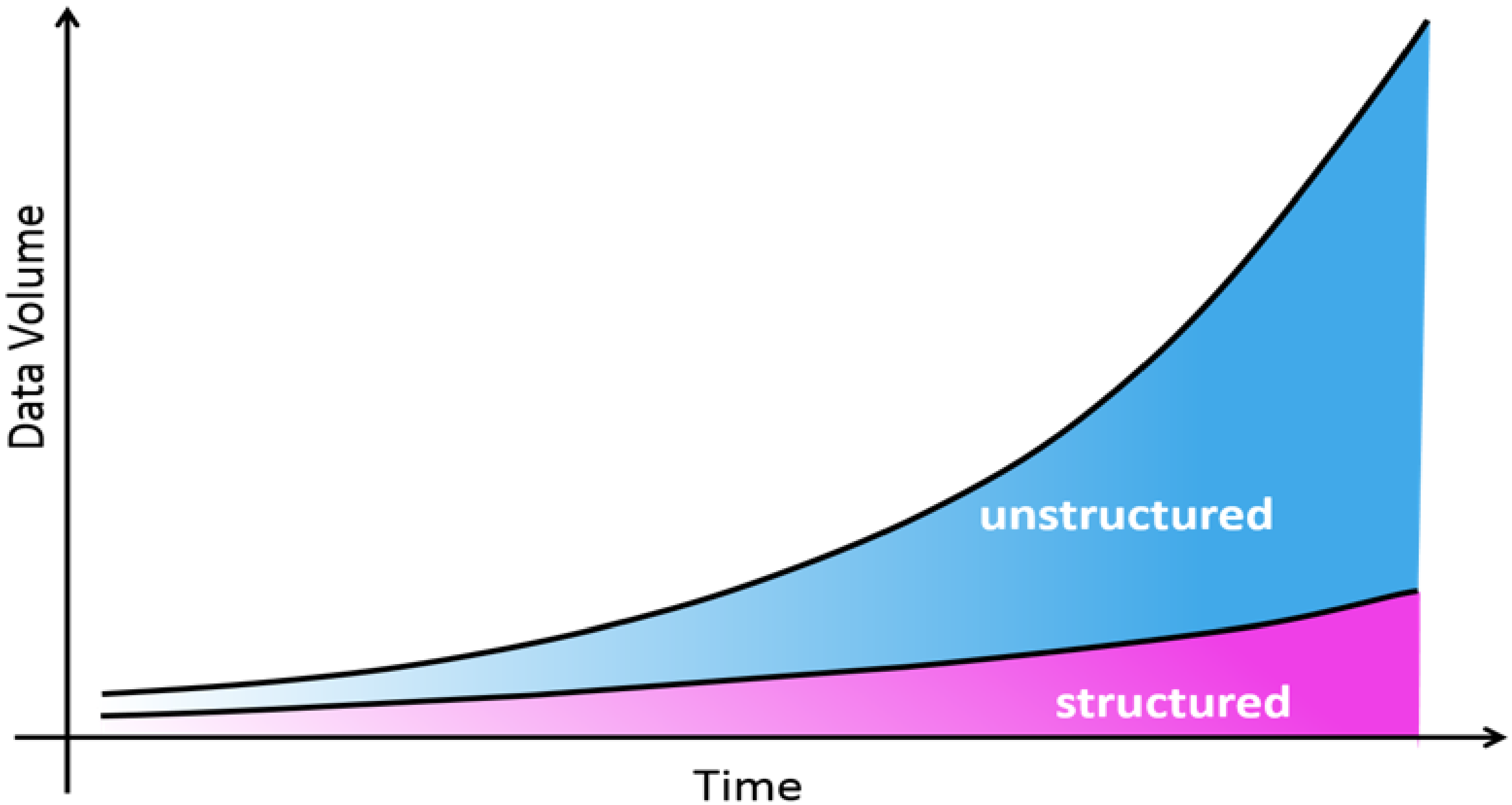
AgileDD's approach – convert unstructured to structured

Instead of 4,000 man hours, 2 weeks

Cleaning up the noise

Oceanit

As time goes on, an increasing need



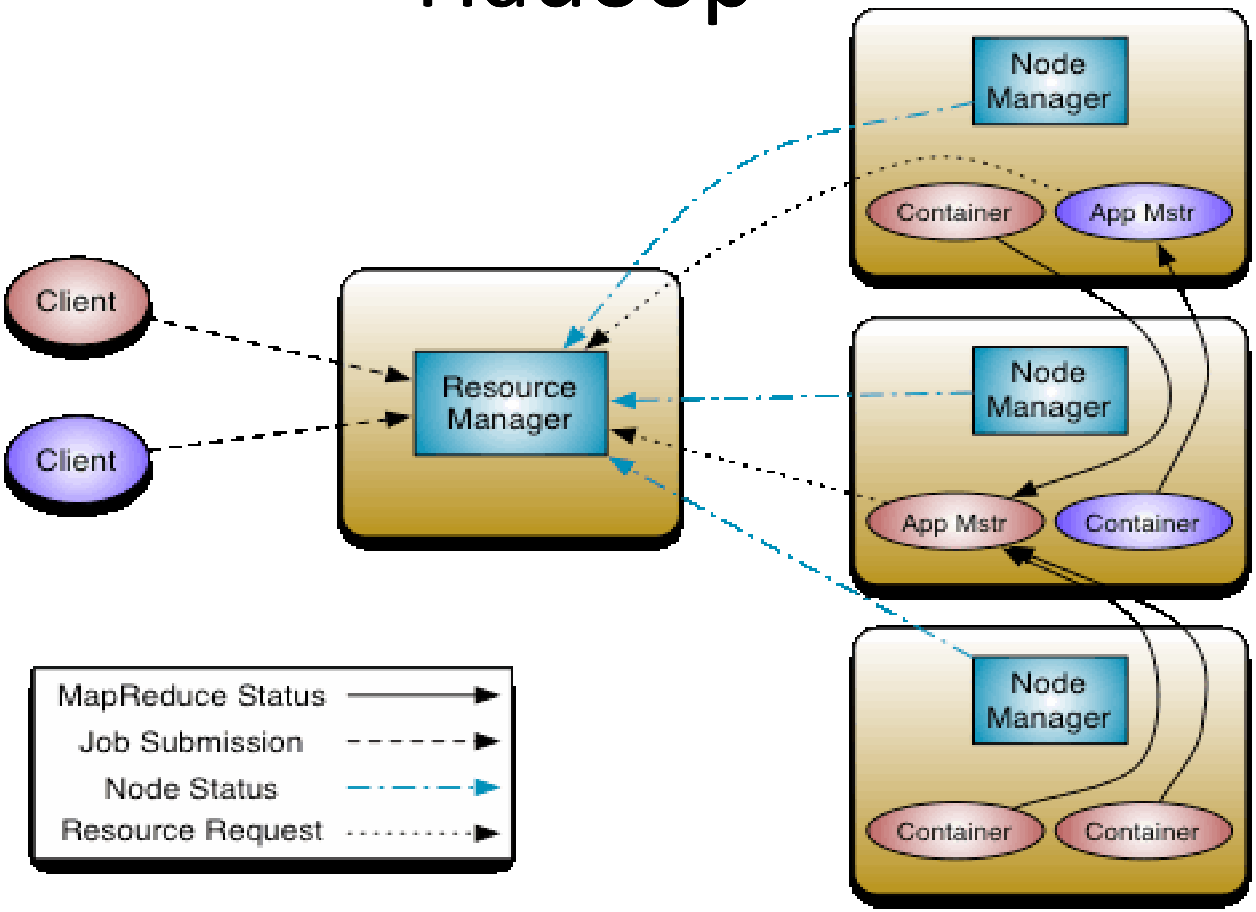
Processing Non-Integrated and Unstructured Data

- Must ask a question: What kind of data do you really want to include? All? Some? None? Why? What will relating completely different types of data potentially reveal to you?

Step 3: Access & Consume Data – How?

- Develop (or incorporate) workbenches
- Hortonworks Hadoop Sandbox

Hadoop



doop (HDFS): The Hadoop Distributed File System (HDFS) is the data storage component of the open source Apache Hadopp project. It can store any type of data (structured, semi-structured and unstructured). Can be run on low-cost computers / hardware and can scale out across thousands of machines for quick responses (and redundancy).

Hadoop YARN: Framework for job scheduling and clustering resources

Part of Hadoop: MapReduce – resource management and processing component of Hadoop. It incorporates YARN.

Can write MapReduce jobs to bring application data into Hadoop and the Hadoop Distributed File System (HDFS) – use Hadoop storage, and increase processing power

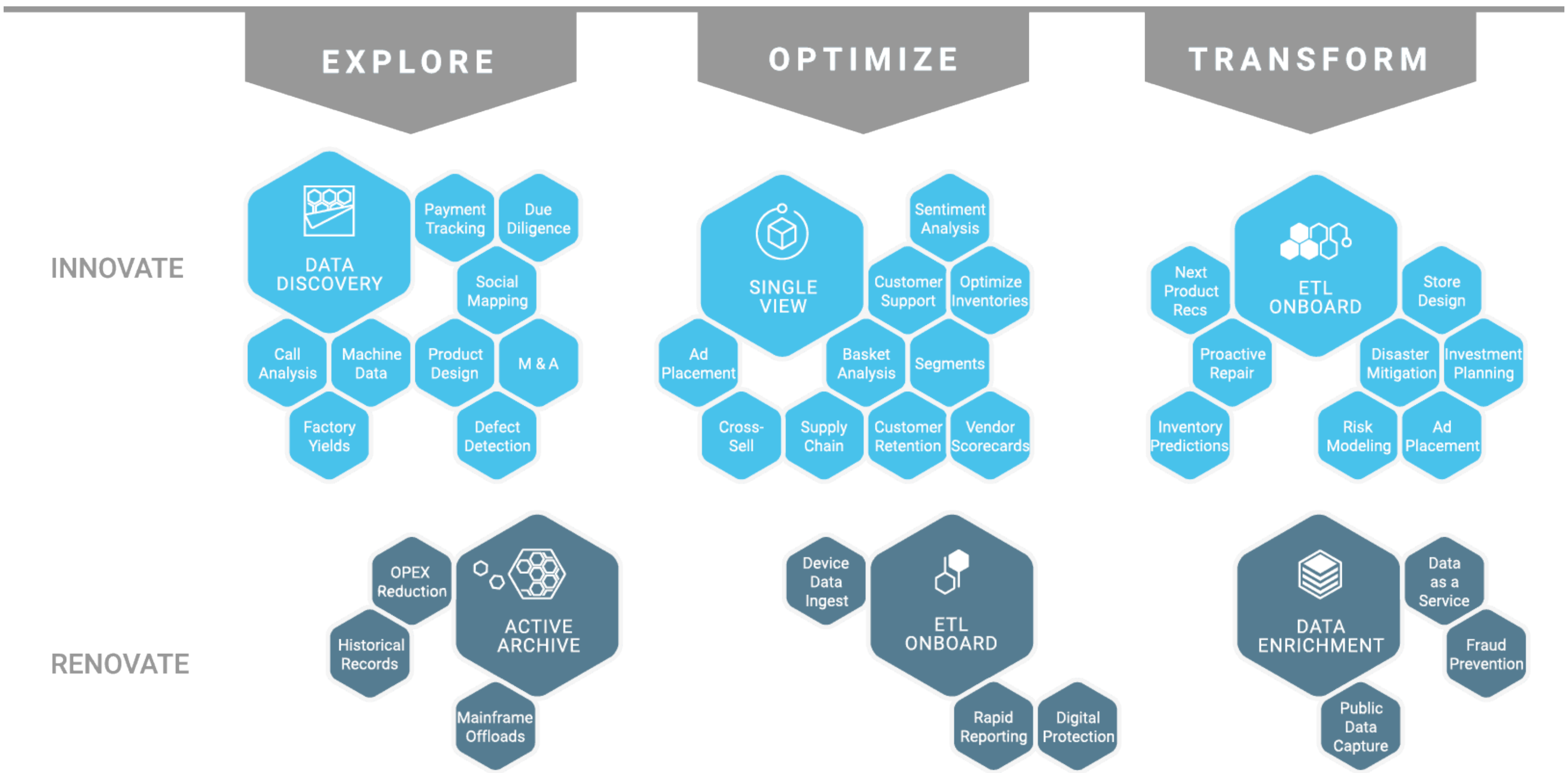
Hadoop

- The Hadoop Distributed File System (HDFS) is the data storage component of the open source Apache Hadoop project. It can store any type of data (structured, semi-structured and unstructured). Can be run on low-cost computers / hardware and can scale out across thousands of machines for quick responses (and redundancy).

Hortonworks

- Put all structured data in Hadoop (don't put it in the Enterprise Data Warehouse - EDW)
- Process unstructured data and integrate it into Hadoop
- Hadoop – can be the place you go to for clean, integrated data that you put in your EDW
- Analytics tools sit outside and go in and access the data as needed (in Hadoop (the raw data) and in the clean data in the EDW)
- Big data apps – can use many of them, and they can be used often

Hortonworks, Cloudera, etc.



Step 4: Building the Algorithms – How?

- Create your own algorithms, or use custom applications
  - Intel Nervana
- Python deep learning modules
  - Petrabytes
  - Halliburton

What are companies doing now?

- Devon and Data
  - Step 1: Identify and store all data
  - Step 2: Clean up the data
    - Structured and unstructured, streaming, real-time
  - Step 3: Structure in better form
    - Dashboards, GIS, mobile, more
  - Step 4: Apply algorithms
    - Machine learning, artificial intelligence, etc.

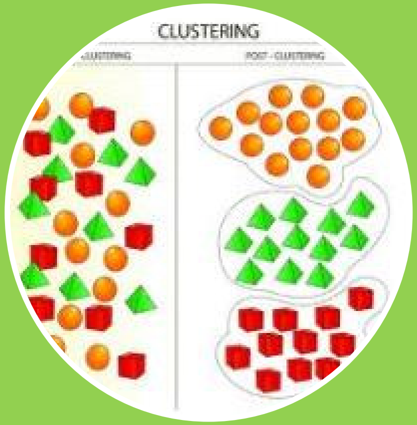
Supervised Learning

TEACH DESIRED BEHAVIOR WITH LABELED DATA

MAKE SENSE OF NEW DATA BASED ON PRIOR DATA

Unsupervised Learning

MAKE INFERENCES WITHOUT LABELED DATA



DISCOVER UNKNOWN OR HIDDEN PATTERNS

Reinforcement Learning

ACT IN AN ENVIRONMENT TO MAXIMIZE REWARD

BUILD AUTONOMOUS AGENTS THAT LEARN

End Product Example:
Deep Learning: DeepFace Architecture
Face = Sweet Spot?

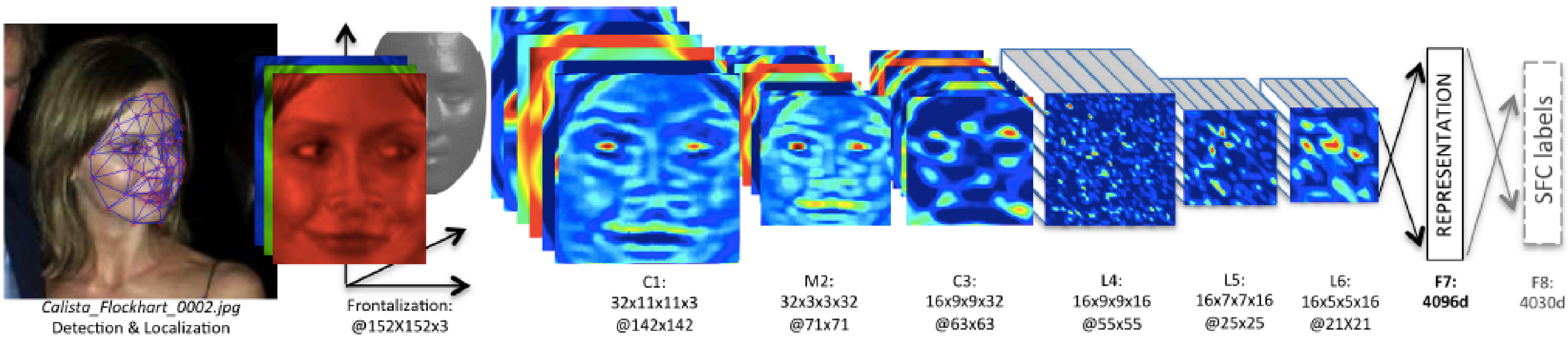


Figure 2. **Outline of the *DeepFace* architecture**. A front-end of a single convolution-pooling-convolution filtering on the rectified input, followed by three locally-connected layers and two fully-connected layers. Colors illustrate outputs for each layer. The net includes more than 120 million parameters, where more than 95% come from the local and fully connected layers.

https://gigaom.com/2015/03/06/how-paypal-uses-deep-learning-and-detective-work-to-fight-fraud/

Unsupervised Learning:

- Detect anomalies
- Social network analysis
- Market segmentation
- Astronomical data analysis

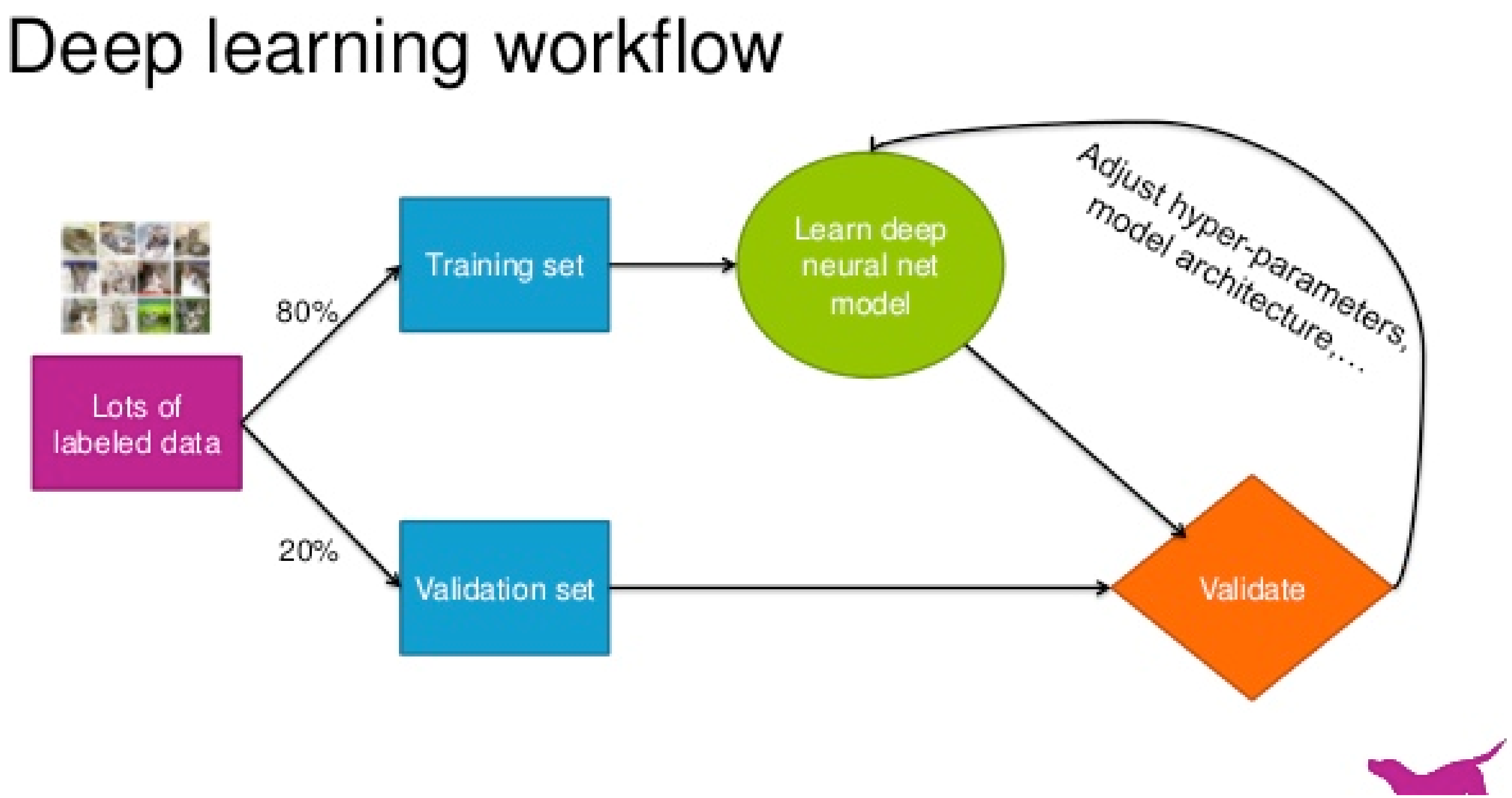
Supervised Learning

Excellent for predictions because there is a training set; good for building predictive models

Two categories of algorithms:

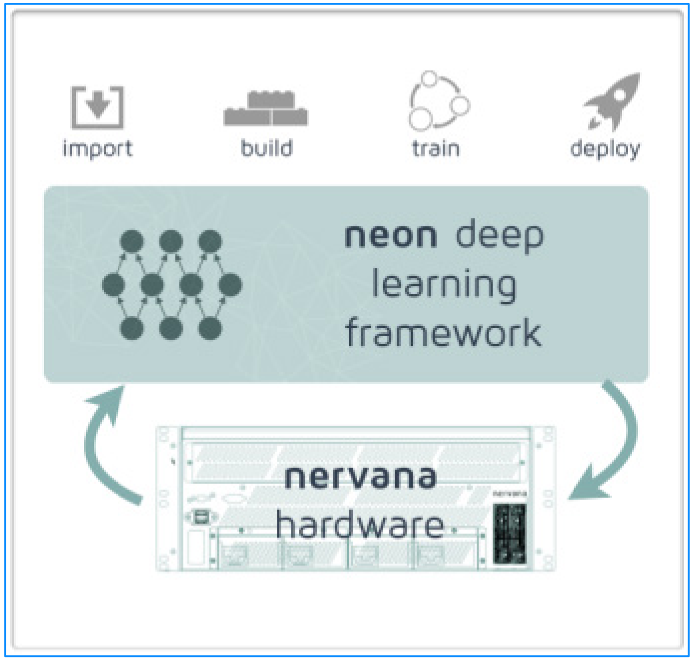
- **Classification: Categories of responses**
  - Neural networks
  - Decision trees
  - Discriminant analysis
  - Nearest neighbors (*k*NN)
- **Regression: Continuous-response values**
  - Linear regressions
  - Non-linear regressions
  - Decision trees
  - Neural networks

How the training sets work

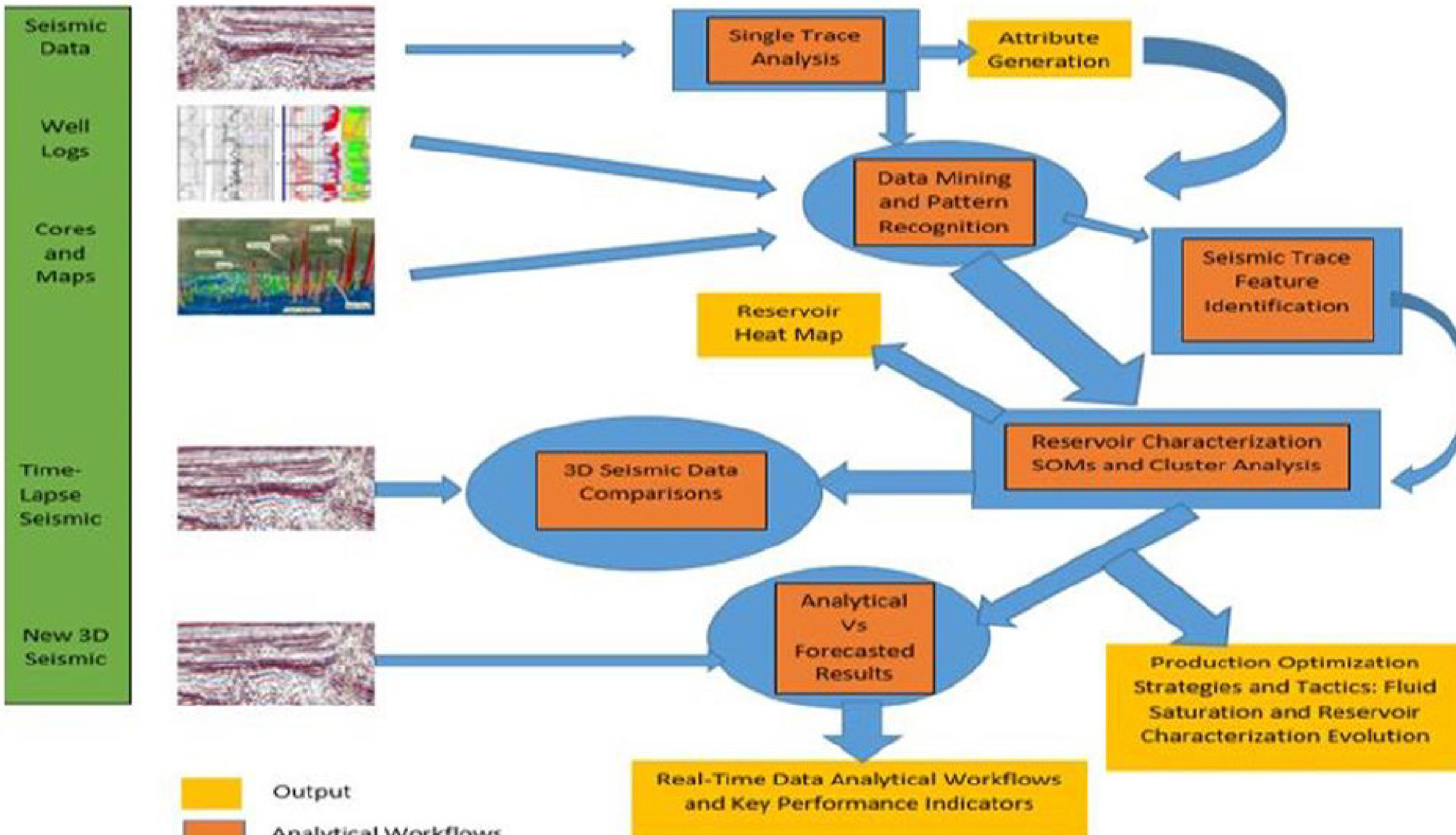


Implement Deep Learning

- Use outsourced services
- Hosting
- Deep Learning framework



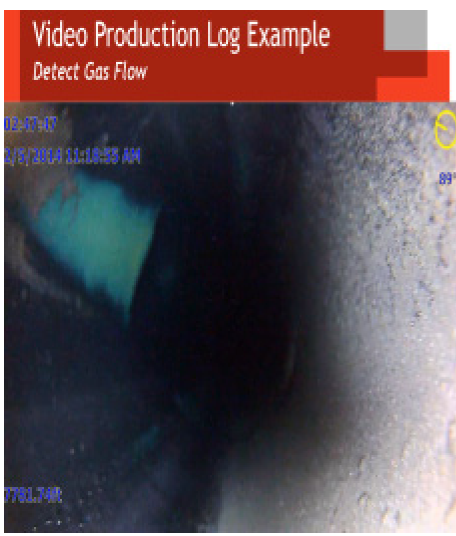
Seismic has been using neural networks and pattern analysis this for a long time. But…. In the past, it was hampered by constraints of hardware, training sets, data.

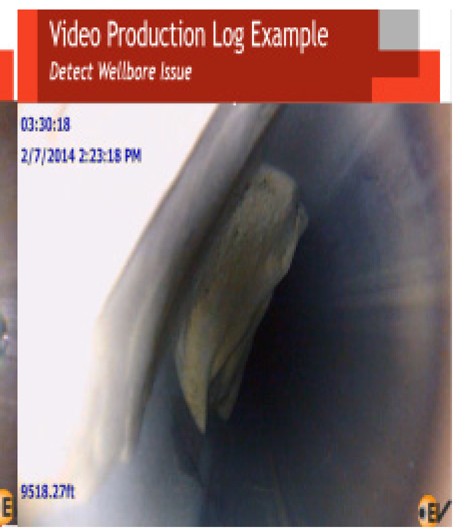


Deep Learning in Exploration and Development

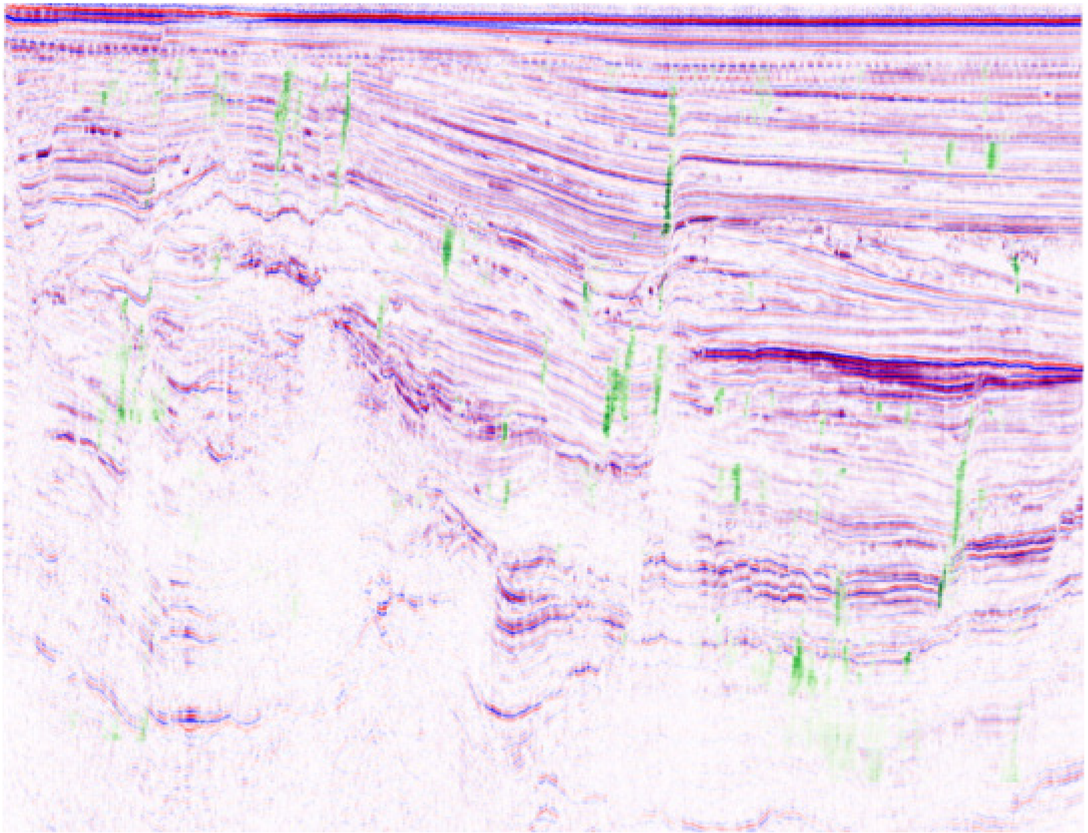
- Better geosteering
- Sweet spot identification
- Refrac zones and planning
- Completion monitoring / smart systems
- Completions and Production machine-based decisions

Devon: Real-time analysis of reservoir fluids during production using video production logging: optimize production

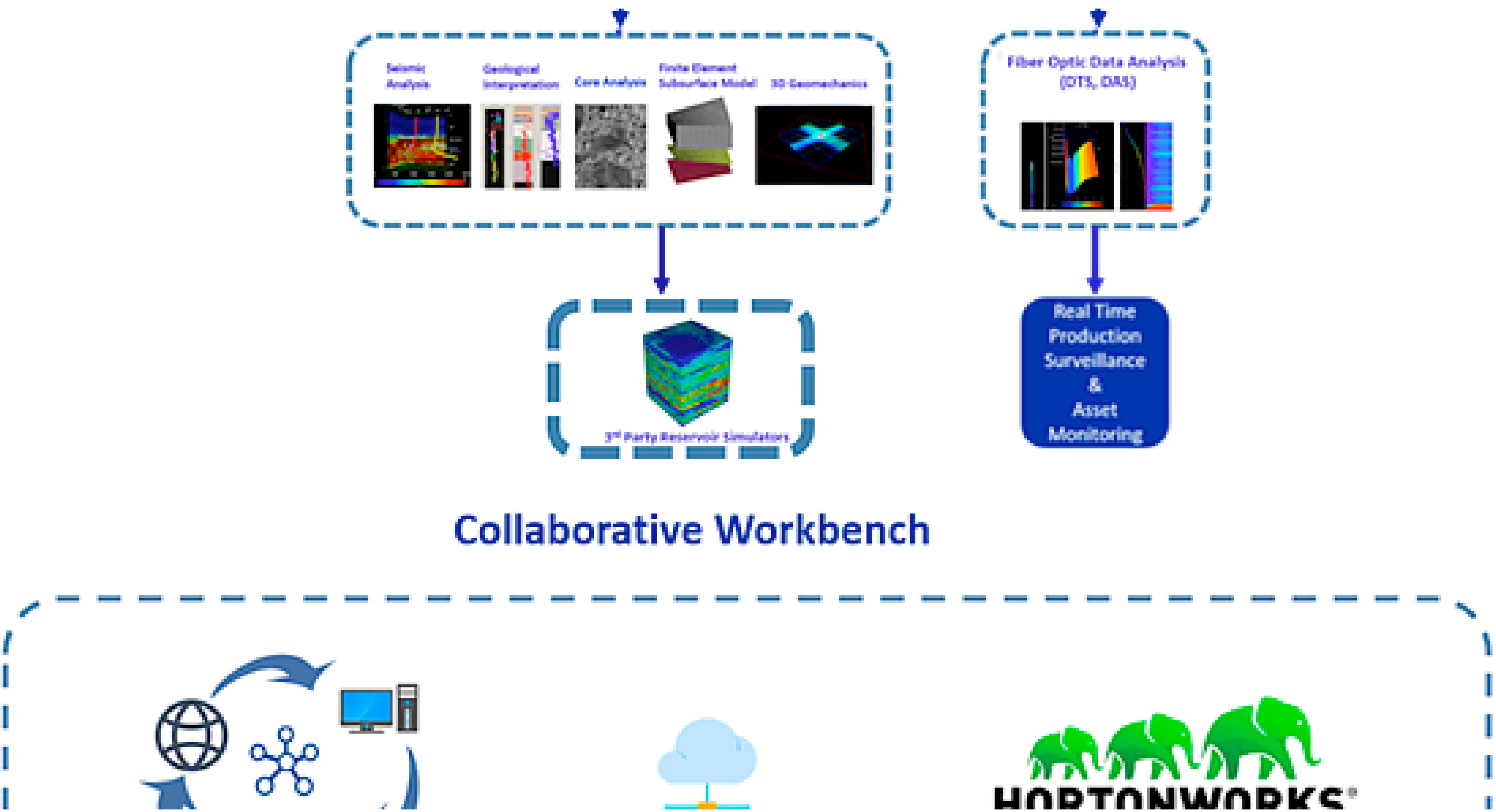




Auto-detecting faults from seismic reflection data helps oil exploration
New Zealand
Paradigm / Nervana project



DIY Deep Learning with Petrabytes



Petrabytes

- With Petrabytes workbench, operators can develop integrated field models using various datasets - seismic, well logs, drilling data, microseismic and sensor measurements.
- Collaboration with Repsol – integrated field operations / history-matched reservoir characterization

With Petrabytes workbench, operators can develop integrated field models using various datasets - seismic, well logs, drilling data, microseismic and sensor measurements.
Analyze all oilfield data in the cloud
Smart data processing of several Gigabytes & Terabytes of oilfield measurement Data
3D / 4D Visualization of large datasets.
Analytics throughout the asset life cycle in areas of seismic, drilling, completions and reservoir monitoring.
Fully Integrated with Industry Standards.
Build Workflows on the fly with encryption.
Real Time Enabled.

The limits of your world are bounded only by the limits of your imagination

- *New bold directions (a few ideas)*
- Tie surface geomorphology and geochemistry to subsurface structure (migration pathways? Frontier exploration?
- Overlooked sweet spots / real-time production optimization
- Enormous databases with predictive value (“Pillow” – Zillow for Petroleum?)
- “Smart” proppants – they change size and also surface tension in response to pressure, temperature, lithology, and they “learn” based on history matching
- “Smart” oilfield chemicals