



Model Velocity at Scale

Presented by Jan Krynauw
6 December 2019

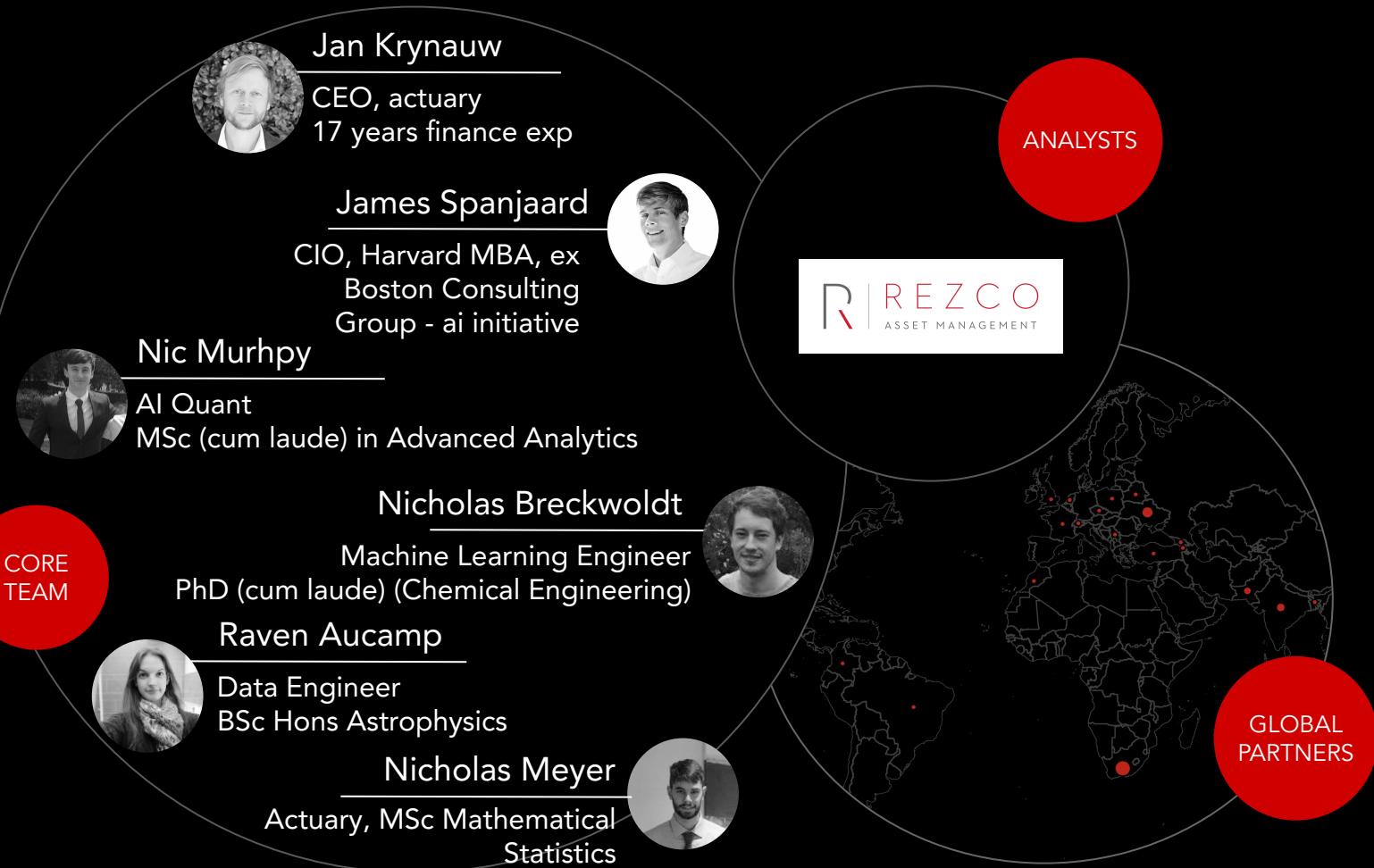


about us

autonomous learning investment systems

independent quant investment team within Rezco Asset Management focusing on the application of AI to investing

close and symbiotic relationship with Rezco investment team



what is investing?

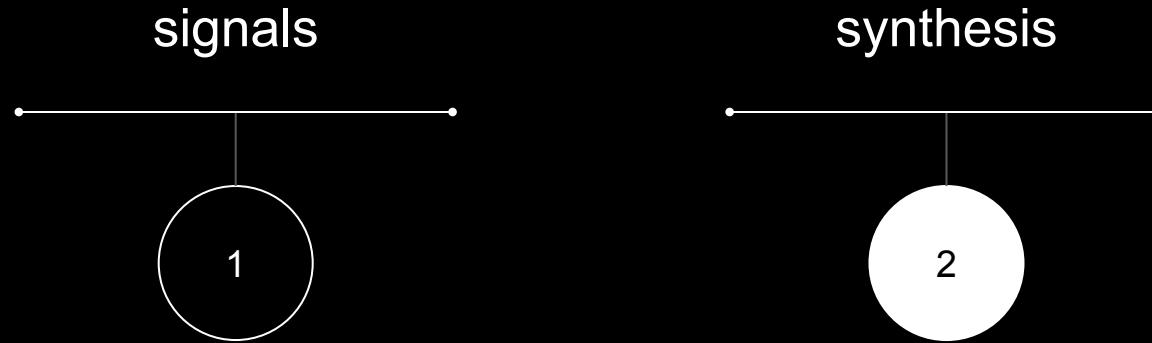


synthesis of information

for the purpose of...

pricing an asset

Two stories to machine learning:



01

Disparate Dirty Datasets

02

Markets are Not
Stationary

03

Interactions are
non-linear and high
dimensional

04

Large Search Space in
Finance

05

Low Signal to Noise Ratio

What we do is Hard.

pick a stock...

company_name	primary_ticker
Ascena Retail Group Inc	ASNA.O
Tailored Brands Inc	TLRD.K
J.Jill Inc	JILL.K
Gap Inc	GPS
Express Inc	EXPR.K
Caleres Inc	CAL
Urban Outfitters Inc	URBN.O
Foot Locker Inc	FL
L Brands Inc	LB
Buckle Inc	BKE
Tillys Inc	TLYS.K
Chico's FAS Inc	CHS
Genesco Inc	GCO
American Eagle Outfitters Inc	AEO
Childrens Place Inc	PLCE.O
Abercrombie & Fitch Co	ANF
TJX Companies Inc	TJX
Boot Barn Holdings Inc	BOOT.K
Ross Stores Inc	ROST.O
Burlington Stores Inc	BURL.K

FCF to EV	PE FY0	PE FY1	PE FY2
	-25.9	-1.4	
	2.3	3.1	2.8
	3.0	10.9	5.8
9.4	7.3	9.1	8.8
171.7	7.7	-32.5	93.3
10.2	8.1	7.5	6.8
6.7	8.4	9.3	8.5
11.1	8.9	8.3	7.7
6.2	9.2	10.4	9.9
	10.1	11.1	12.6
16.0	10.3	11.2	9.8
27.2	10.5	131.9	48.7
11.5	11.6	10.9	10.1
9.8	12.3	11.3	10.3
4.9	14.7	16.1	11.6
7.3	16.2	18.8	12.1
4.1	22.5	20.8	19.1
1.3	24.1	21.6	17.7
4.1	25.4	23.6	21.4
3.1	27.9	25.7	22.6

pick a stock...

company_name	primary_ticker
Ascena Retail Group Inc	ASNA.O
Tailored Brands Inc	TLRD.K
J.Jill Inc	JILL.K
Gap Inc	GPS
Express Inc	EXPR.K
Caleres Inc	CAL
Urban Outfitters Inc	URBN.O
Foot Locker Inc	FL
L Brands Inc	LB
Buckle Inc	BKE
Tillys Inc	TLYS.K
Chico's FAS Inc	CHS
Genesco Inc	GCO
American Eagle Outfitters Inc	AEO
Childrens Place Inc	PLCE.O
Abercrombie & Fitch Co	ANF
TJX Companies Inc	TJX
Boot Barn Holdings Inc	BOOT.K
Ross Stores Inc	ROST.O
Burlington Stores Inc	BURL.K

SGR	CAGR (-2,2)	FCF to EV	PE FY0	PE FY1	PE FY2
			-25.9	-1.4	
5.0	1.6		2.3	3.1	2.8
7.3			3.0	10.9	5.8
5.0	1.4	9.4	7.3	9.1	8.8
5.0	-57.6	171.7	7.7	-32.5	93.3
11.4	7.3	10.2	8.1	7.5	6.8
5.0	9.4	6.7	8.4	9.3	8.5
7.0	3.0	11.1	8.9	8.3	7.7
5.0	-8.6	6.2	9.2	10.4	9.9
5.0	-6.1		10.1	11.1	12.6
6.9	18.1	16.0	10.3	11.2	9.8
25.0	-46.4	27.2	10.5	131.9	48.7
6.2	-2.2	11.5	11.6	10.9	10.1
9.0	9.0	9.8	12.3	11.3	10.3
10.4	12.0	4.9	14.7	16.1	11.6
19.5		7.3	16.2	18.8	12.1
9.4	12.8	4.1	22.5	20.8	19.1
13.5	35.6	1.3	24.1	21.6	17.7
9.6	15.2	4.1	25.4	23.6	21.4
12.2	25.3	3.1	27.9	25.7	22.6

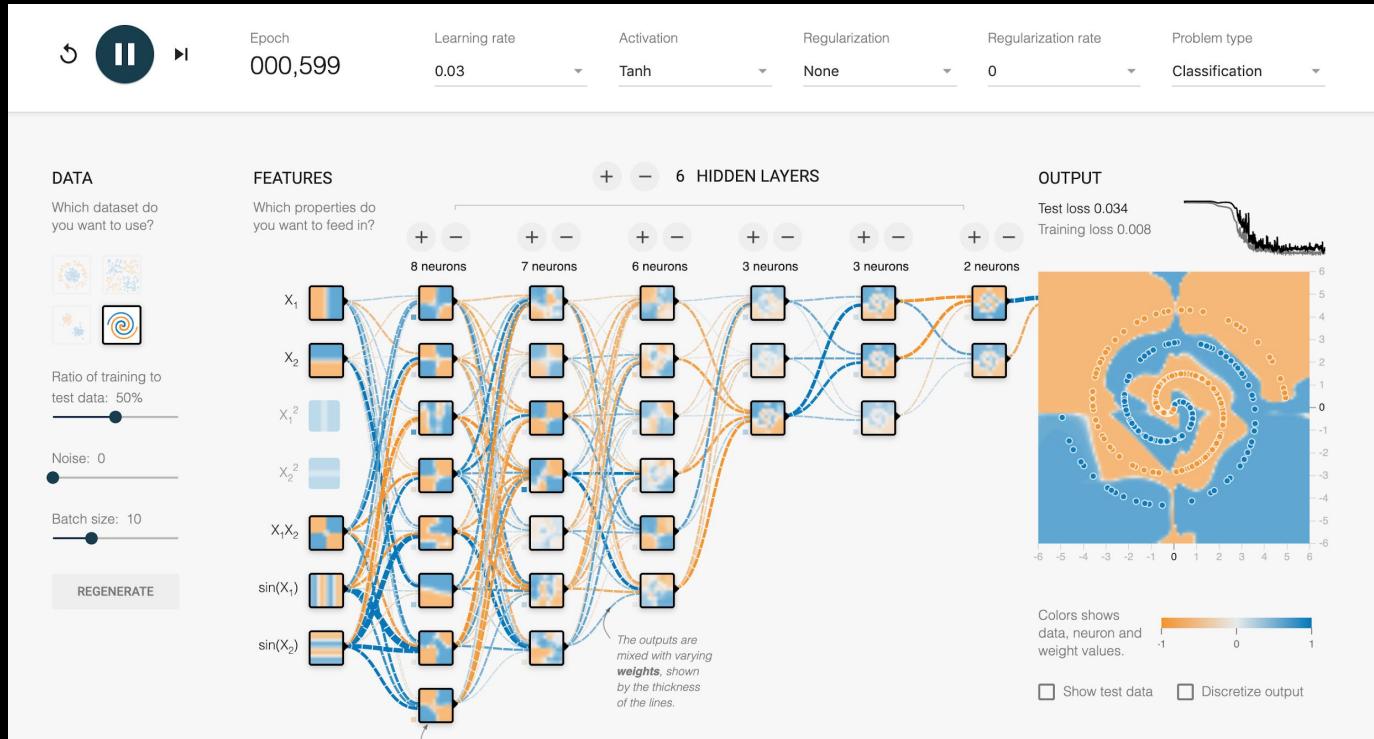
pick a stock...

company_name	primary_ticker
Ascena Retail Group Inc	ASNA.O
Tailored Brands Inc	TLRD.K
J.Jill Inc	JILL.K
Gap Inc	GPS
Express Inc	EXPR.K
Caleres Inc	CAL
Urban Outfitters Inc	URBN.O
Foot Locker Inc	FL
L Brands Inc	LB
Buckle Inc	BKE
Tillys Inc	TLYS.K
Chico's FAS Inc	CHS
Genesco Inc	GCO
American Eagle Outfitters Inc	AEO
Childrens Place Inc	PLCE.O
Abercrombie & Fitch Co	ANF
TJX Companies Inc	TJX
Boot Barn Holdings Inc	BOOT.K
Ross Stores Inc	ROST.O
Burlington Stores Inc	BURL.K

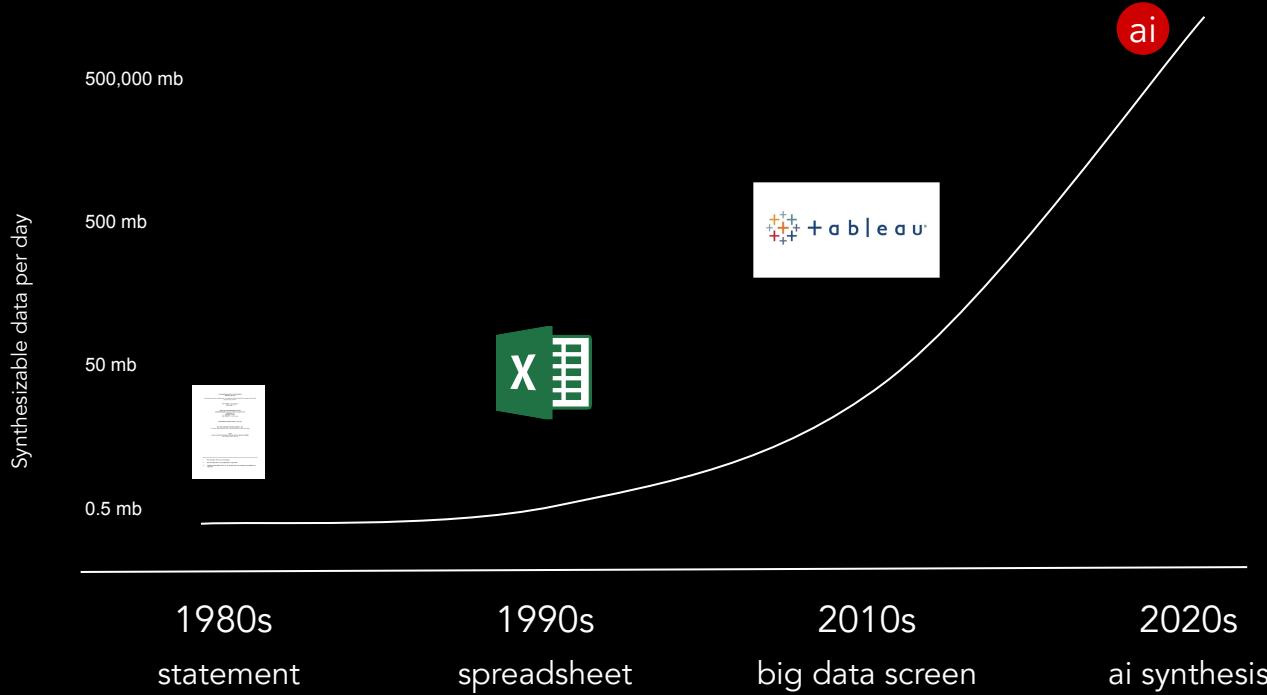
SGR	CAGR (-2,2)	FCF to EV	PE FY0	PE FY1	PE FY2	GPE FY1	GPE FY1 CAGR(..)	Growth (-3,-2)	Growth (-2,-1)	Growth (-1,0)	Growth (0,1)	Growth (1,2)
Ascena Retail Group Inc	ASNA.O		-25.9	-1.4			1.7	-63.3	-109.1			
Tailored Brands Inc	TLRD.K	5.0	1.6	2.3	3.1	2.8	1.5957	3.67	-2.2	25.0	5.0	-27.1
J.Jill Inc	JILL.K	7.3		3.0	10.9	5.8	0.6719	7.99			-8.9	-72.4
Gap Inc	GPS	5.0	1.4	9.4	7.3	9.1	0.5507	0.37	-16.9	5.5	21.6	-20.1
Express Inc	EXPR.K	5.0	-57.6	171.7	7.7	-32.5	93.3	-0.1539	-44.1	-55.6	-11.1	-123.6
Caleres Inc	CAL	11.4	7.3	10.2	8.1	7.5	6.8	1.5126	1.47	0.0	8.0	2.3
Urban Outfitters Inc	URBN.O	5.0	9.4	6.7	8.4	9.3	8.5	0.5349	1.01	1.1	-15.6	73.3
Foot Locker Inc	FL	7.0	3.0	11.1	8.9	8.3	7.7	0.8385	0.95	12.4	-14.7	14.6
L Brands Inc	LB	5.0	-8.6	6.2	9.2	10.4	9.9	0.4809	0.47	-6.3	-14.4	-11.9
Buckle Inc	BKE	5.0	-6.1		10.1	11.1	12.6	0.4506	-1.07	-33.7	-8.9	6.5
Tillys Inc	TLYS.K	6.9	18.1	16.0	10.3	11.2	9.8	0.6179	1.20	2.4	46.5	27.0
Chico's FAS Inc	CHS	25.0	-46.4	27.2	10.5	131.9	48.7	0.1896	1.30	8.0	-16.1	-54.4
Genesco Inc	GCO	6.2	-2.2	11.5	11.6	10.9	10.1	0.5666	0.78	0.9	-27.7	10.5
American Eagle Outfitters Inc	AEO	9.0	9.0	9.8	12.3	11.3	10.3	0.7957	0.85	14.7	-7.2	27.6
Childrens Place Inc	PLCE.O	10.4	12.0	4.9	14.7	16.1	11.6	0.6469	2.37	50.8	45.7	-14.7
Abercrombie & Fitch Co	ANF	19.5		7.3	16.2	18.8	12.1	1.038	2.93	-105.4	-1,183.3	76.9
TJX Companies Inc	TJX	9.4	12.8	4.1	22.5	20.8	19.1	0.452	0.43	6.0	12.2	22.7
Boot Barn Holdings Inc	BOOT.K	13.5	35.6	1.3	24.1	21.6	17.7	0.6264	0.99	-20.3	40.0	77.9
Ross Stores Inc	ROST.O	9.6	15.2	4.1	25.4	23.6	21.4	0.4074	0.43	12.8	18.0	25.5
Burlington Stores Inc	BURL.K	12.2	25.3	3.1	27.9	25.7	22.6	0.4745	0.55	40.3	34.9	47.4

how does ML tackle this challenge...

<https://playground.tensorflow.org>



Our observation

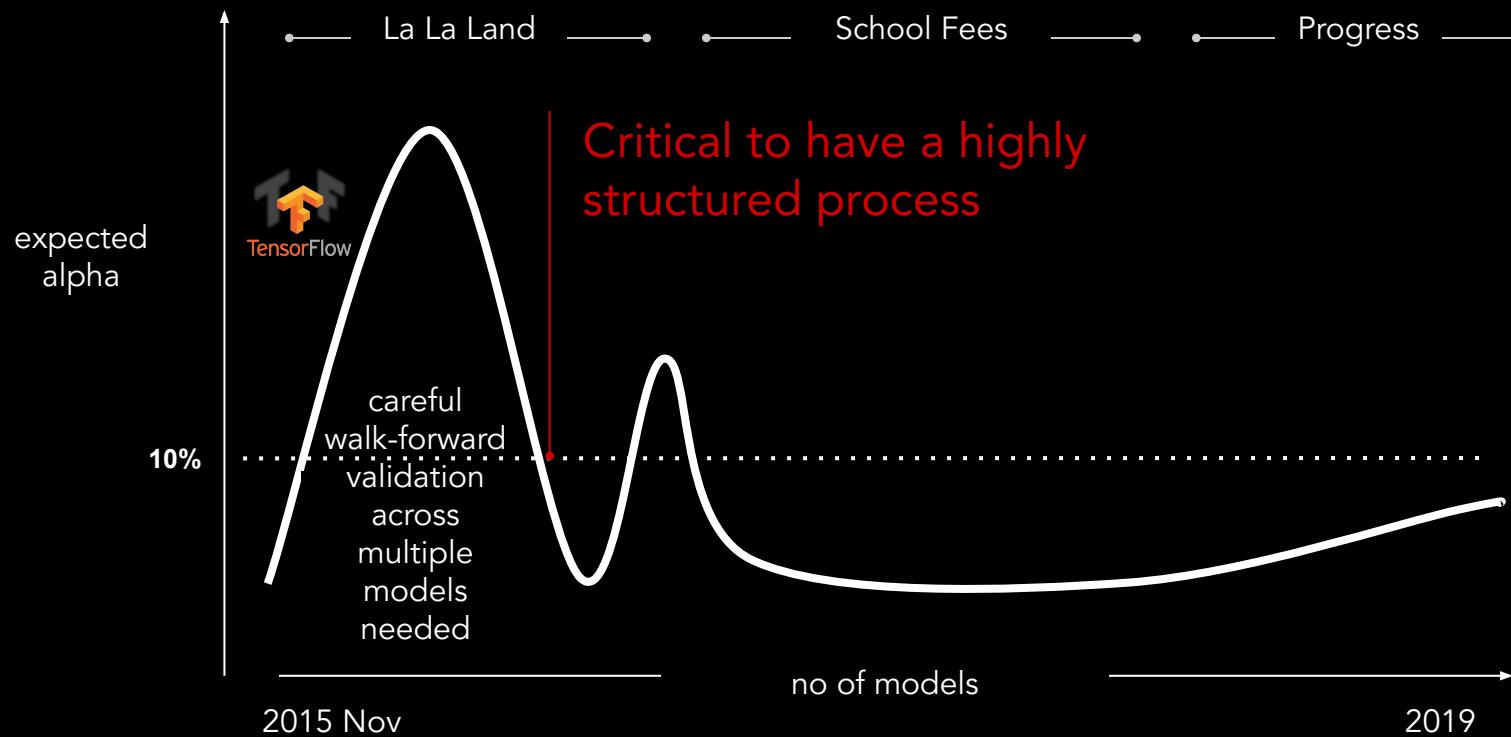


The amount of information an investment manager must ingest to remain competitive grows exponentially

Machines are shifting from presenting to synthesizing information

autonomous learning investment systems are the excel spreadsheets of the next decade

Our journey with machine learning....



In tech everything
reduces to scale.



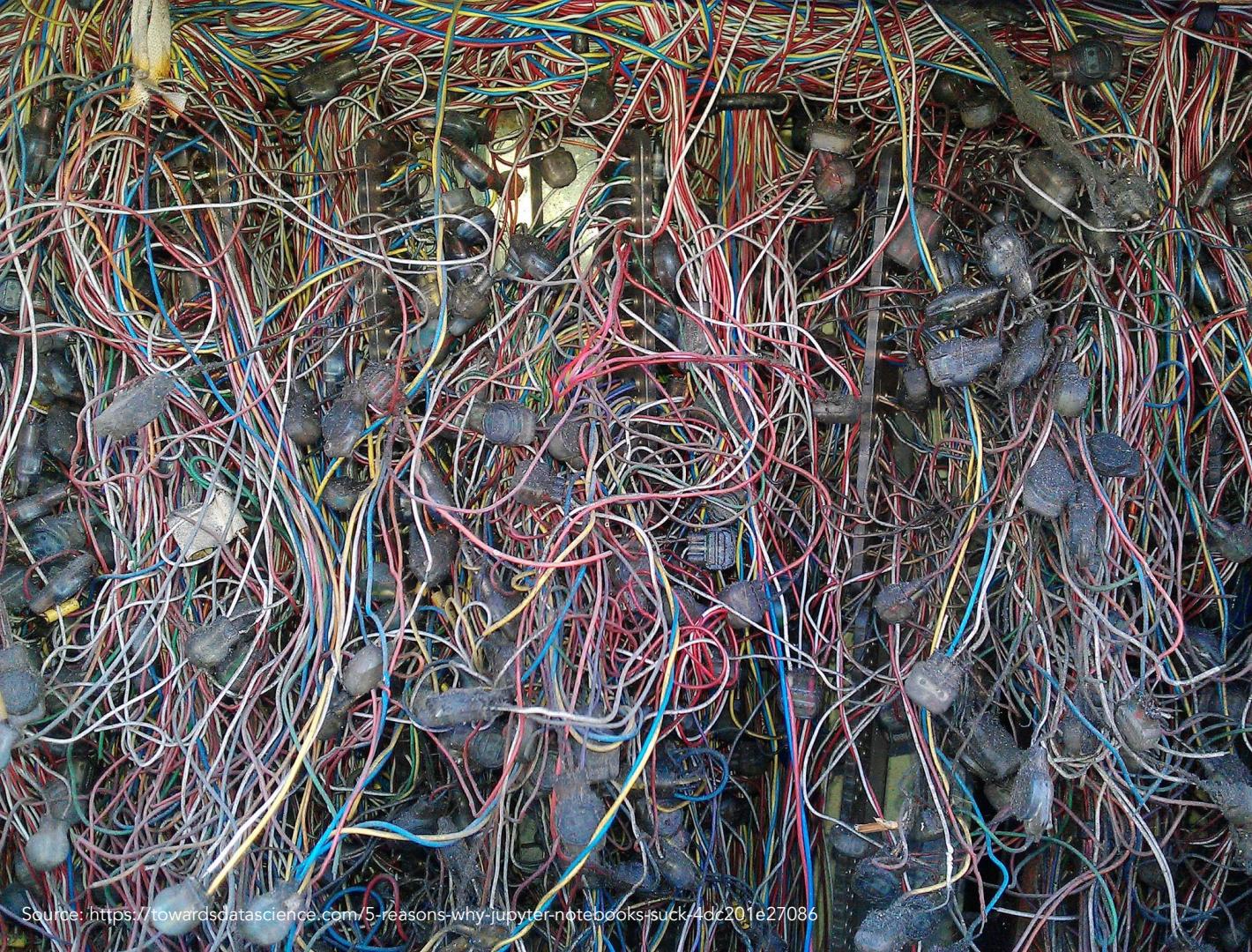
scaling:

01 DevOps

02 Team

03 Model Research Cycles

04



Jupyter Notebooks as a Development Environment

Source: <https://towardsdatascience.com/5-reasons-why-jupyter-notebooks-suck-4dc201e27086>

- Easily switch between Production and Development environments
- IDEs makes it easy to enable good code versioning
- Integrated CI/CD
- Integrated linting and code-style correction
- Jupyter is really bad for running long asynchronous tasks
- Dockerize!
- non-linear workflow of jupyter

Use
an
IDE
(integrated
development
environment)



pandas?



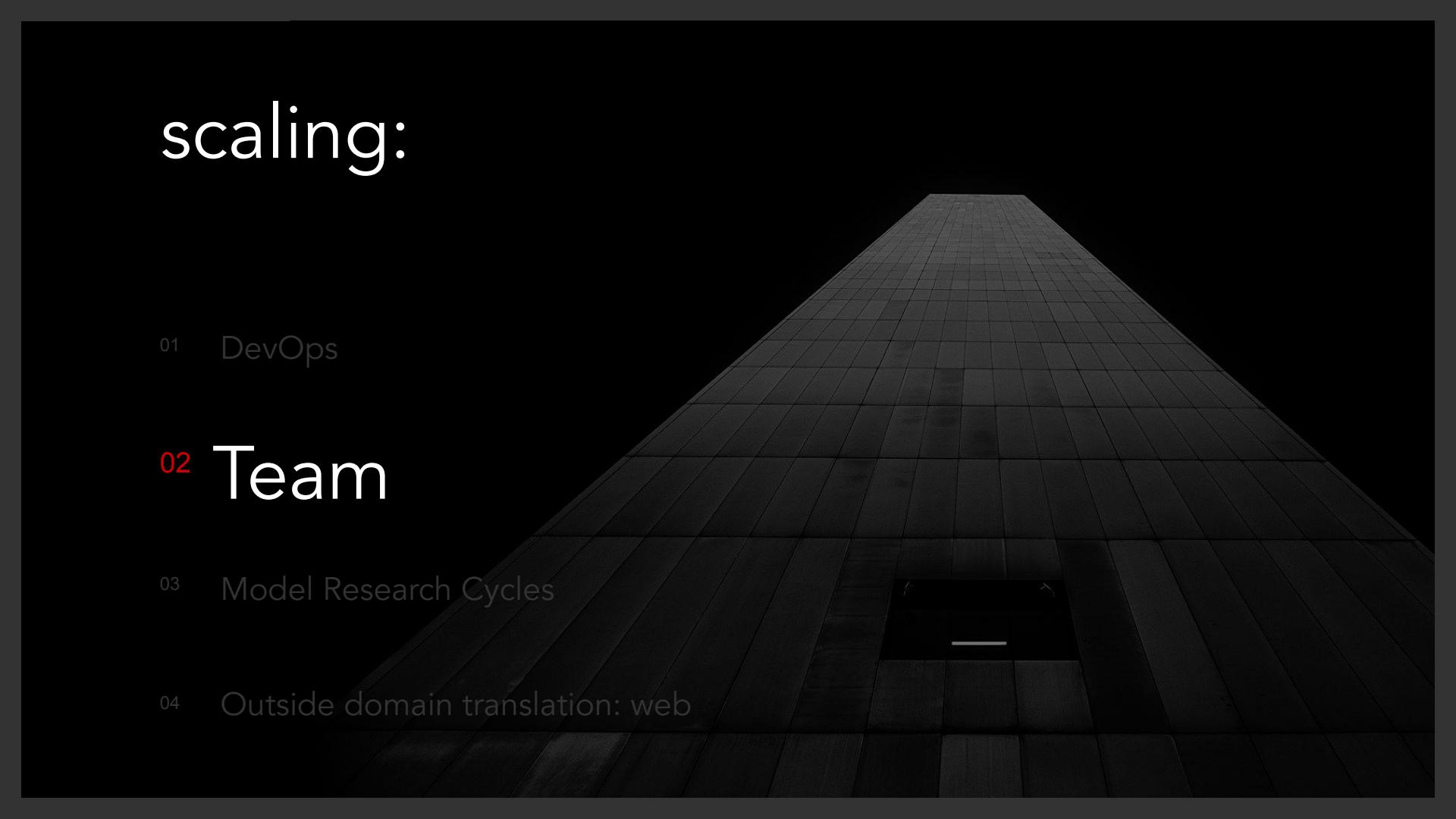
scaling:

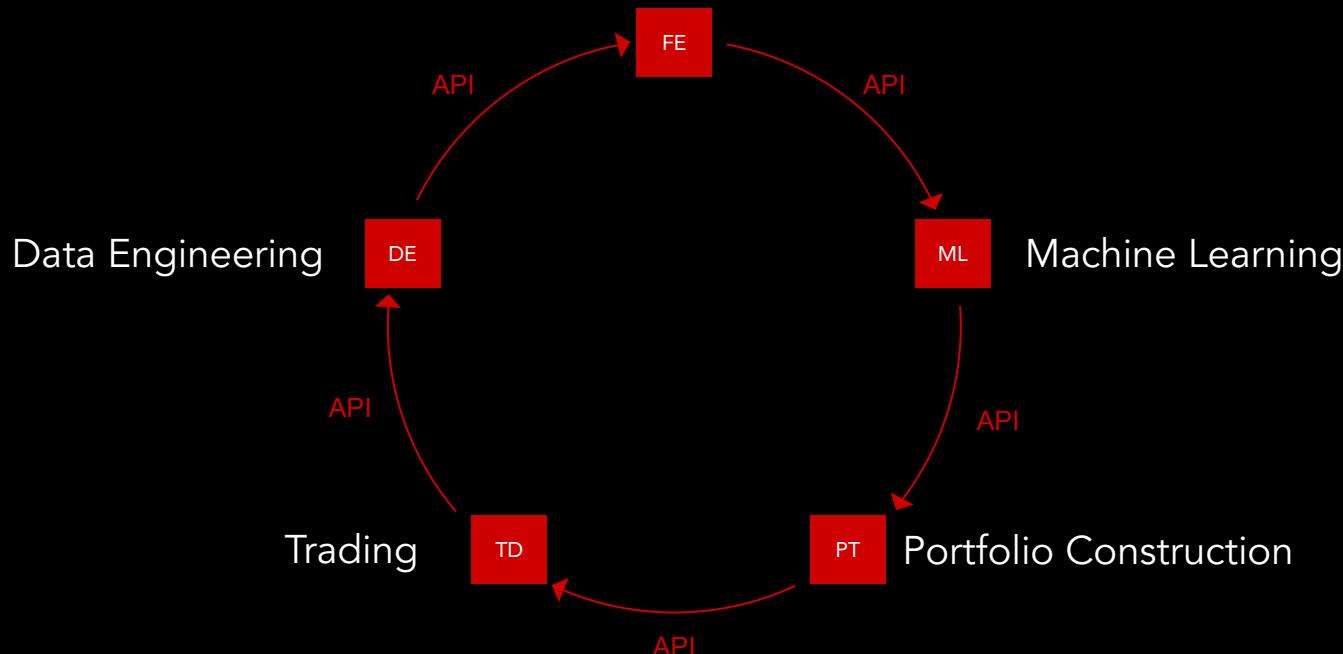
01 DevOps

02 Team

03 Model Research Cycles

04 Outside domain translation: web





How we
structured
the team

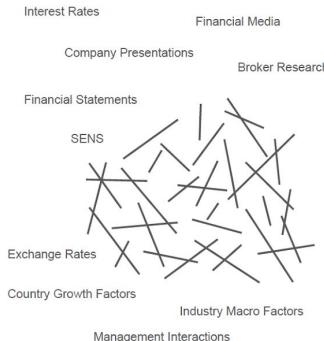
50+ External Developers
worked with during the last 2 years



Seamless
integration
with
Investment
Team

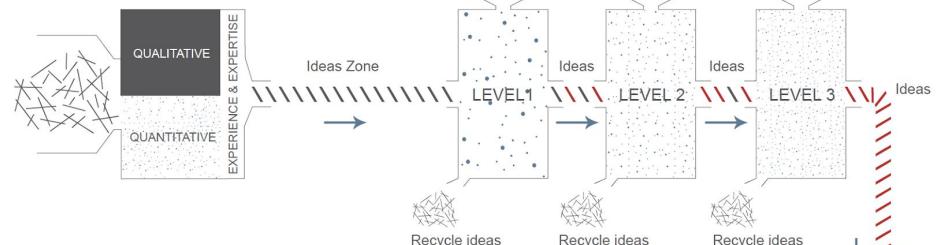
THE REZCO INVESTMENT PROCESS = APPLICATION OF OUR INVESTMENT PHILOSOPHY

01 SOURCES OF INFORMATION



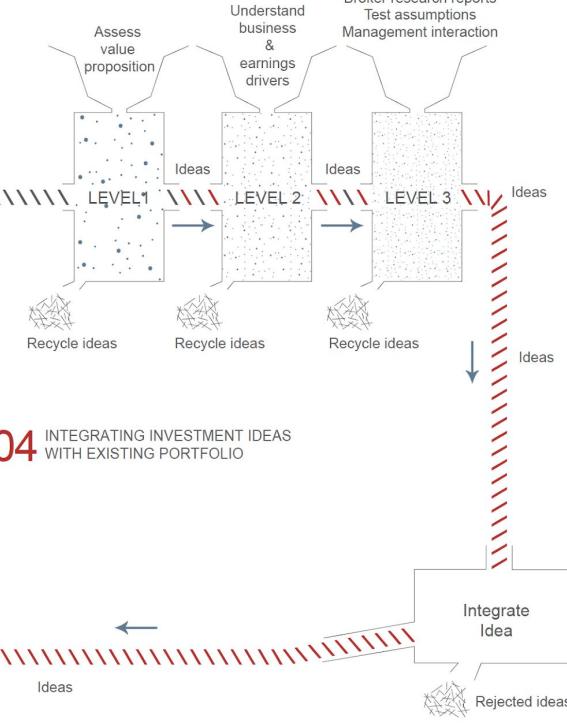
02 SORT, PROCESS & FILTER

- Generate ideas using wealth of experience
- Expertise and experience enable us to establish importance
- Search for value gems
- Objective is to find mispriced sectors, asset classes or securities
- Identify ideas

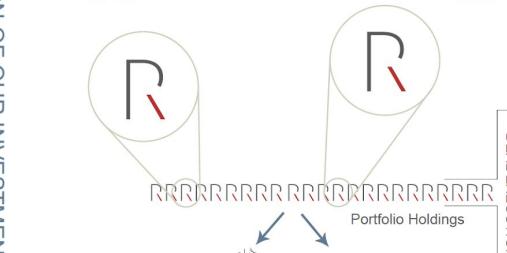


03 REFINE IDEA

- Internal research reports
- Detailed company study
- Exposing potential flaws
- Broker research reports
- Test assumptions
- Management interaction



06 ASSESSMENT OF DECISIONS



05 PORTFOLIO CONSTRUCTION

- Allocate ideas to conviction buckets
- Apply discretion/judgement within buckets

04 INTEGRATING INVESTMENT IDEAS WITH EXISTING PORTFOLIO

- Review relative values
- Apply macro views
- Review suitability of ideas vs current holdings
- Review exposure to industries, countries, themes



alis_

Human

alis_ as a member of the Investment Team

	Add Task	Rob	Simon	Industry	Mode...	Prediction Vi...	
8	✓ Cobham PLC	Idea Fact... alis_ 🤔	✓	Indus...	196	https://analyt...	
9	✓ Lockheed Martin Corp	Idea Fact... alis_ 🤗	✓			https://analyt...	
10	✓ Amazon.com Inc	Idea Fact... alis_ 🤔	✓			https://analyt...	
11	✓ PNC Financial Services Gr	Idea Fact... ali... 🤗	✓			https://analyt...	
12	✓ Regions Financial Corp	Idea Fact... alis_ 🤗	✓			https://analyt...	
13	✓ Citizens Financial Group Ir	Idea Fact... alis_ 🤗	✓			https://analyt...	
14	✓ UnitedHealth Group Inc	Idea Fact... alis_ 🤗	✓			https://analyt...	
15	✓ eBay Inc	Idea Fact... alis_ 🤗	✓			https://analyt...	
16	✓ Comerica Inc	Idea Fact... alis_ 🤗	✓	Finan...	183	https://analyt...	

scaling:

01 DevOps

02 Team

03 Model Research Cycles

04 Outside domain translation: web

Curse of Dimensionality

- 01 Problem type (regression vs classification)
- 02 Label choice (say 10 label types)
- 03 Model architecture (MLPs, CNNs or RNNs)
- 04 Model hyperparameters (number of layers, number of hidden units per layer, pooling, dropout)
- 05 Activation functions (ReLU, leaky ReLU, swish, GELU)
- 06 More complex architectures (SNN, ResNet, Gated CNN, LSTM, GRUs, Dilated CNNs, Separable CNNs)
- 07 5 Markets, 11 Sectors
- 08 Little 't'...say last 10 years, quarterly models

Number of Models

2 m

Cloud Compute &
Distributed ML &
Workflow Framework
(100 000+ teraflops)

alis_ operating system



scaling:

01 DevOps

02 Team

03 Model Research Cycles

04 Outside domain translation: web

Scaling a simple
task makes it
complex



An example of a simple data exercise

<https://colab.research.google.com/drive/19Wu6nx3wjnn2EuZ9DxyaartJXiOdpLov>

▼ let's get some data

first let's find the identifying data for the company we are interested in, Tombstone

```
[33] # https://www.openfigi.com/api#post-v2-search
openfigi_url = "https://api.openfigi.com/v2/search"
company = "Tombstone"
body = {
    'query': company,
    'exchCode': 'US',
    'securityType': 'Common Stock',
    'includeUnlistedEquities': True,
}

# run request
headers = {
    'Content-Type': 'application/json'
}
requestResponse = requests.post(openfigi_url, json.dumps(body), headers=headers)
openfigi_data = requestResponse.json()
data = openfigi_data['data']

company_identifiers = data[0]

# uncomment this to see the full response
# print(json.dumps(company_identifiers, indent=2))

ticker = company_identifiers['ticker']
print(f'The ticker for {company} is {ticker}')
```

→ The ticker for Tombstone is TMBXF

great, now let's get some price information

One company, one day is easy

How about 65 000 companies, 5,000 days, pulled daily?

Per single REST API Call	
DNS Resolution Time	1 ms
Connection Time	328 ms
SSL Handshake Time	832 ms
Header Response Time	2.44 s
Total Response Time	3.61 s
Response Body Size	939 bytes

$$\times 65\,536 = 2.7 \text{ days}$$

we'll have to get more creative...

Let's get some data for 65 536 tickers...

```
1 # my long list of tickers
2 my_tickers = ['ticker1', 'ticker2', 'ticker3', 'ticker4', 'ticker5', 'ticker6']
3
4 if len(my_tickers) == 1:
5     # Make the REST API request on my_tickers[0]
6     print(f'Making a REST API request for ${my_tickers[0]}... ')
7 else:
8     # Splitting the list in two and repeat
9     print(f'Splitting the list in two... ')
10
```

Splits

Requests

1

1

2

2

2

3

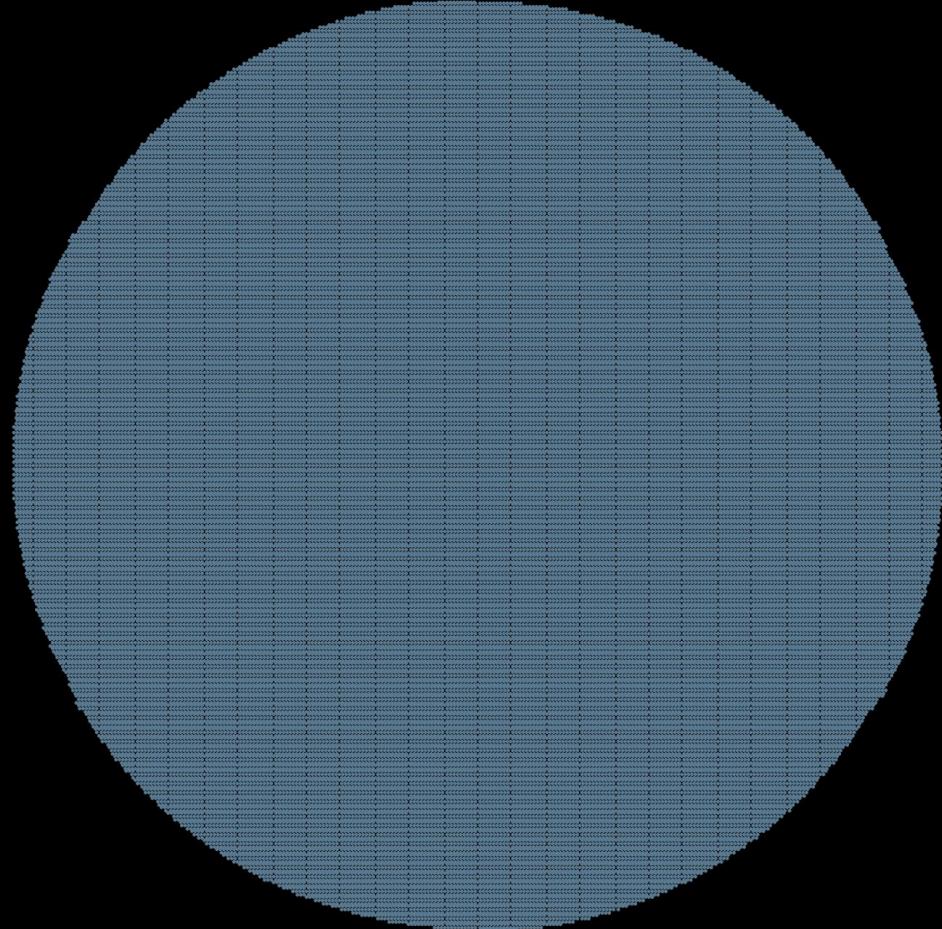
4

5

6

7

8



65 536
REST API
Requests

Split in 4



alis_

Human

Data Pipelines are Hard

We source data from various sources /
technologies / humans
alis_ handles 99.9% of pipelines.



What about inconsistencies?

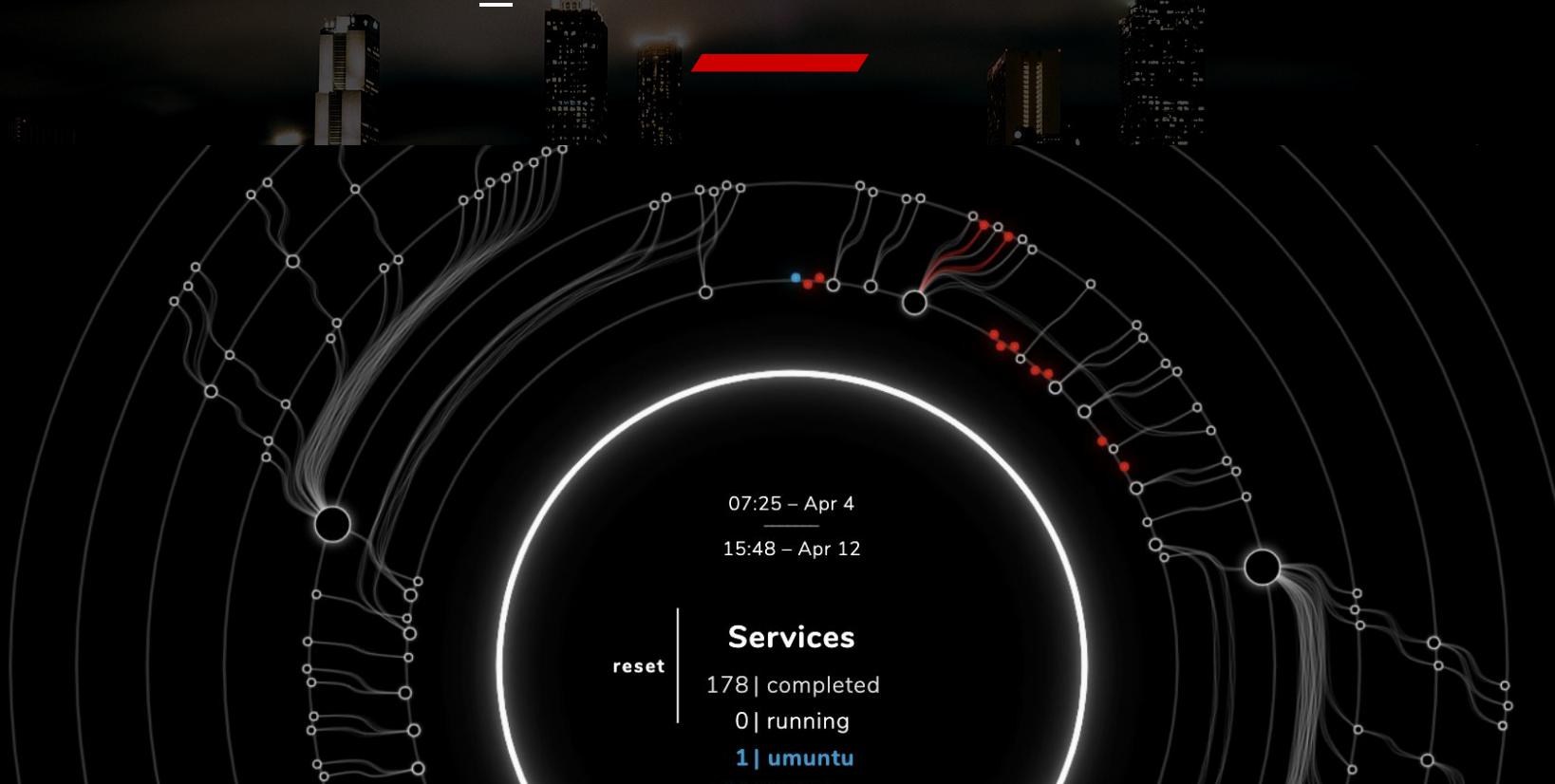
CSV Latest original.csv ×

1	Bond Data;;;	1	Bond Data,,
2	Trade Date:;2019/04/03;;;	2	Trade Date:,2019/04/03,,
3	;;;;;	3	;;;;;
4	Bond Code;Pricing Class Code;ISIN No;Issuer;Issuer Cl	4	Bond Code,Pricing Class Code,ISIN No,Issuer,Issuer
5	;;;;;	5	;;;;;
6	AA06;FRN;ZAG000114836;ANGLO AMERICAN SA FINANCE LIMITED	6	AA06,FRN,ZAG000114836,ANGLO AMERICAN SA FINANCE LIMITED
7	AA07;V;ZAG000114828;ANGLO AMERICAN SA FINANCE LIMITED	7	AA07,V,ZAG000114828,ANGLO AMERICAN SA FINANCE LIMITED
8	ABCPI3;CPI;ZAG000073669;ABSA BANK LIMITED;H10;ZAG;ZAF	8	ABCPI3,CPI,ZAG000073669,ABSA BANK LIMITED,H10,ZAG,ZAF
9	ABFN08;FRN;ZAG000084757;ABSA BANK LIMITED;H10;ZAG;ZAF	9	ABFN08,FRN,ZAG000084757,ABSA BANK LIMITED,H10,ZAG,ZAF
10	ABFN09;FRN;ZAG000084765;ABSA BANK LIMITED;H10;ZAG;ZAF	10	ABFN09,FRN,ZAG000084765,ABSA BANK LIMITED,H10,ZAG,ZAF
11	ABFN10;FRN;ZAG000109919;ABSA BANK LIMITED;H10;ZAG;ZAF	11	ABFN10,FRN,ZAG000109919,ABSA BANK LIMITED,H10,ZAG,ZAF
12	ABFN12;FRN;ZAG000114455;ABSA BANK LIMITED;H10;ZAG;ZAF	12	ABFN12,FRN,ZAG000114455,ABSA BANK LIMITED,H10,ZAG,ZAF
13	ABFN13;FRN;ZAG000126160;ABSA BANK LIMITED;H10;ZAG;ZAF	13	ABFN13,FRN,ZAG000126160,ABSA BANK LIMITED,H10,ZAG,ZAF
14	ABFN14;FRN;ZAG000126178;ABSA BANK LIMITED;H10;ZAG;ZAF	14	ABFN14,FRN,ZAG000126178,ABSA BANK LIMITED,H10,ZAG,ZAF
15	ABFN15;FRN;ZAG000130949;ABSA BANK LIMITED;H10;ZAG;ZAF	15	ABFN15,FRN,ZAG000130949,ABSA BANK LIMITED,H10,ZAG,ZAF
16	ABFN16;FRN;ZAG000130964;ABSA BANK LIMITED;H10;ZAG;ZAF	16	ABFN16,FRN,ZAG000130964,ABSA BANK LIMITED,H10,ZAG,ZAF
17	ABFN17;FRN;ZAG000136607;ABSA BANK LIMITED;H10;ZAG;ZAF	17	ABFN17,FRN,ZAG000136607,ABSA BANK LIMITED,H10,ZAG,ZAF
18	ABFN18;FRN;ZAG000139684;ABSA BANK LIMITED;H10;ZAG;ZAF	18	ABFN18,FRN,ZAG000139684,ABSA BANK LIMITED,H10,ZAG,ZAF
19	ABFN20;FRN;ZAG000144213;ABSA BANK LIMITED;H10;ZAG;ZAF	19	ABFN20,FRN,ZAG000144213,ABSA BANK LIMITED,H10,ZAG,ZAF
20	ABFN21;FRN;ZAG000144221;ABSA BANK LIMITED;H10;ZAG;ZAF	20	ABFN21,FRN,ZAG000144221,ABSA BANK LIMITED,H10,ZAG,ZAF

CSV Latest corrected.csv ×



alis_ enters UMUNTU state





alis_ enters UMUNTU state

Show Project Description

DE



FE

jse-bond-data



Today

UM...

jse-bon...

jse-bond-data



Assigned To
Jan Krynauw



Due Date
Today

ALIS Task State

UMUNTU

Domain

DE | Data Engineering

Service

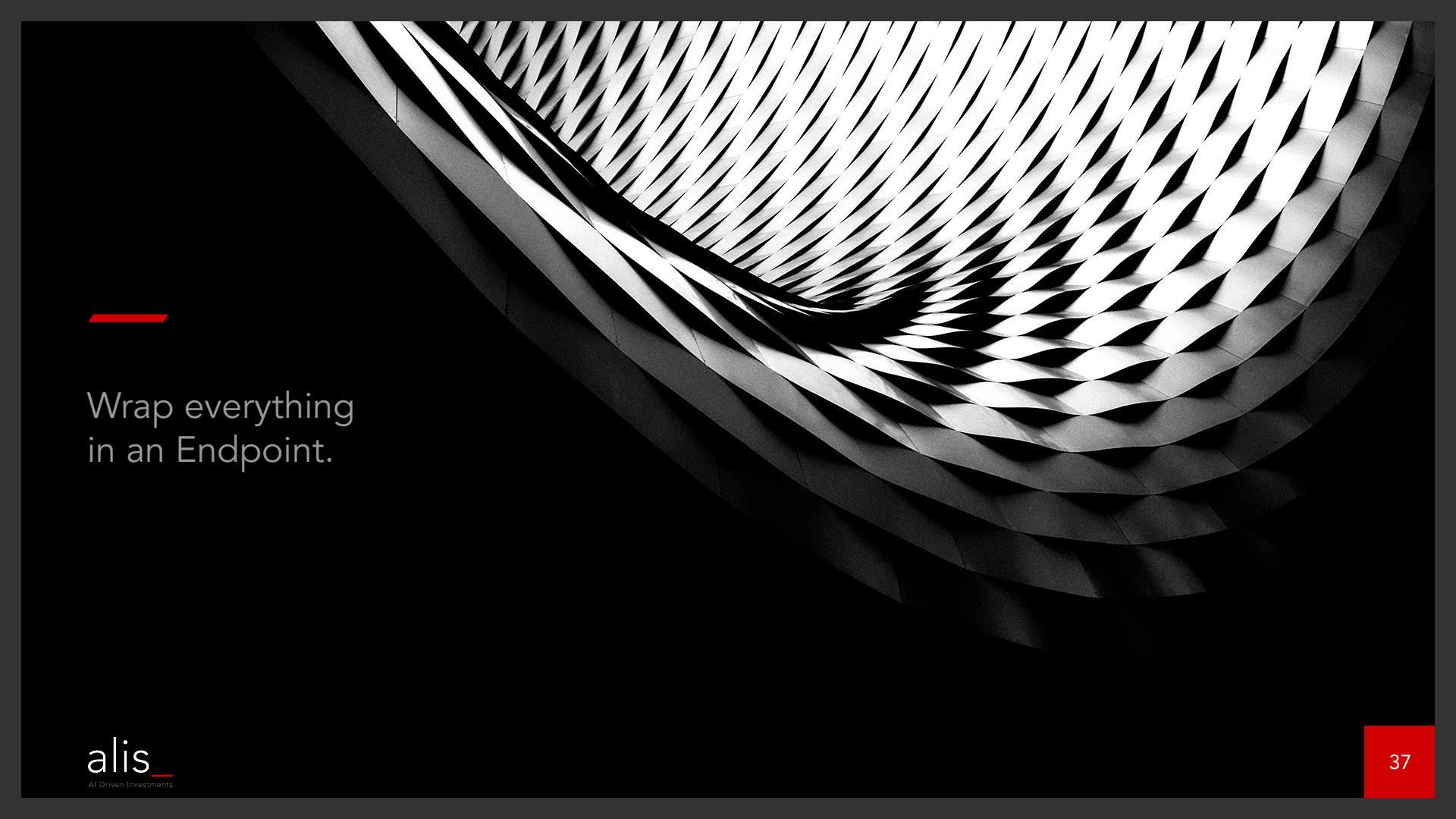
jse-bond-data

Version

r1-0

Parent Service

Expected delimiter "," was not found in .csv file.



—

Wrap everything
in an Endpoint.



Thank you

Contact Us

Jan Krynauw | jan@alis.fund | +27 82 928 4511 | www.alis.fund