

Machine Learning Deep Dive

Thomas Grabowski

Principal Product Manager for Machine Learning



Agenda

- 1 About Elastic
- 2 Machine Learning background
- 3 Demo
- 4 Q&A



85,000+
Community
Members

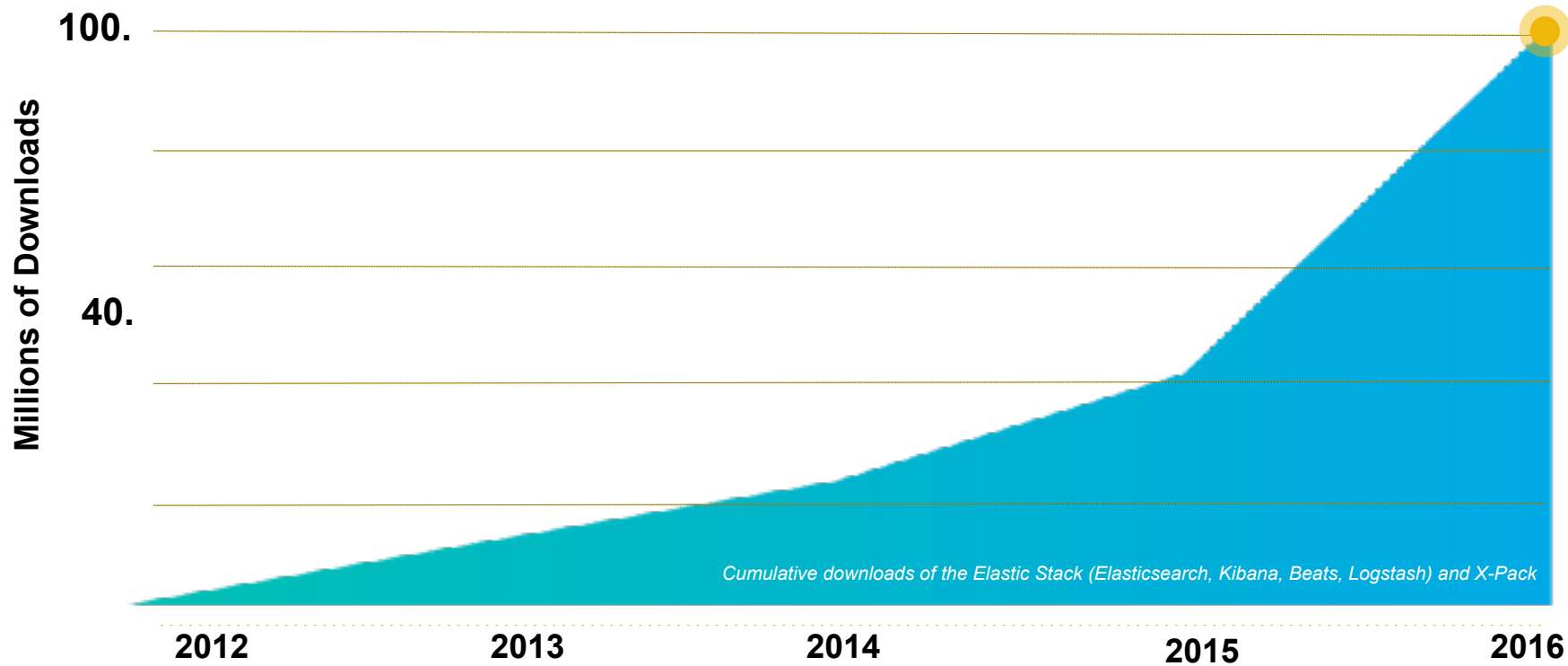


140M+
Product
Downloads



3,000+
Subscription
Customers

Statistics since 2012, founding of Elastic





Horizontal Scale



Real-Time Data
Availability



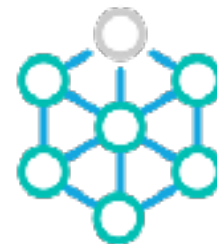
Flexible Data Model



Rapid Query Execution



Sophisticated Query Language



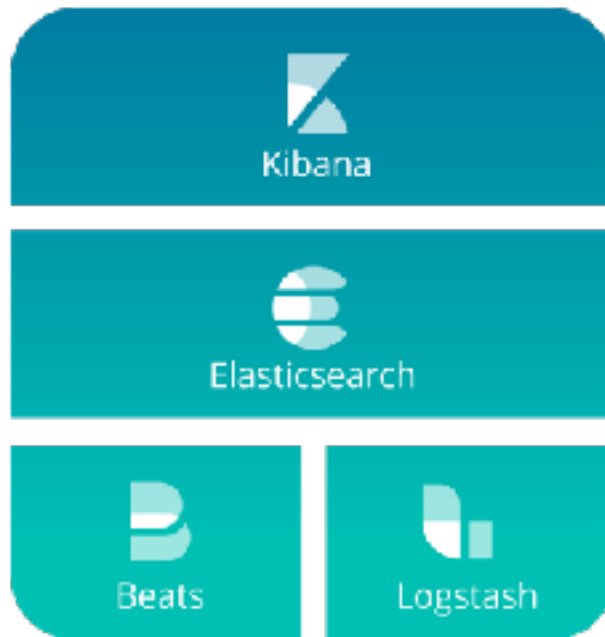
Schemaless

Today's Developer Requirements



Elastic Stack

100% open source
No enterprise edition
All new versions with 5.0





X-Pack

Single install
Extensions for the Elastic Stack
Subscription pricing



Security



Alerting



Monitoring



Reporting



Graph



Machine Learning



Elastic Cloud

Hosted Elasticsearch & Kibana
Includes X-Pack features
Starts at \$45/mo



Available in AWS today and Google Cloud
Machine Learning and Azure platform available soon
Available as a private cloud/on-premise solution
(Elastic Cloud Enterprise)

**75% of our customers
use our products for
multiple use cases**



LOG ANALYTICS



METRICS



SECURITY



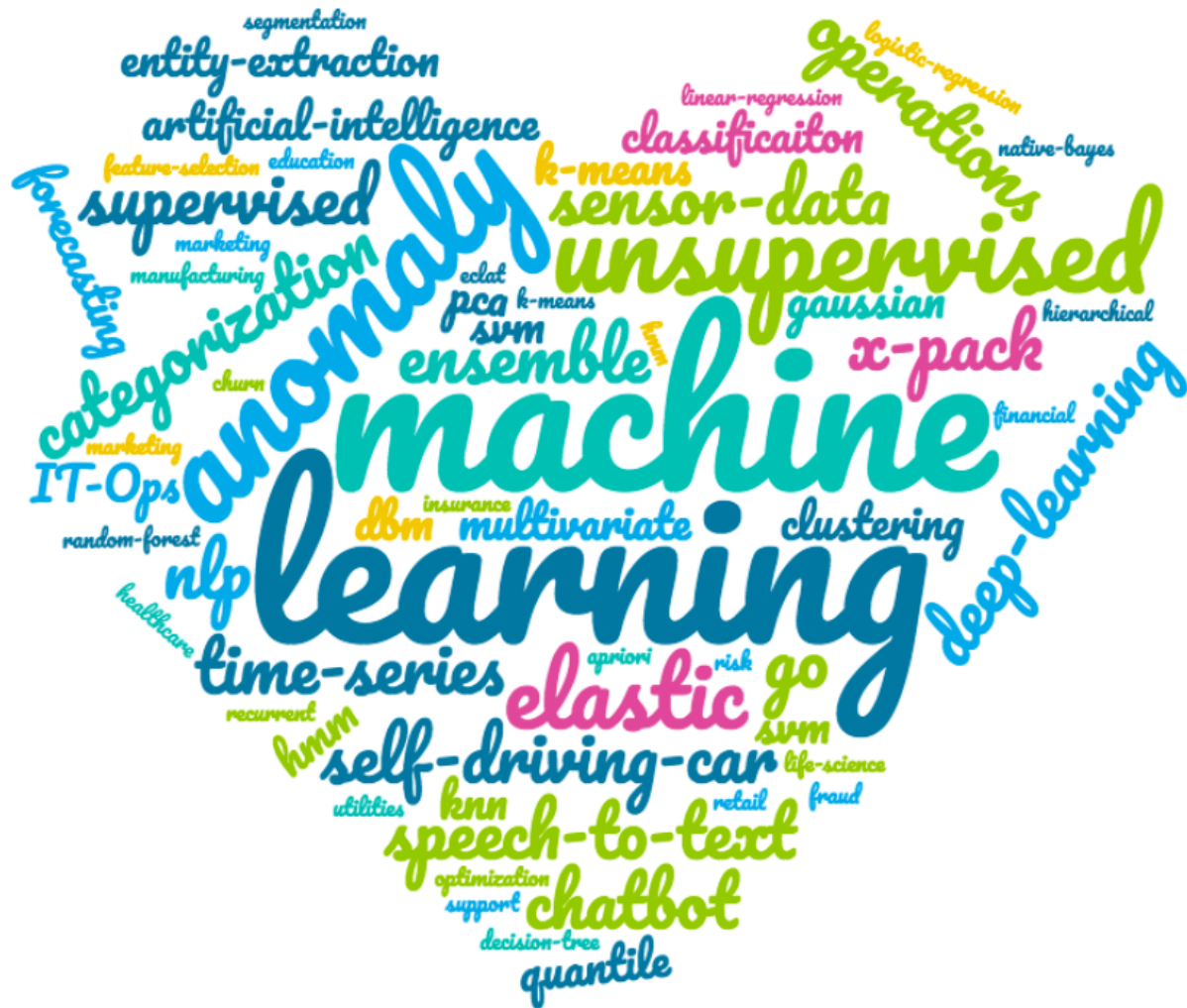
SEARCH



OPERATIONAL
ANALYTICS



CUSTOM APPS



Discover

Visualize

Dashboard

Timelion

Graph

Dev Tools

Monitoring

Management

Guest User

Logout

Collapse

Apache - Total Visitors

11,979,137

Count

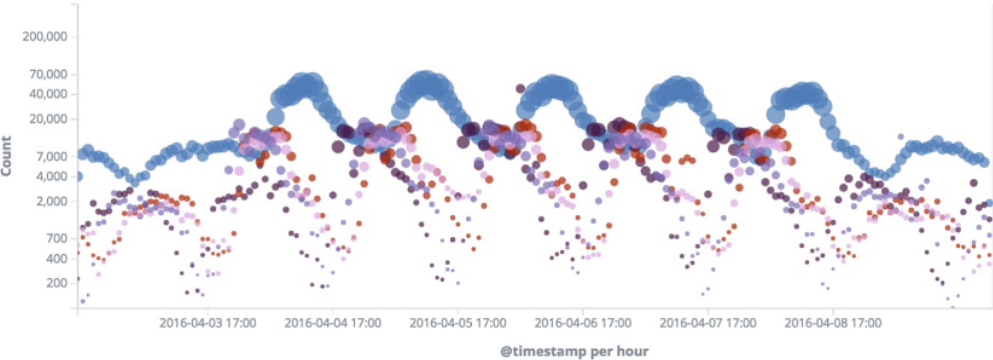
Apache - Unique Visitors

128,088

Unique Visitors

Apache - Bytes and Count


United StatesChinaFranceGermanyIndia



Apache - Country and Status

United StatesChinaFranceGermanyIndia

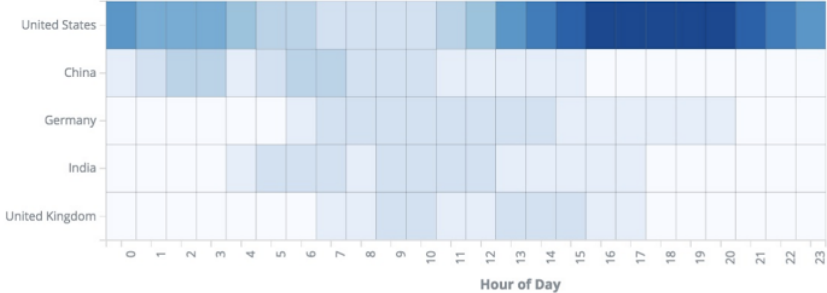
200304301404416206412




Apache - Unique Visitors by City

City	Unique Visitors
Beijing	243,706
Woodbridge	134,791
Seattle	123,545
Seoul	111,663
Mumbai	109,801
Tokyo	106,182
New York	102,947
Bangalore	102,738

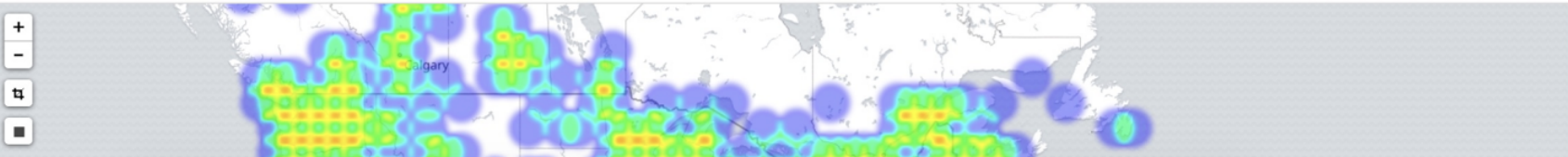
Apache - Country traffic by hour



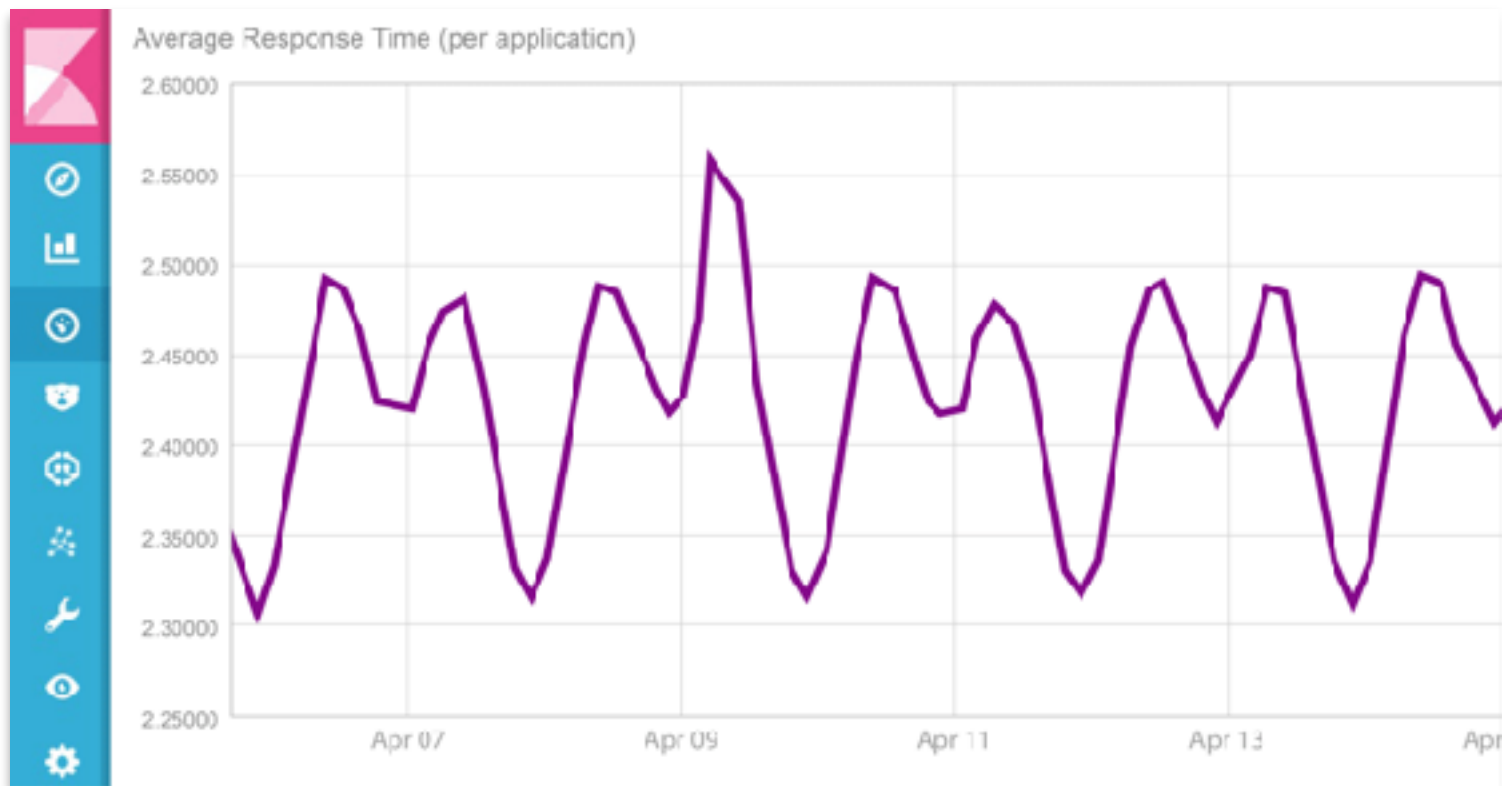
Apache - Top OS



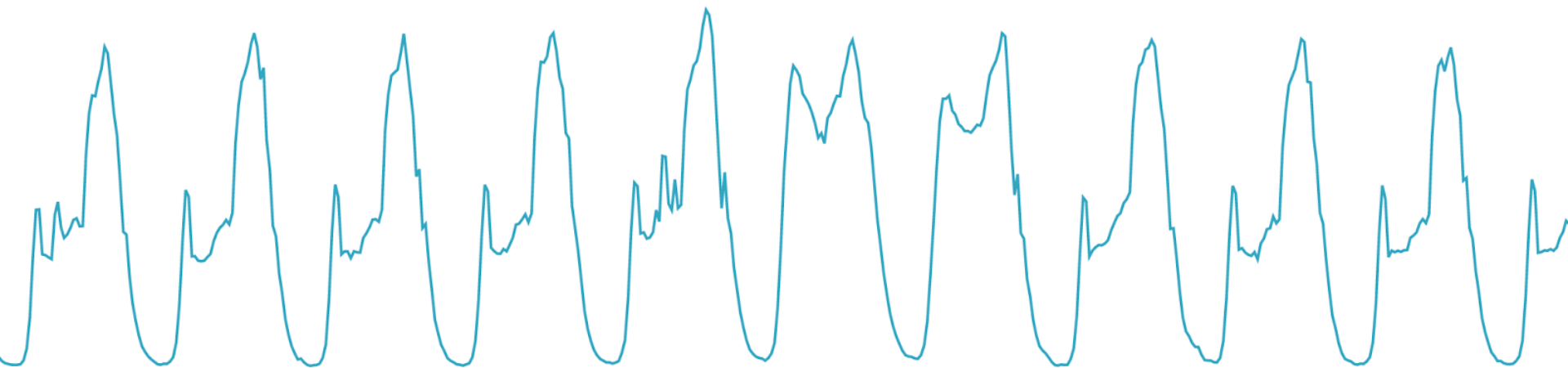
Apache - Visitor Map



Dashboards aren't enough

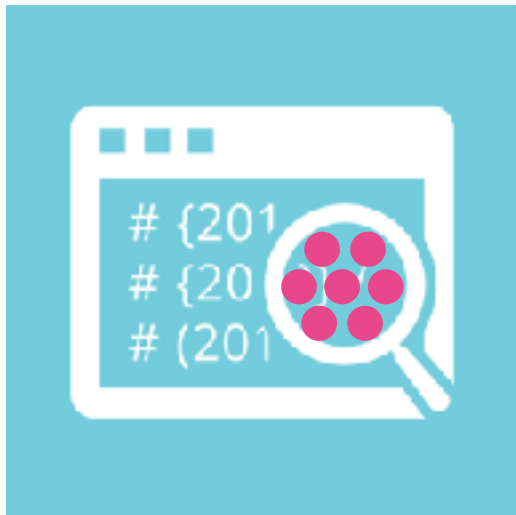


Rules Don't Scale



- Where do you set the threshold?
- Who updates the rules?
- False positives are costly

Extracting useful, valuable information is hard



Search

Aggregations

Visualization

Machine Learning

Machine Learning¹

Algorithms and methods for data driven prediction, decision making, and modeling

Supervised Learning

Prediction based on examples of correct behavior

Unsupervised Learning

No explicit target, only data, goal to model/discover

Semi-supervised Learning

Supplement limited annotations with unsupervised learning

Transfer Learning

How to apply what you have learned from A to B

Active Learning

Learn to query the examples actually needed for learning

Reinforcement Learning

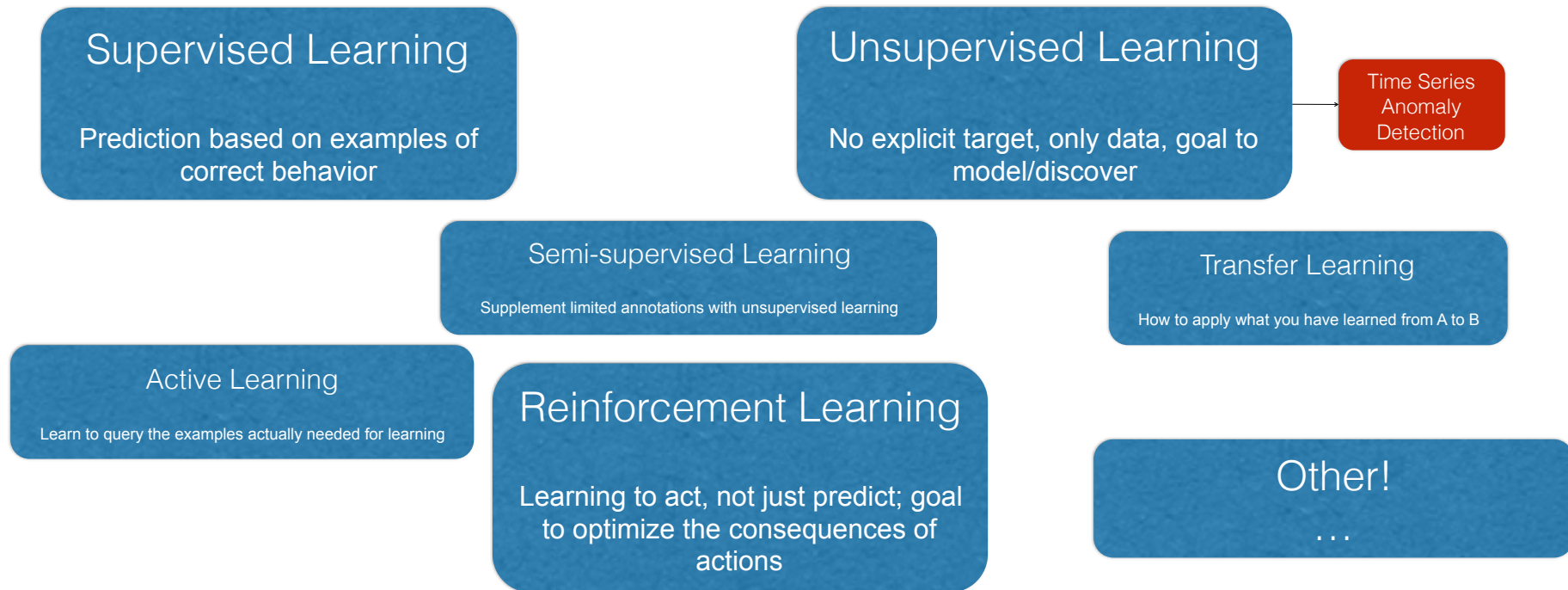
Learning to act, not just predict; goal to optimize the consequences of actions

Other!

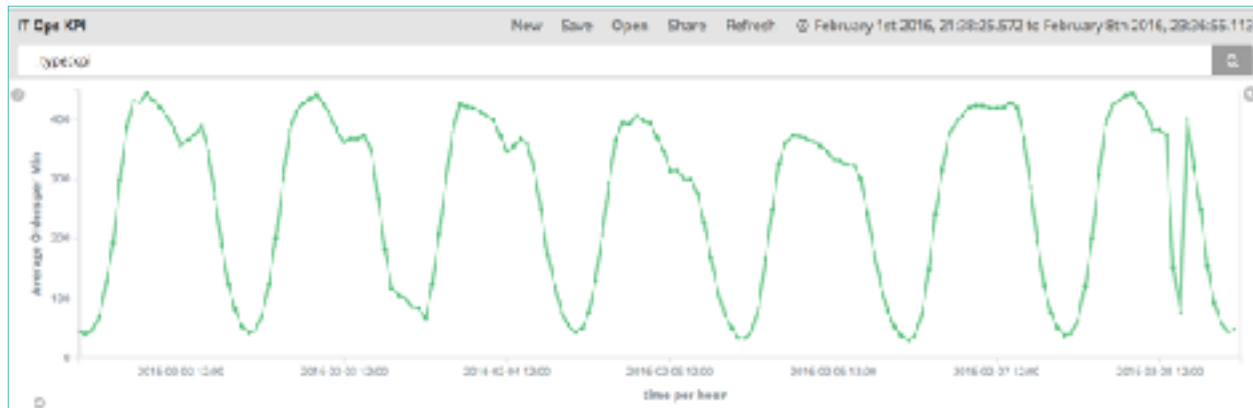
...

Machine Learning¹

Algorithms and methods for data driven prediction, decision making, and modeling

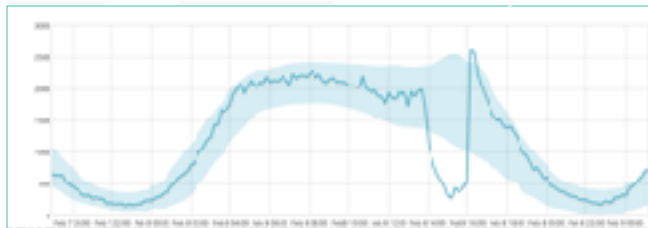
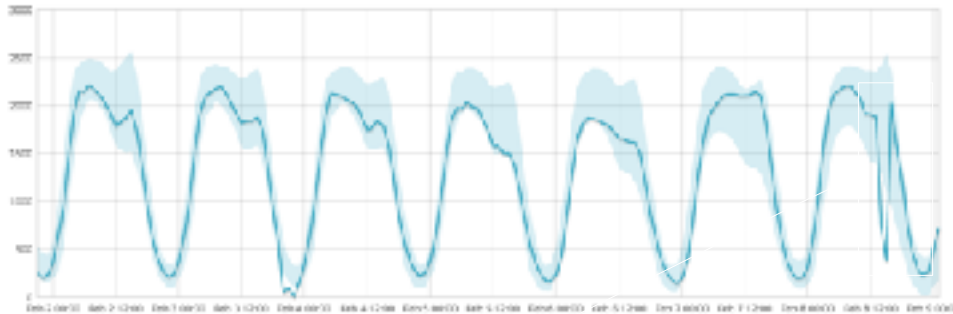


Has my order rate dropped significantly?



Has my order rate dropped significantly?

- Learn models from past behavior (training, modeling)
- Use models to predict future behavior (prediction)
- Use predictions to make decisions



Actual value @ 15:05 = 280

Probability = 0.0000174025



Machine Learning Technical Debt

All the additional plumbing that is required for Machine Learning systems

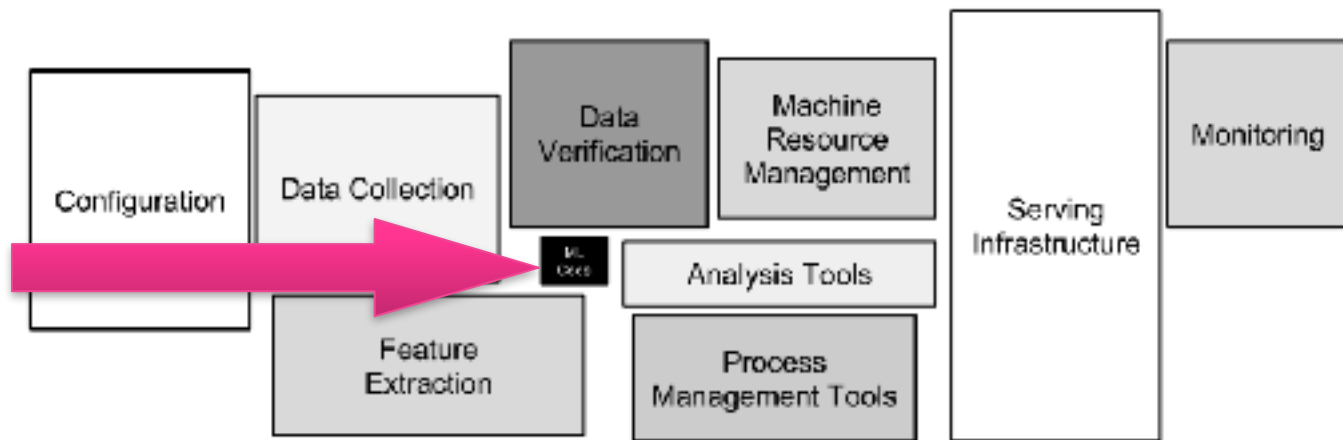


Figure 1: Only a small fraction of real-world ML systems is composed of the ML code, as shown by the small black box in the middle. The required surrounding infrastructure is vast and complex.

From Google Paper: Scully, D., et al. "Machine learning - The high-interest credit card of technical debt." (2014)

It All Begins with Data

Discovering information in NGINX logs

```
68.75.44.178, 172.68.146.54, 127.0.0.1 - - [15/May/  
2017:12:16:27 +0200] "GET /sites/default/files/styles/  
company_profile_cover_crop/public/1500x500_1_10.jpg?  
itok=RUgim2UQ&sc=297009042628d7de3f0eb50e807d29e4 HTTP/1.1" 200  
92763 "https://www.startus.cc/company/finleap" "Mozilla/5.0  
(Windows NT 6.1; Win64; x64) AppleWebKit/537.36 (KHTML, like  
Gecko) Chrome/58.0.3029.110 Safari/537.36"
```

68.75.44.178, 172.68.146.54, 127.0.0.1 - - [15/May/2017:12:16:27 +0200] "GET /sites/default/files/styles/
company_profile_cover_crop/public/1500x500_1_10.jpg?itok=RUgim2UQ&sc=297009042628d7de3f0eb50e807d29e4 HTTP/1.1" 200
92763 "https://www.startus.cc/company/finleap" "Mozilla/5.0 (Windows NT 6.1; Win64; x64) AppleWebKit/537.36 (KHTML,
like Gecko) Chrome/58.0.3029.110 Safari/537.36"

221.247.242.171, 162.158.166.51, 127.0.0.1 - - [15/May/2017:12:16:27 +0200] "GET /sites/default/files/styles/
company_profile_logo/public/company_logos/aaaaaagaaaaaaaawvaaajdk3n2vkzme0lte0zjctngy3mslinmm4lta4ntnhzqwqymzvmog.png?
itok=H2B05xX0 HTTP/1.1" 200 9296 "https://www.startus.cc/company/finleap" "Mozilla/5.0 (Windows NT 6.1; Win64; x64)
AppleWebKit/537.36 (KHTML, like Gecko) Chrome/58.0.3029.110 Safari/537.36"

192.228.32.190, 108.162.246.21, 127.0.0.1 - - [15/May/2017:12:16:27 +0200] "GET /jobs/24237/it-back-end HTTP/1.1" 301 5
-" "Mozilla/5.0 (compatible; Googlebot/2.1; +http://68.75.44.178, 172.68.146.54, 127.0.0.1 - - [15/May/2017:12:16:27
+0200] "GET /sites/default/files/styles/company_profile_cover_crop/public/1500x500_1_10.jpg?
itok=RUgim2UQ&sc=297009042628d7de3f0eb50e807d29e4 HTTP/1.1" 200 92763 "https://www.startus.cc/company/finleap"
"Mozilla/5.0 (Windows NT 6.1; Win64; x64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/58.0.3029.110 Safari/537.36"

221.247.242.171, 162.158.166.51, 127.0.0.1 - - [15/May/2017:12:16:27 +0200] "GET /sites/default/files/styles/
company_profile_logo/public/company_logos/aaaaaagaaaaaaaawvaaajdk3n2vkzme0lte0zjctngy3mslinmm4lta4ntnhzqwqymzvmog.png?
itok=H2B05xX0 HTTP/1.1" 200 9296 "https://www.startus.cc/company/finleap" "Mozilla/5.0 (Windows NT 6.1; Win64; x64)
AppleWebKit/537.36 (KHTML, like Gecko) Chrome/58.0.3029.110 Safari/537.36"

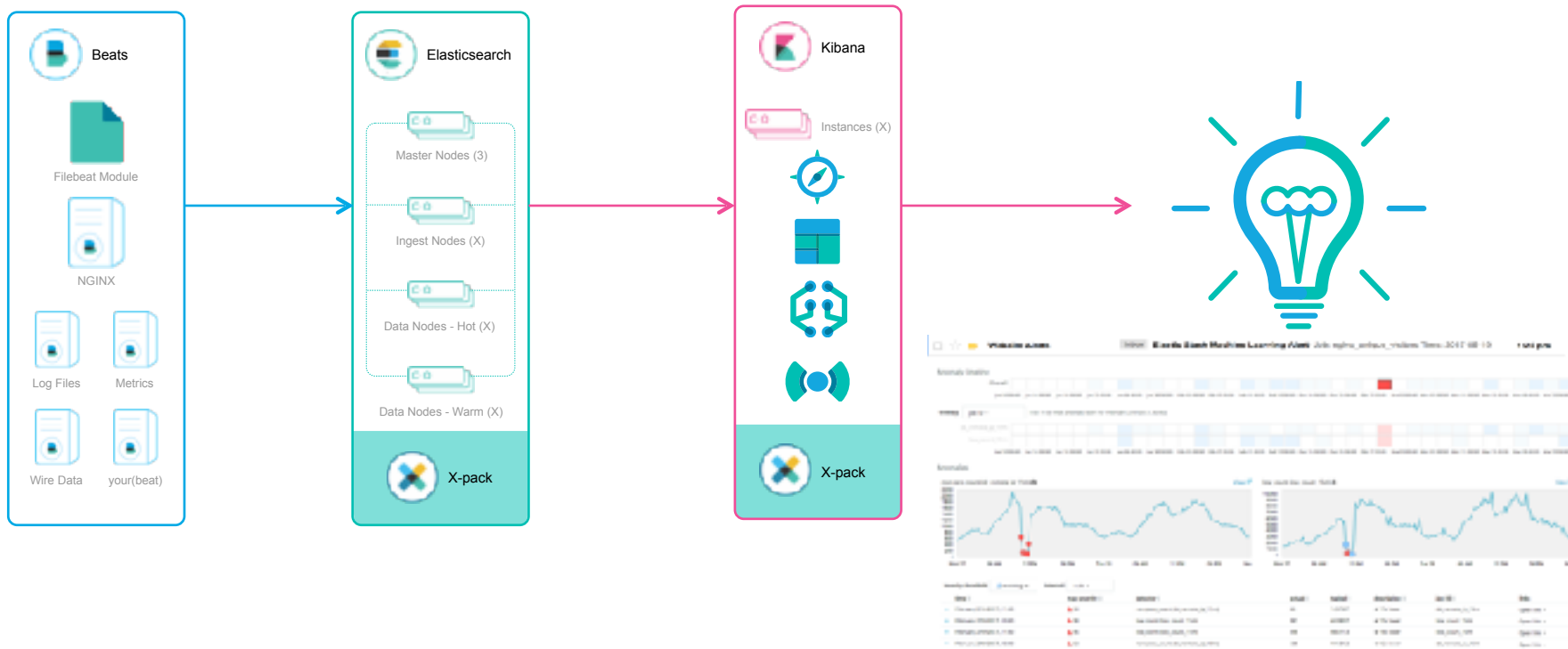
192.228.32.190, 108.162.246.21, 127.0.0.1 - - [15/May/2017:12:16:27 +0200] "GET /jobs/24237/it-back-end HTTP/1.1" 301 5
-" "Mozilla/5.0 (compatible; Googlebot/2.1; +http://www.google.com/bot.html)"

137.56.184.63, 162.158.165.50, 127.0.0.1 - - [15/May/2017:12:16:27 +0200] "GET /sites/default/files/styles/
company_profile_cover/public/1500x500_1_10.jpg?itok=1cNqdGYK HTTP/1.1" 200 102268 "https://www.startus.cc/company/
finleap" "Mozilla/5.0 (Windows NT 6.1; Win64; x64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/58.0.3029.110 Safari/
537.36"

92.222.165.172, 162.158.167.202, 127.0.0.1 - - [15/May/2017:12:16:27 +0200] "POST /jstats.php HTTP/1.0" 200 13
"https://www.startus.cc/company/finleap" "Mozilla/5.0 (Windows NT 6.1; Win64; x64) AppleWebKit/537.36 (KHTML, like
Gecko) Chrome/58.0.3029.110 Safari/537.36" 68.75.44.178, 172.68.146.54, 127.0.0.1 - - [15/May/2017:12:16:27 +0200] "GET
/sites/default/files/styles/company_profile_cover_crop/public/1500x500_1_10.jpg?
itok=RUgim2UQ&sc=297009042628d7de3f0eb50e807d29e4 HTTP/1.1" 200 92763 "https://www.startus.cc/company/finleap"
"Mozilla/5.0 (Windows NT 6.1; Win64; x64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/58.0.3029.110 Safari/537.36"

221.247.242.171, 162.158.166.51, 127.0.0.1 - - [15/May/2017:12:16:27 +0200] "GET /sites/default/files/styles/
company_profile_logo/public/company_logos/aaaaaagaaaaaaaawvaaajdk3n2vkzme0lte0zjctngy3mslinmm4lta4ntnhzqwqymzvmog.png?
itok=H2B05xX0 HTTP/1.1" 200 9296 "https://www.startus.cc/company/finleap" "Mozilla/5.0 (Windows NT 6.1; Win64; x64)
AppleWebKit/537.36 (KHTML, like Gecko) Chrome/58.0.3029.110 Safari/537.36"

Ingest, Enrich, Visualize, Analyze, Alert



Getting Started Resources

Try it out yourself: [Download Elastic Stack & X-Pack](#)

Get the product tour in Elastic's [Machine Learning webinar](#)

Try Elastic's Machine Learning video series:

[Single metric jobs](#)

[Multi metric jobs](#)

[Detect outliers in a population](#)

Take Elastic's [Machine Learning training course](#)

Demo Time!

Questions?

