

splunk> .conf2017

Harnessing Robotic Microscopes, AI, And Machine Learning To Increase The Speed Of Drug Discovery

August Allen | Automation Scientist

Ben Miller | Director High Throughput Screening

September 26, 2017 | Washington, DC

Forward-Looking Statements

During the course of this presentation, we may make forward-looking statements regarding future events or the expected performance of the company. We caution you that such statements reflect our current expectations and estimates based on factors currently known to us and that actual events or results could differ materially. For important factors that may cause actual results to differ from those contained in our forward-looking statements, please review our filings with the SEC.

The forward-looking statements made in this presentation are being made as of the time and date of its live presentation. If reviewed after its live presentation, this presentation may not contain current or accurate information. We do not assume any obligation to update any forward looking statements we may make. In addition, any information about our roadmap outlines our general product direction and is subject to change at any time without notice. It is for informational purposes only and shall not be incorporated into any contract or other commitment. Splunk undertakes no obligation either to develop the features or functionality described or to include any such feature or functionality in a future release.

Splunk, Splunk>, Listen to Your Data, The Engine for Machine Data, Splunk Cloud, Splunk Light and SPL are trademarks and registered trademarks of Splunk Inc. in the United States and other countries. All other brand names, product names, or trademarks belong to their respective owners. © 2017 Splunk Inc. All rights reserved.

What we'll be talking about

- Mission: Accelerate Drug Discovery
- Process: Robotics, Deep Learning, AI

- Scaling and improving existing processes
- Implementing new procedures and automation

- ETL
- Manufacturing Metrics Tracking

August Allen

► Work Experience

- Joined Recursion in June 2016
- Johnson and Johnson Co-op 2015
- Bristol Myers-Squibb Co-op 2013

► Splunk

- Dec 2016-Present

► Education

- Rochester Institute of Technology
 - B.S. in Biomedical Engineering

► Recreation

- Soccer, snowboarding, biking
- Co-host podcast on space exploration



<https://www.linkedin.com/in/august-allen-a8239841>

Ben Miller

► Work:

- Recursion: Director of High Throughput Science
- Myriad Genetics
- Velocity11
- GlycoFi

► Splunk:

- Dec 2014-Present
- 2016 Innovation award

► School:

- Dartmouth College: MS Genetic Engineering
- BA Engineering

► Fun:

- Biking, Skiing, Hiking, Climbing



[linkedin.com/in/bentonjmiller](https://www.linkedin.com/in/bentonjmiller)

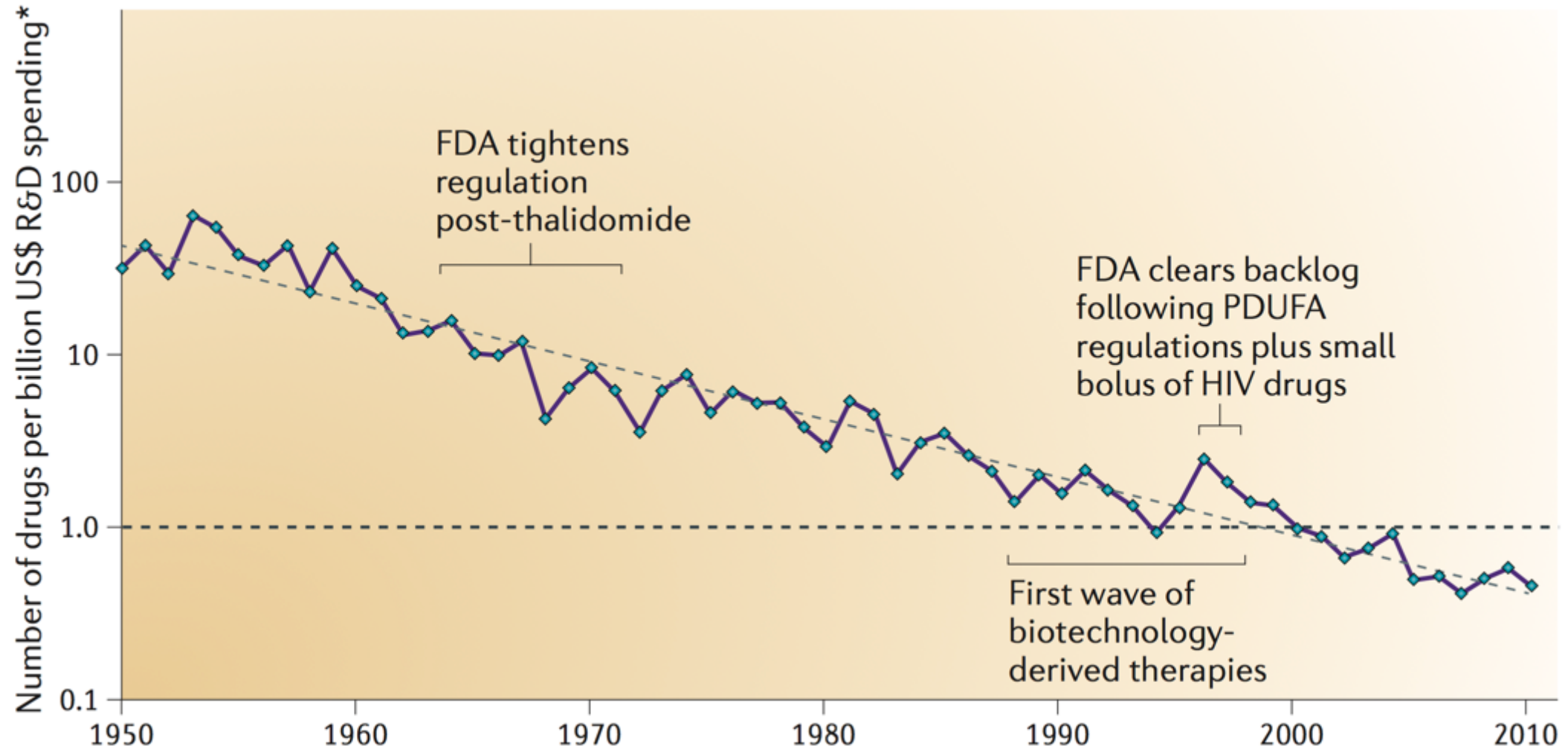


Recursion Pharmaceuticals

Our Mission and Process

Our Mission

Reverse Eroom's Law of Drug Development



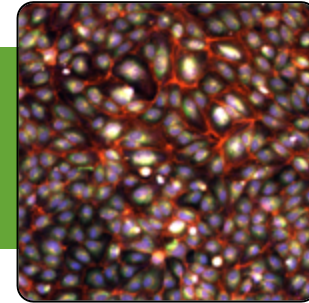
Scannell, et al. Diagnosing the decline in pharmaceutical R&D efficiency. Nature Reviews Drug Discovery, 2012

Recursion Technology

Turning cell biology into a facial recognition problem



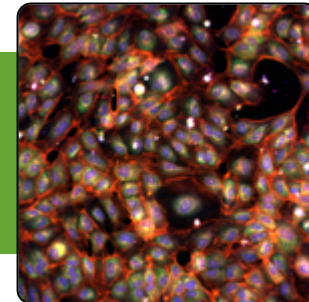
Healthy child



Healthy cells



**Child with rare
genetic disease**
(Cornelia de Lange Syndrome)



**Genetic disease
model cells**
(Cornelia de Lange Syndrome)

Automated high-dimensional feature extraction

Computer vision and machine learning enable the efficient extraction and meaningful representation of the data

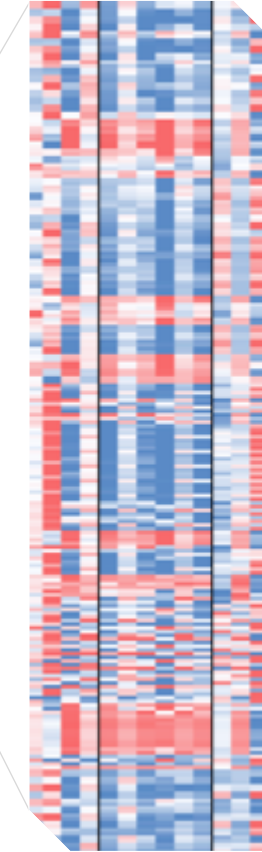
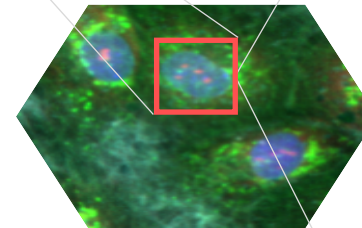
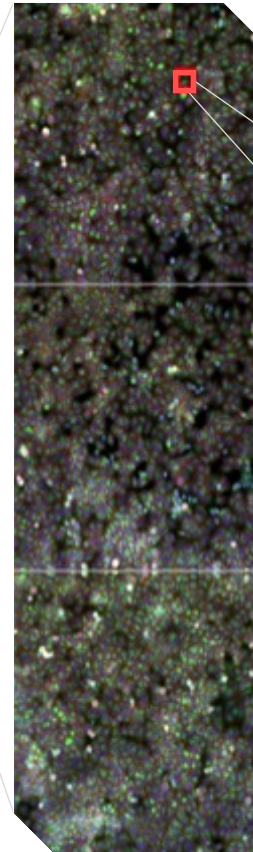
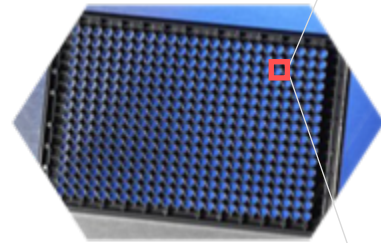
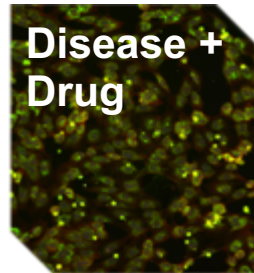
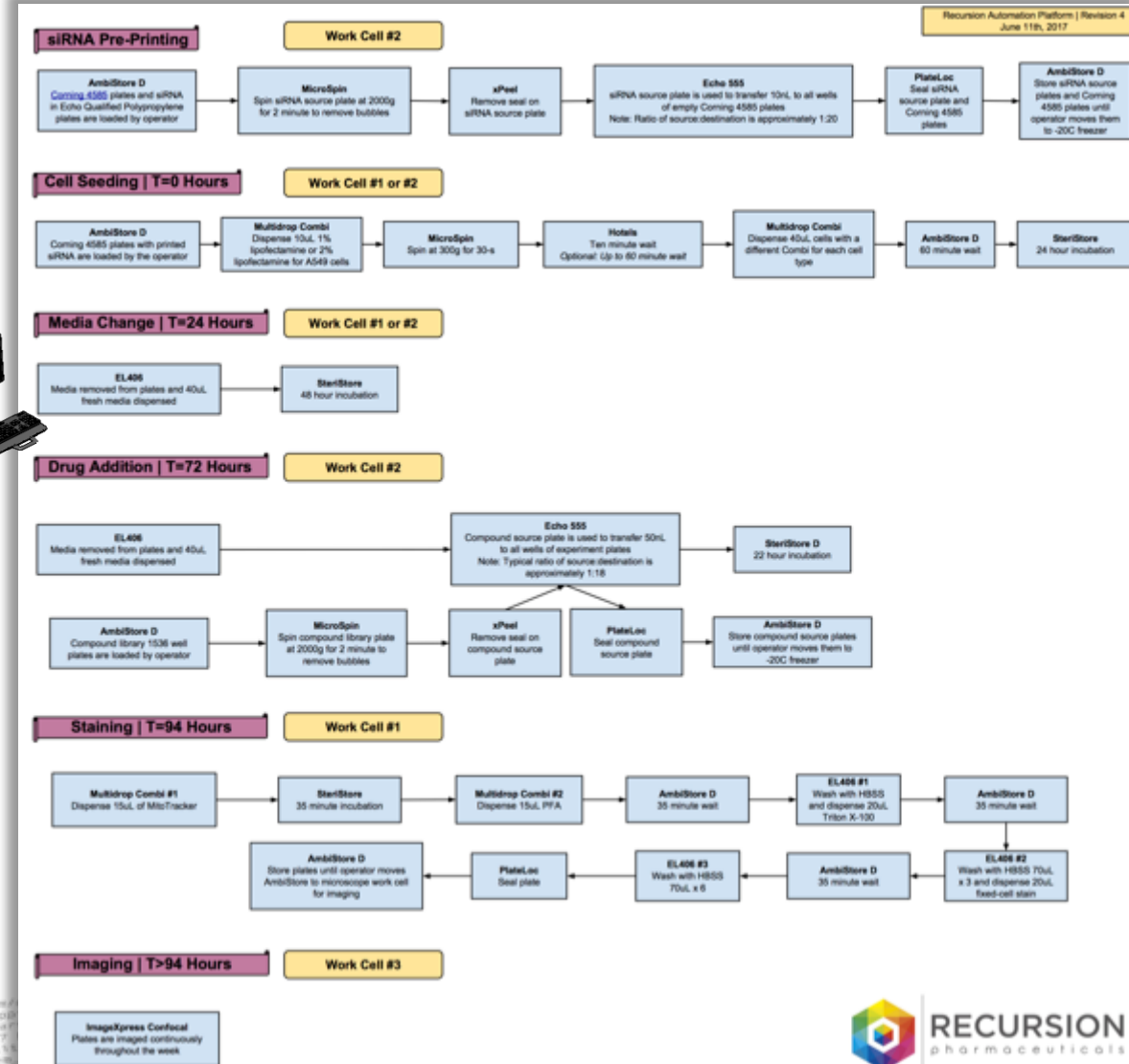
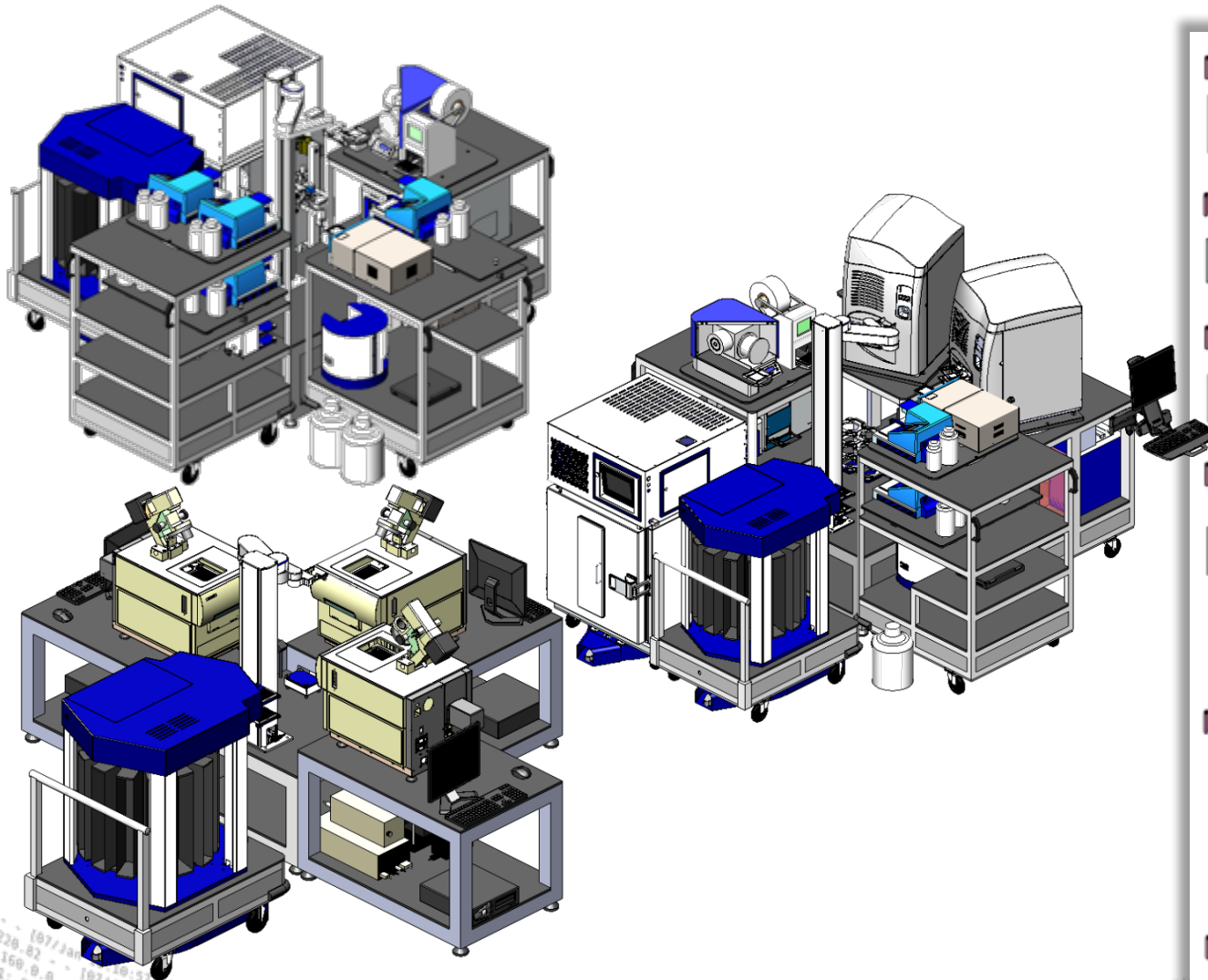


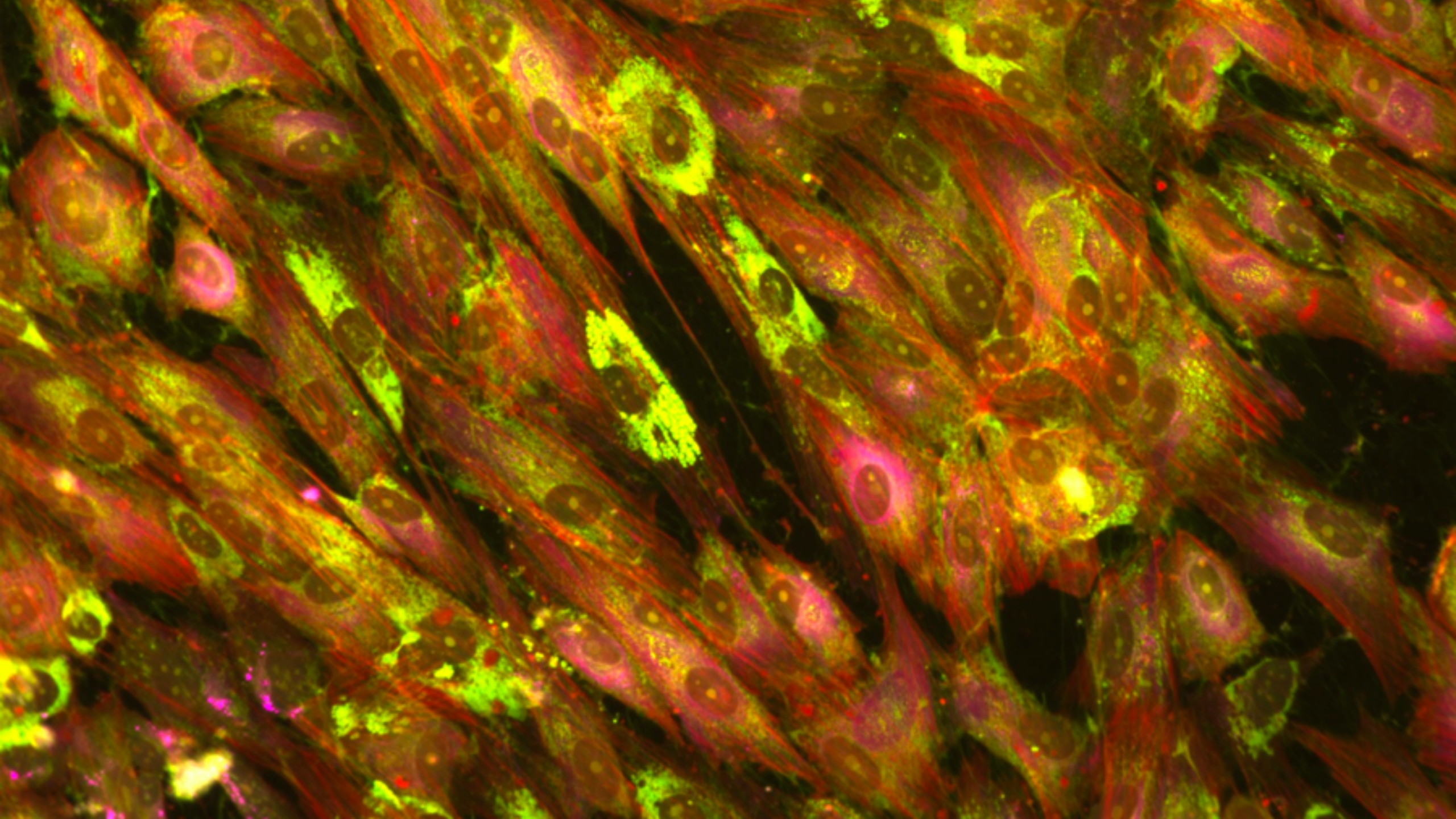
Diagram illustrating the relationship between the five features and the five cell types:

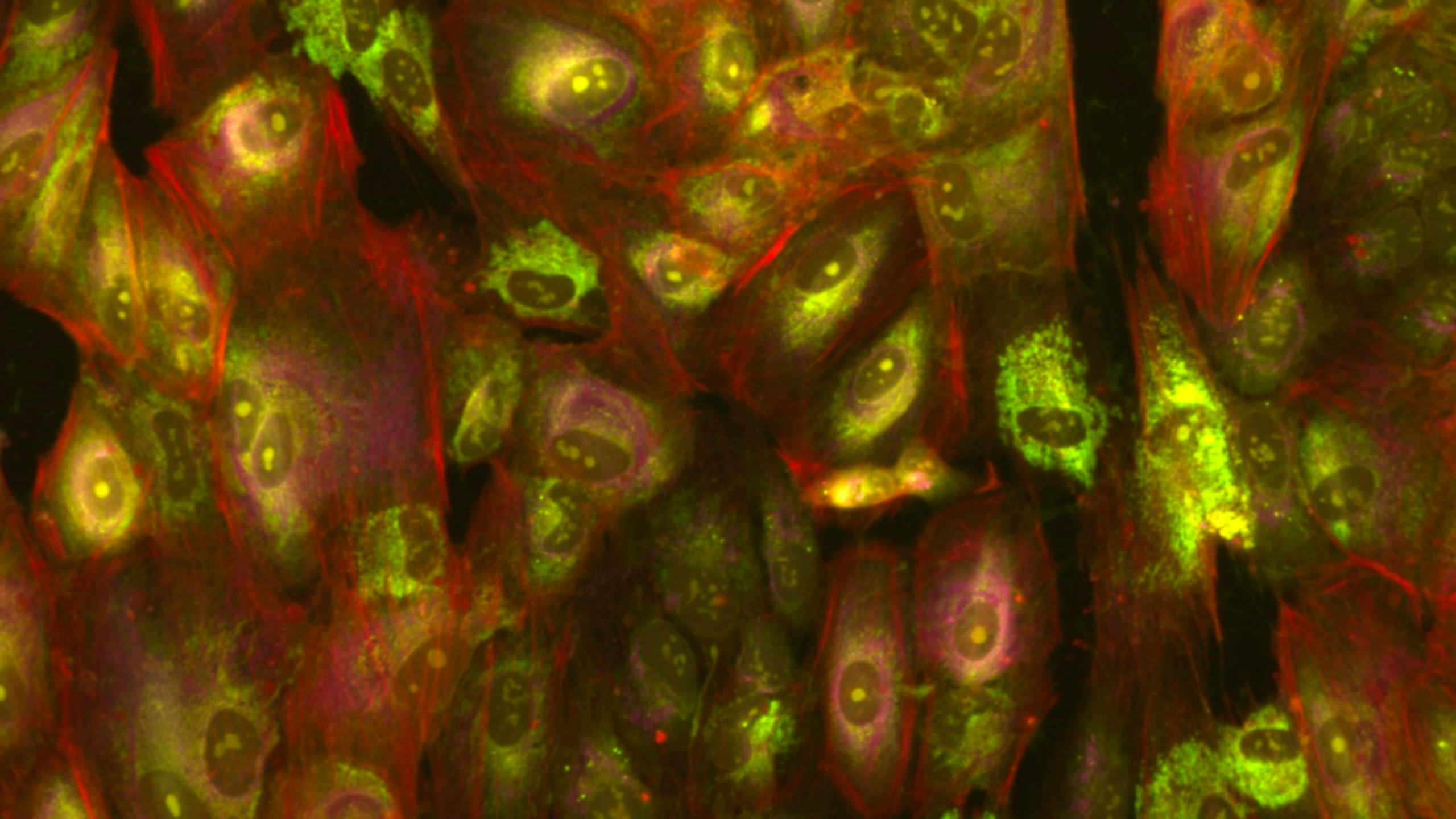
- Nuclear Texture
- Nuclear Radius
- ER Intensity
- RNA Intensity
- Mitochondria

Generating Massive Data Sets

Automated biology leveraging advanced robotics







Our Vision



SHORT TERM

RARE GENETIC DISEASE

100+ genetic disease treatments in 10 years



INTERMEDIATE TERM

EXPAND FOCUS

Disrupt drug discovery across new disease areas and new applications



VISION

SYSTEMS BIOLOGY

Leverage technology to build a map of human biology

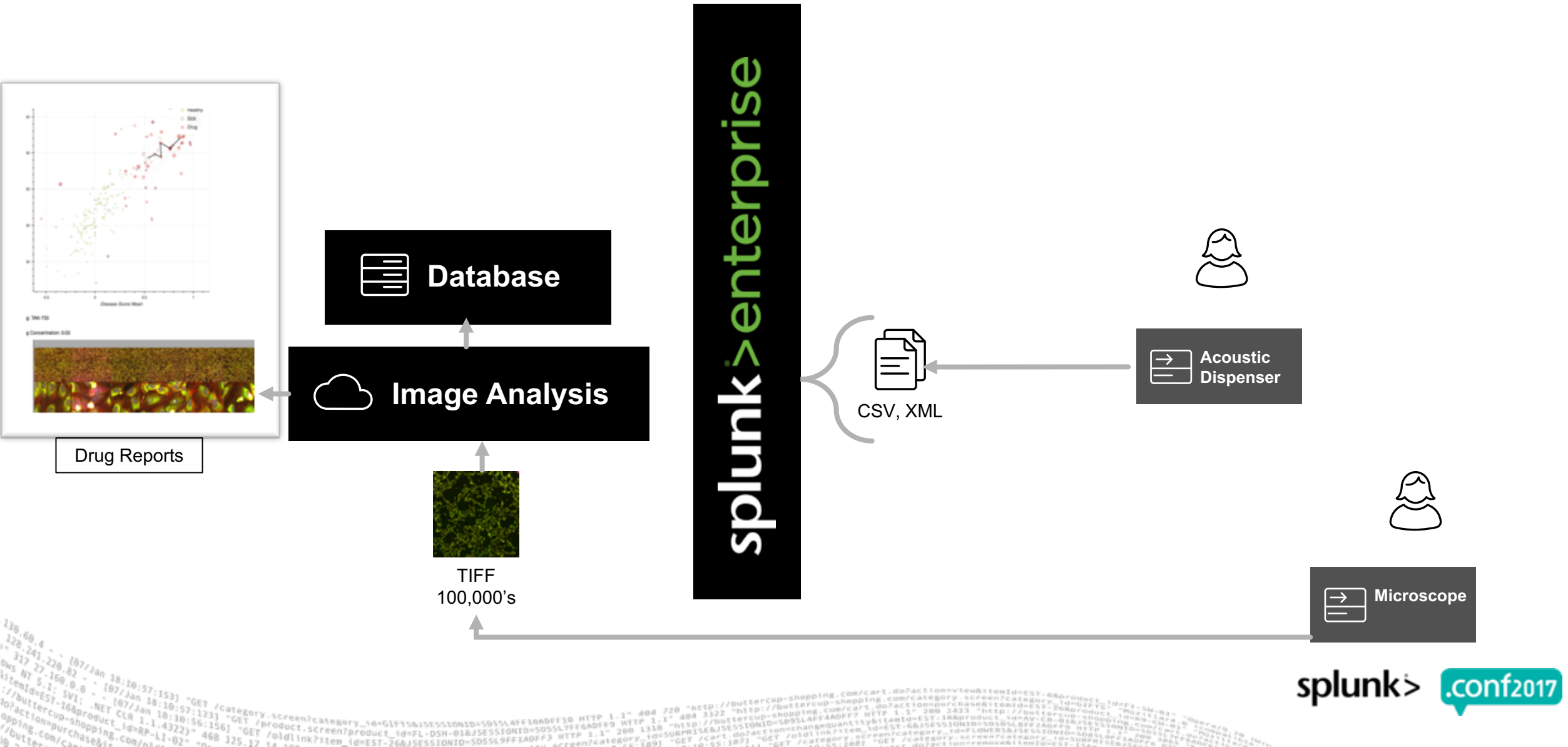
Recursion Pharmaceuticals

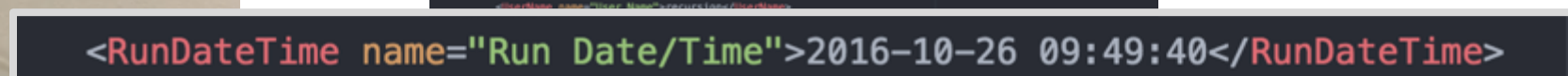
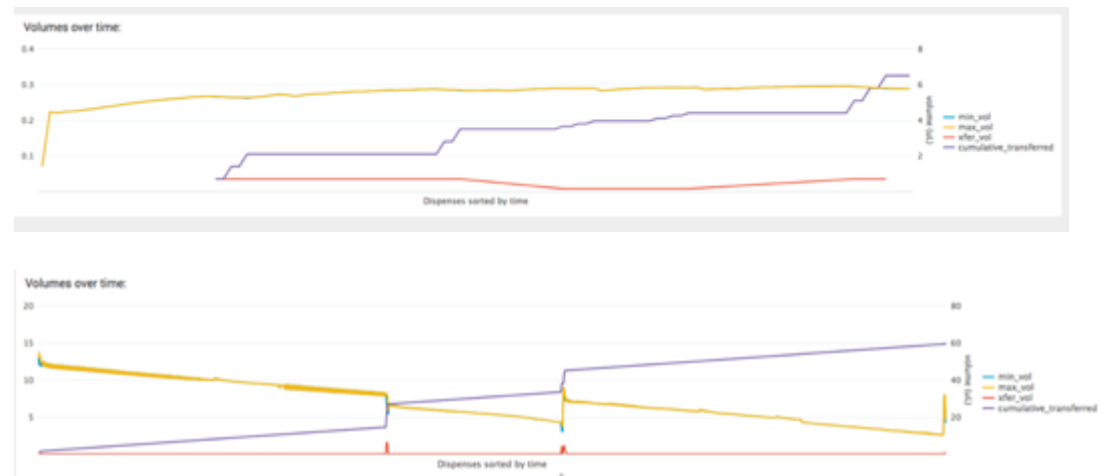
Challenges of Rapid Growth and the Evolution of a
Splunk implementation



Oct 2016: Platform 1.0

- ▶ 34 Plates per week max
- ▶ Limited instrumentation
- ▶ Manual labor





```
<SrcPlateBarcode name= "Source Plate Barcode" />  
<SrcWell name="Source Well">D5</SrcWell>  
DestPlateName name="Destination Plate Name"
```

```
<CurrentFluidVolume name="Current Fluid Volume">11.32</CurrentFluidVolume>
```

```
<DestPlateName name="Destination Plate Name">Destination Plate[1]-01</DestPlateName>
<DestPlateBarcode name="Destination Plate Barcode">
```

All time ▾ 🔍

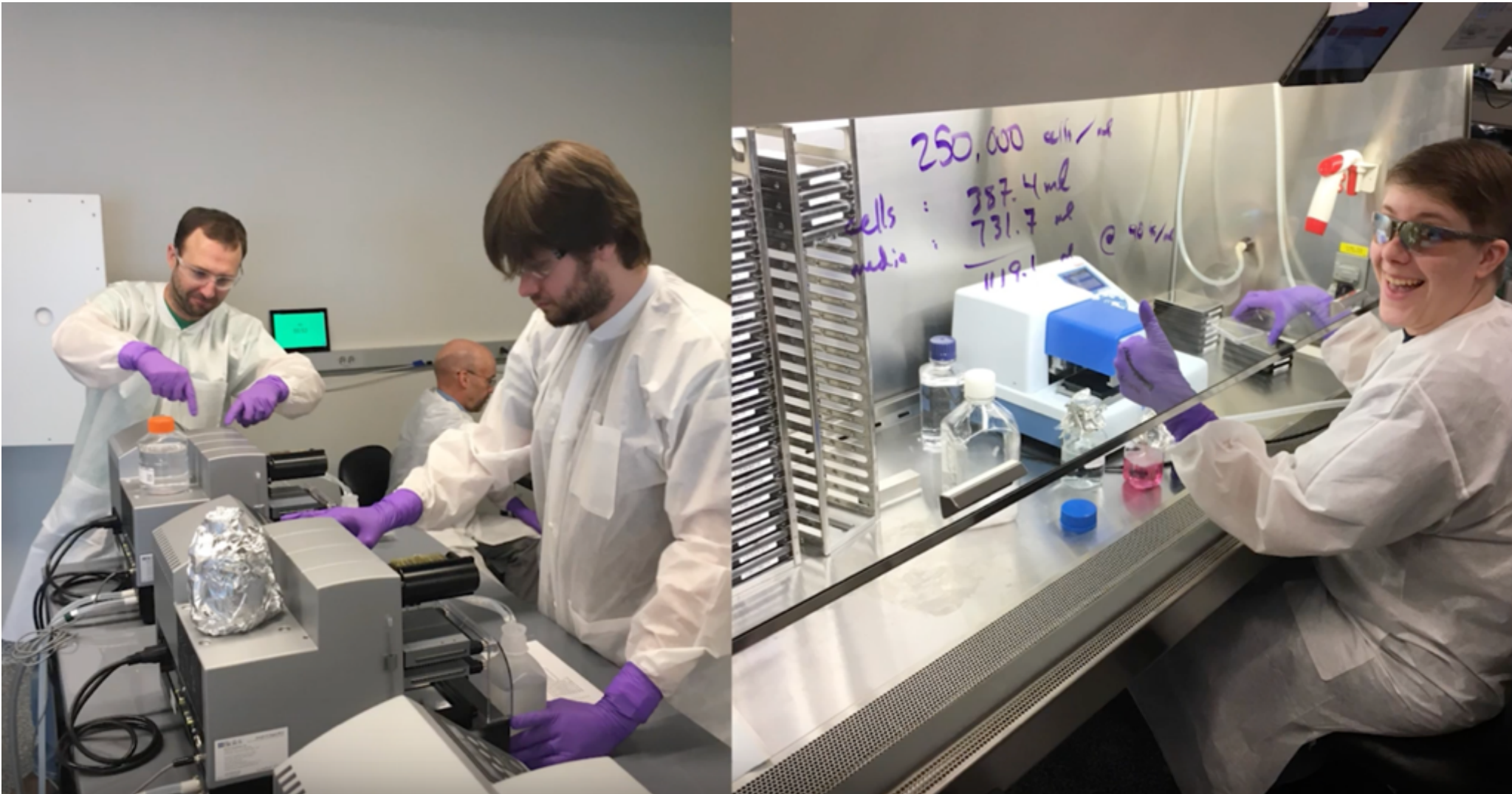
Job ▾ || ■ ↶ 🖨 ⬇️ 🗨 Verbose Mode ▾

Line Chart Format

All time ▾ 🔍

Job Verbose Mode

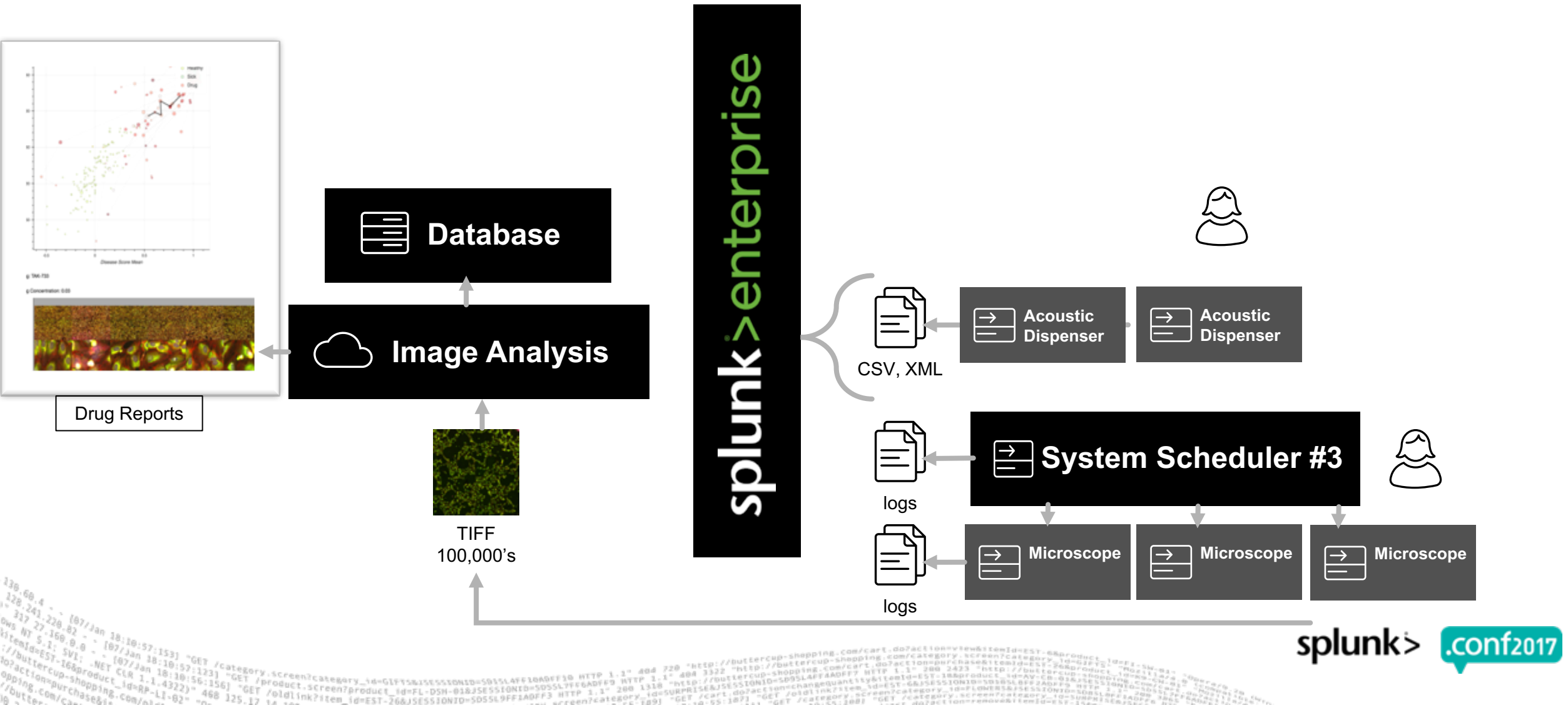
Line Chart Format



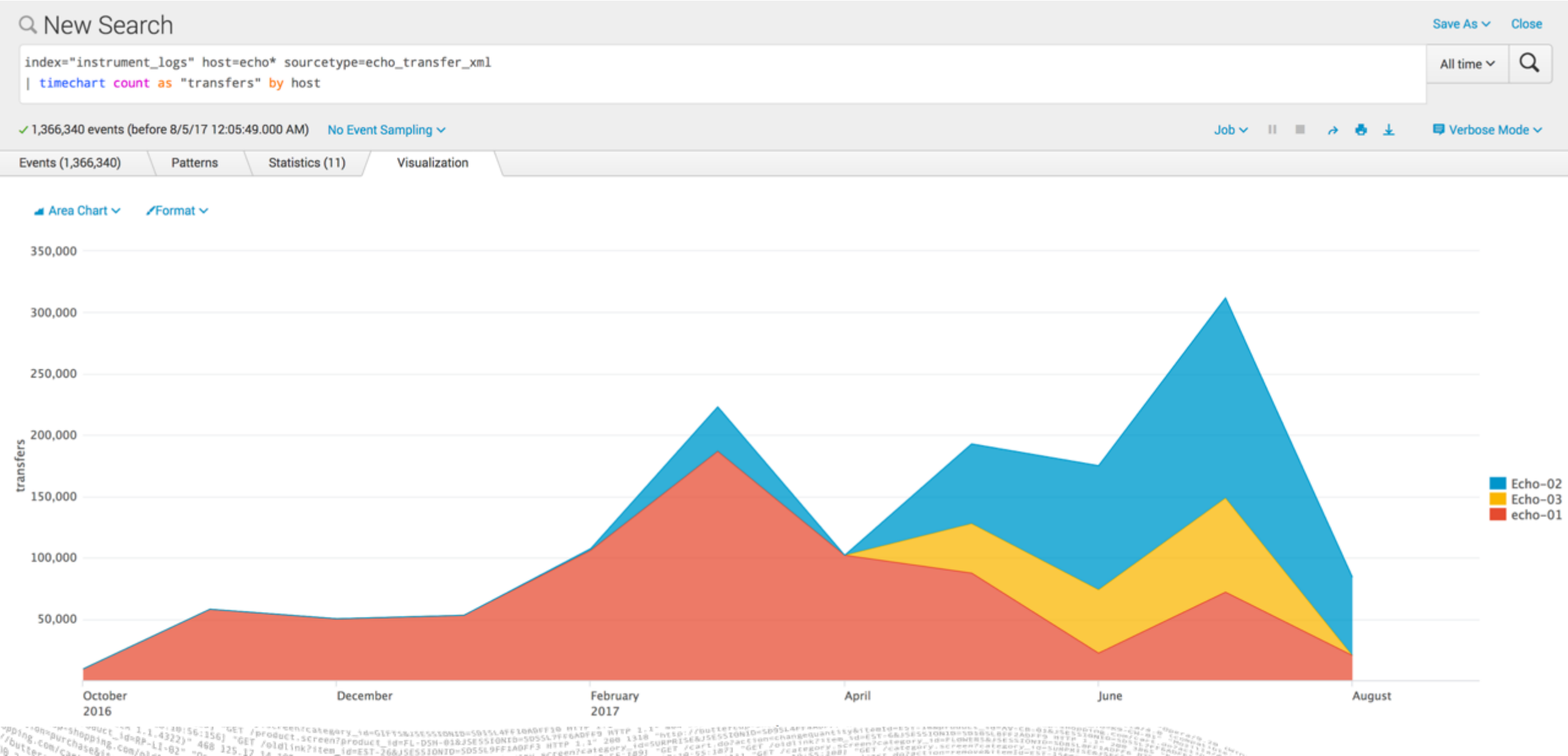
April 2017: Platform 1.5

- ▶ 100 Plates per week max
- ▶ Multiple instruments allow parallel processing
- ▶ Manual processes streamlined

Laboratory Information Management



Acoustic Transfers Performed Over Time By Machine

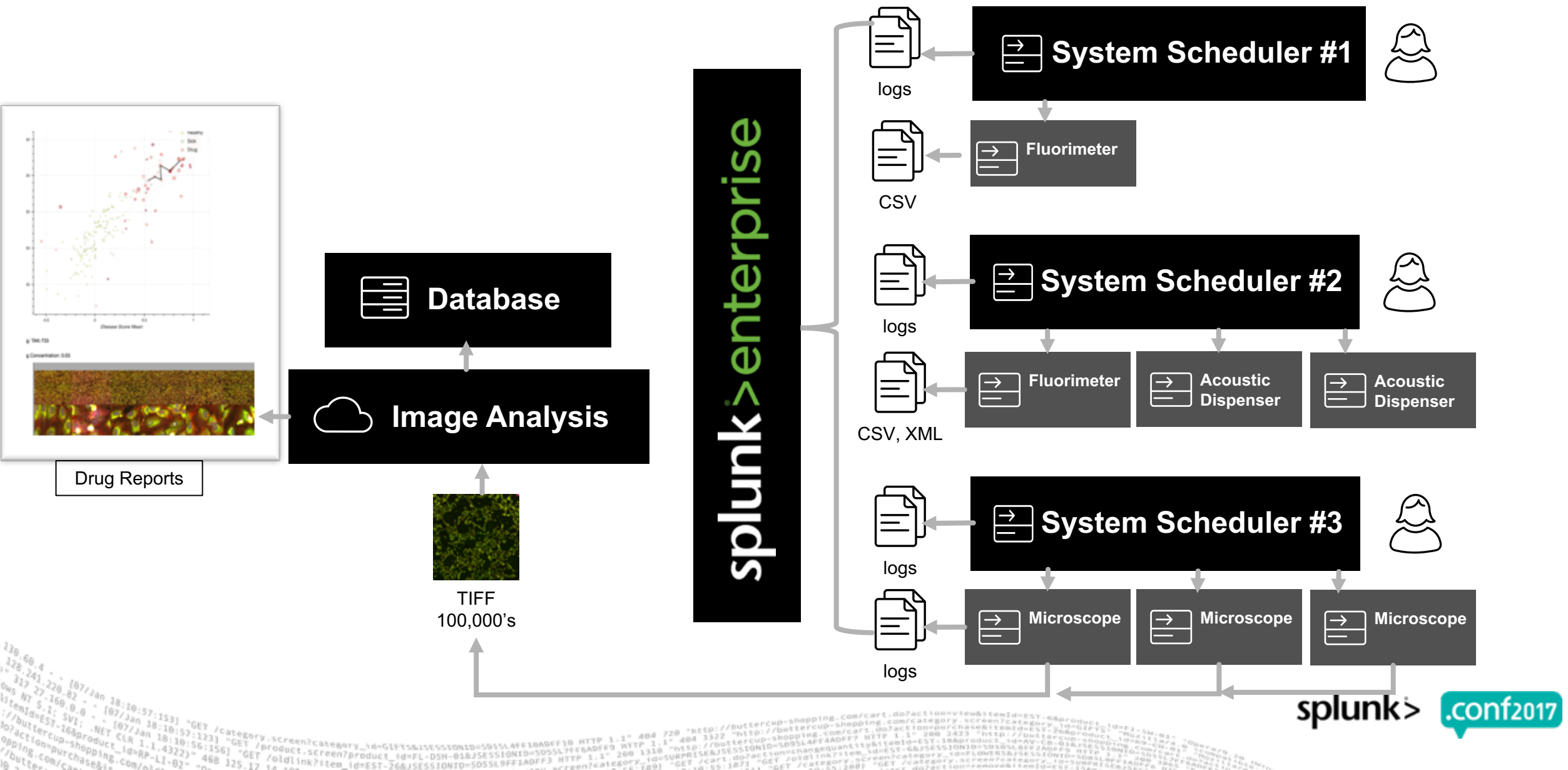


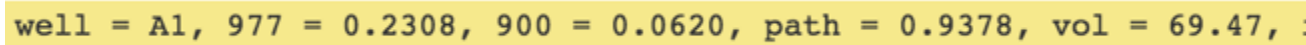


July 2017: Platform 2.0

- ▶ 165 Plates per week Max
- ▶ Instrumentation re-allocated
- ▶ Processes automated

Laboratory Information Management





row	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
A	69.47	71.36	70.21	70.62	70.62	71.19	70.58	70.95	69.34	69.34	69.96	70.95	68.93	69.38	68.93	70.99	69.71	69.88	70.45	71.03	68.60	69.55	70.53	71.07
B	71.19	70.95	70.16	71.15	71.03	71.44	71.52	71.11	70.99	69.86	72.26	75.42	70.37	72.59	68.89	72.88	69.22	70.91	70.58	69.67	70.86	72.67	71.36	71.31
C	69.55	72.22	70.70	71.77	70.49	72.14	70.29	71.98	70.70	72.35	70.33	71.40	70.45	72.59	70.37	72.43	70.00	72.14	70.82	72.02	70.49	71.73	70.58	72.06
D	71.11	70.53	72.22	70.62	70.65	70.86	70.86	71.28	74.86	71.69	71.89	70.08	71.36	70.45	70.37	71.19	72.14	71.36	71.81	71.48	72.55	70.66	72.22	70.29
E	70.08	70.70	70.21	70.95	70.62	71.23	70.91	71.73	70.58	72.39	70.41	72.51	71.15	72.40	70.37	71.98	70.86	72.22	70.91	71.85	70.62	71.73	70.49	71.98
F	70.49	71.28	71.56	71.15	71.32	70.29	71.58	70.58	71.60	71.19	70.45	71.07	72.30	70.91	72.02	69.51	71.73	70.49	71.81	73.54	71.32	69.92	71.23	70.45
G	68.27	72.39	68.77	71.56	70.74	71.69	69.88	72.02	70.04	72.26	70.25	70.45	70.99	72.22	69.55	70.12	68.48	72.14	70.86	71.28	69.55	71.23	70.12	70.58
H	70.99	71.11	72.30	71.03	71.60	74.90	71.28	98.77	70.45	71.93	72.39	70.70	72.43	70.25	71.40	69.96	65.19	66.75	71.93	71.23	71.89	70.37	70.45	70.53
I	70.37	71.36	71.11	72.02	70.53	72.72	70.78	70.35	69.59	72.55	71.23	72.96	70.12	72.76	70.62	72.59	70.25	72.06	71.28	72.30	69.67	71.81	71.32	72.22
J	71.48	71.28	72.22	72.14	72.18	71.77	72.22	70.82	72.88	71.98	72.35	71.44	70.99	70.21	69.05	71.36	71.03	71.32	72.55	71.85	72.06	71.07	71.93	70.49
K	69.63	72.18	70.74	72.67	70.41	71.40	70.74	71.40	70.86	72.18	69.59	72.02	71.07	73.09	69.75	71.81	69.84	72.30	70.49	71.73	71.07	72.84	70.33	72.10
L	69.88	71.85	72.76	70.95	72.18	71.36	72.92	70.95	71.60	71.44	72.51	71.36	71.85	71.81	72.51	71.77	71.52	71.69	70.49	71.73	72.96	71.15	71.56	70.74
M	69.42	71.28	70.21	72.10	69.34	72.22	69.09	71.07	70.16	71.11	69.55	71.65	70.25	71.48	70.95	71.98	70.95	71.40	70.12	72.06	70.41	71.85	70.53	70.86
N	69.63	70.62	72.80	71.02	72.51	69.67	71.44	70.78	70.41	71.07	70.99	71.36	71.28	70.58	71.73	71.44	70.95	70.62	71.81	70.66	70.99	71.19	71.77	68.27
O	70.33	71.15	70.16	72.88	71.03	72.06	70.66	72.22	71.77	70.79	70.58	71.69	71.07	72.63	72.30	72.30	71.73	72.63	71.40	72.76	70.16	73.06	70.99	90.90
P	70.58	71.60	69.88	71.60	71.85	71.40	72.30	72.51	72.92	71.73	72.18	72.47	71.65	70.25	72.35	71.48	71.36	71.77	71.89	71.98	71.73	71.77	72.18	70.74

2017-08-04	08:38:56.9249	TRACE Diagnostic	-well = B1, 977 = 0.2350, 900 = 0.0620, path = 0.9411, vol = 71.19, resource = Combi_02, dispenser = 836-39440M, cassette = 832057609
2017-08-04	08:38:56.9249	INFO Diagnostic	-well = C1, 977 = 0.2302, 900 = 0.0612, path = 0.9389, vol = 69.55, resource = Combi_02, dispenser = 836-39440M, cassette = 832057609
2017-08-04	08:38:56.9409	TRACE Diagnostic	-well = C1, 977 = 0.2302, 900 = 0.0612, path = 0.9389, vol = 69.55, resource = Combi_02, dispenser = 836-39440M, cassette = 832057609
2017-08-04	08:38:56.9409	INFO Diagnostic	-well = D1, 977 = 0.2348, 900 = 0.0620, path = 0.9600, vol = 71.11, resource = Combi_02, dispenser = 836-39440M, cassette = 832057609
2017-08-04	08:38:56.9579	TRACE Diagnostic	-well = D1, 977 = 0.2348, 900 = 0.0620, path = 0.9600, vol = 71.11, resource = Combi_02, dispenser = 836-39440M, cassette = 832057609
2017-08-04	08:38:56.9579	INFO Diagnostic	-well = E1, 977 = 0.2333, 900 = 0.0630, path = 0.9461, vol = 70.08, resource = Combi_02, dispenser = 836-39440M, cassette = 832057609
2017-08-04	08:38:56.9740	TRACE Diagnostic	-well = E1, 977 = 0.2333, 900 = 0.0630, path = 0.9461, vol = 70.08, resource = Combi_02, dispenser = 836-39440M, cassette = 832057609
2017-08-04	08:38:56.9740	INFO Diagnostic	-well = F1, 977 = 0.2317, 900 = 0.0604, path = 0.9517, vol = 70.49, resource = Combi_02, dispenser = 836-39440M, cassette = 832057609
2017-08-04	08:38:56.9910	TRACE Diagnostic	-well = F1, 977 = 0.2317, 900 = 0.0604, path = 0.9517, vol = 70.49, resource = Combi_02, dispenser = 836-39440M, cassette = 832057609
2017-08-04	08:38:56.9910	INFO Diagnostic	-well = G1, 977 = 0.2260, 900 = 0.0601, path = 0.9217, vol = 68.27, resource = Combi_02, dispenser = 836-39440M, cassette = 832057609
2017-08-04	08:38:57.0071	TRACE Diagnostic	-well = G1, 977 = 0.2260, 900 = 0.0601, path = 0.9217, vol = 68.27, resource = Combi_02, dispenser = 836-39440M, cassette = 832057609
2017-08-04	08:38:57.0071	INFO Diagnostic	-well = H1, 977 = 0.2317, 900 = 0.0592, path = 0.9583, vol = 70.99, resource = Combi_02, dispenser = 836-39440M, cassette = 832057609
2017-08-04	08:38:57.0201	TRACE Diagnostic	-well = H1, 977 = 0.2317, 900 = 0.0592, path = 0.9583, vol = 70.99, resource = Combi_02, dispenser = 836-39440M, cassette = 832057609

Save Save As View Close

All time

Job      Verbose Mode 

Area Chart ▾ Format ▾



Save Save As View Close

All time  

 Verbose Mode 

Job ▾ || ■ ↻ 🖨 ⬇️ 🔊 Verbose Mode ▾

Visualization

The line plot displays the count of tweets over time. The x-axis represents time from January 2017 to August 2017, with labels for January 2017, March, May, and July. The y-axis represents the count, ranging from 0 to 7,000 in increments of 1,000. The count starts at approximately 100 in January, rises to a peak of about 6,500 in March, drops to around 4,100 in May, and then declines to approximately 400 by August.

time	count
January 2017	100
February 2017	4000
March 2017	6500
April 2017	6300
May 2017	4100
June 2017	4400
July 2017	2100
August 2017	400

experiment_id

Last 30 days

*

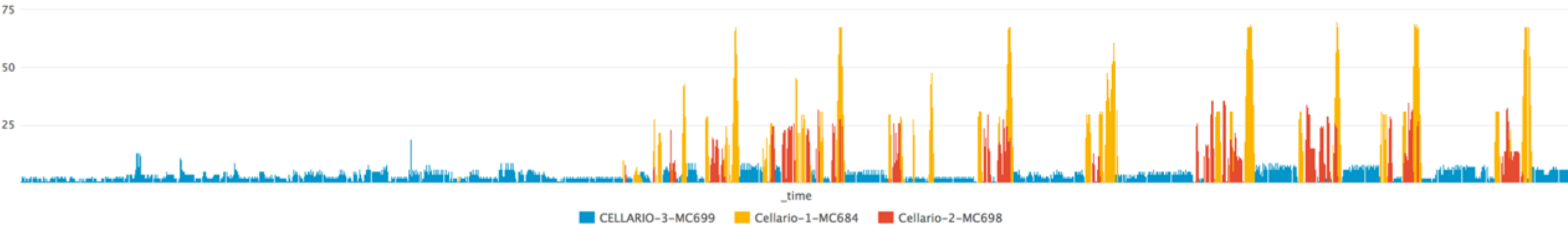
Hide Filters

Summary Stats by Workstation:

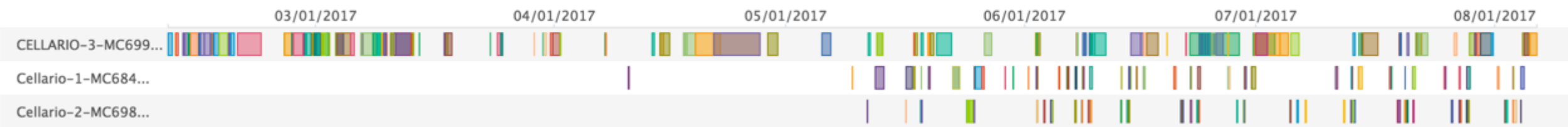
	host	total_plate_orders	total_duration_hr	uptime_percent	experiments	orders	average_plates_per_hour	average_minutes_per_plate
1	CELLARIO-3-MC699	691	11+19:11:11.3238	38.63918490	114	23	2.4400882516	24.589274573
2	Cellario-1-MC684	1435	3+01:10:18.8534	9.983890224	24	30	19.61135254	3.059452420
3	Cellario-2-MC698	1457	4+02:33:35.3197	13.44792583	35	34	14.78290172	4.058743076

All Time:

Distinct count of barcodes by host per hour:



Timeline view of Cellario orders by workstation:



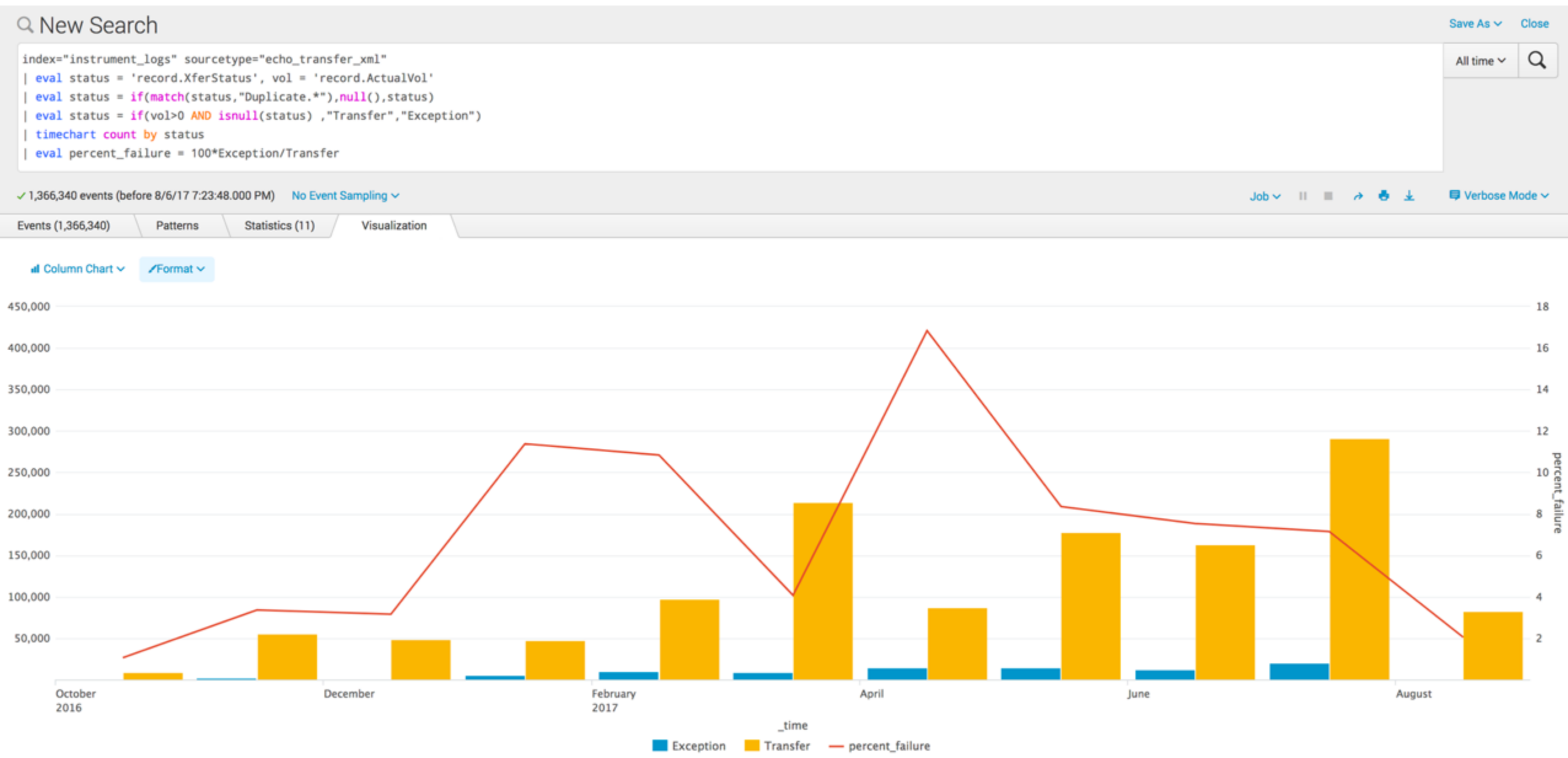
Save Save As View Close

All time ▾ 🔍

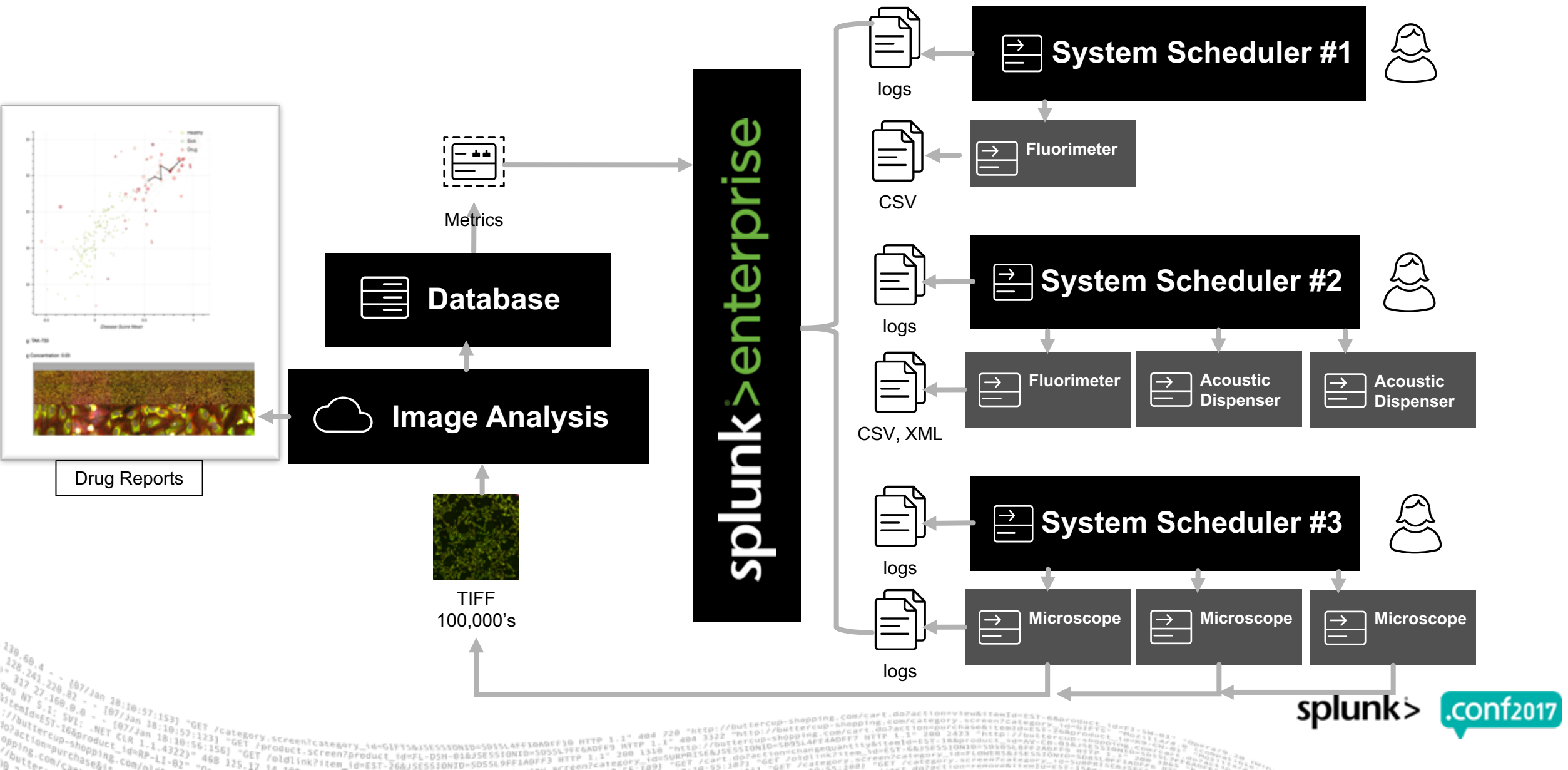
Job       Verbose Mode

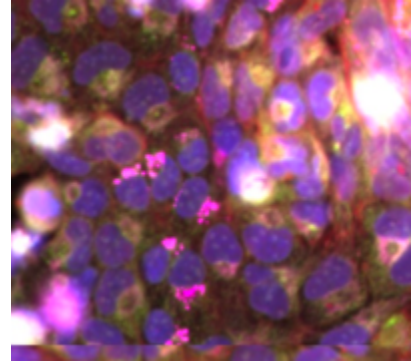
- count
- values(record.XferStatus)

Echo Exceptions, Transfers, And Error Rate Over Time



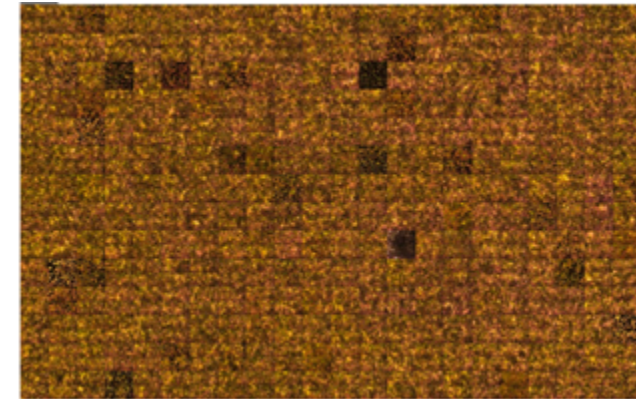
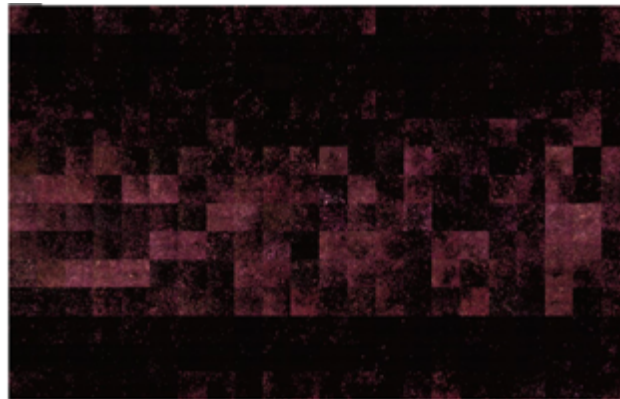
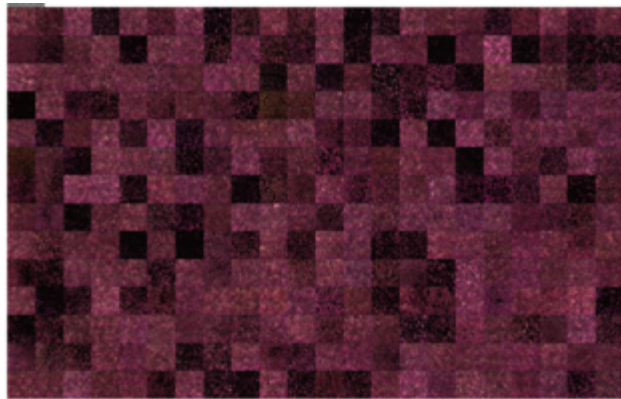
Laboratory Information Management





avg_channel1_intensity ▾	avg_channel2_intensity ▾	avg_channel3_intensity ▾	avg_channel4_intensity ▾	avg_channel5_intensity ▾	avg_channel6_intensity ▾	barcode ▾	cell_type ▾	disease_reagent ▾
25808.1469727	12194.8449707	19789.0744629	16661.4907227	28730.1542969	24949.9677734	mSBXhh	U2OS	s27484
21852.8208008	10148.588623	18187.9821777	13255.1413574	32997.8261719	23067.0234375	XK4smS	U2OS	s29290
20721.7006836	13160.0712891	18416.5869141	14442.5844727	25518.8139648	22767.4267578	cGELoq	U2OS	s195324
19697.0734863	12475.0661621	16913.873291	13995.5244141	20908.534668	22823.34375	LPSf2G	U2OS	s18292
19629.8933105	11648.0418701	19421.2907715	11859.1444092	18833.6164551	20804.0861816	vkNcq	U2OS	s533397

Stain Intensity Principle Component



```
| foreach *_intensity [ eval norm_<<MATCHSTR>> = <<FIELD>> / n_cells]
| fit StandardScaler norm_*
| fit PCA SS_* k=1
| stats avg(metric) as average max(_time) as _time by barcode
```

PC_1	SS_norm_avg_channel1	SS_norm_avg_channel2	SS_norm_avg_channel3	SS_norm_avg_channel4	SS_norm_avg_channel5	SS_norm_avg_channel6
0.542047025991	0.749198975749	0.130088960516	0.130088733826	0.012478124907	0.17494807463	0.215975889902
0.924731645781	1.10612104196	0.191420269008	0.30237843671	0.0665123348751	0.34333317535	0.3999880173
0.281453474654	0.358040510164	-0.0376854954594	0.147796218867	-0.04272317287	0.209605270491	0.093136104118
2.87406083426	2.9549136419	0.7299301176	1.12386480544	0.390252271219	1.07225118623	1.00907488965
3.99845383485	4.0856795404	1.07457372014	1.58964554713	0.589012351913	1.39426279245	1.41795721152
0.0317045244602	0.139728749176	-0.109895263214	0.082216796215	-0.0858320219279	0.0534199246105	0.0182246995409
2.5325122559	2.8388803995	0.572072756197	0.82486665554	0.31430953908	0.913454876995	1.03006538608
0.89421681731	1.19051077965	0.065844292073	0.152014612144	0.023898268711	0.42099863388	0.45431160404
0.668301276731	0.384123255628	0.00709978504416	0.528111465498	-0.010872470252	0.520675382317	0.221012535847
9.16510791085	9.0899603883	1.91543597656	2.74200725041	1.20826653254	5.12292171962	3.16205059482
3.4362062676	3.82621631112	0.551166237993	0.931581662676	0.378034307618	1.51246661703	1.56598287216
1.17998489235	1.44047158894	0.160342356995	0.30590575893	0.0638949104728	0.584319367248	0.473436992279

PC_1 by barcode over time







```
Iteration, 0, Fine Position, 0, 0.00,  
Iteration, 0, Fine Position, 1, 10269.92,  
Site Elapsed Time: 2"  
  
7146.9 0352949801, LAF Parameters, Result, Result, Fine Found, Surface, 0, Iteration, 0, Fine Posit  
-  
0, Iteration, 0, Fine Position, 0, 0.00, (Posit  
0, Iteration, 0, Fine Position, 1, 10269.98, C  
Site Elapsed Time: 3"  
0, Iteration, 0, Fine Position, 1, 10268.80, C  
Site Elapsed Time: 4"  
Well Elapsed Time: B07"
```

What's Next:

- ▶ **More functionality**
 - Alerts with Slack integration
 - Scheduled Reports
- ▶ **More Metrics:**
 - Define control limits
 - By Experiment
 - By Well
- ▶ **More connections**
 - Incubation timing
 - Environmental monitoring
- ▶ **More Data**
 - Inventory control
 - Protocol changes
 - Application performance
- ▶ **More automation**
 - Integrating reports to schedule workloads
 - Using alerts for automation control

Thank You

Don't forget to **rate this session** in the
.conf2017 mobile app

splunk> .conf2017