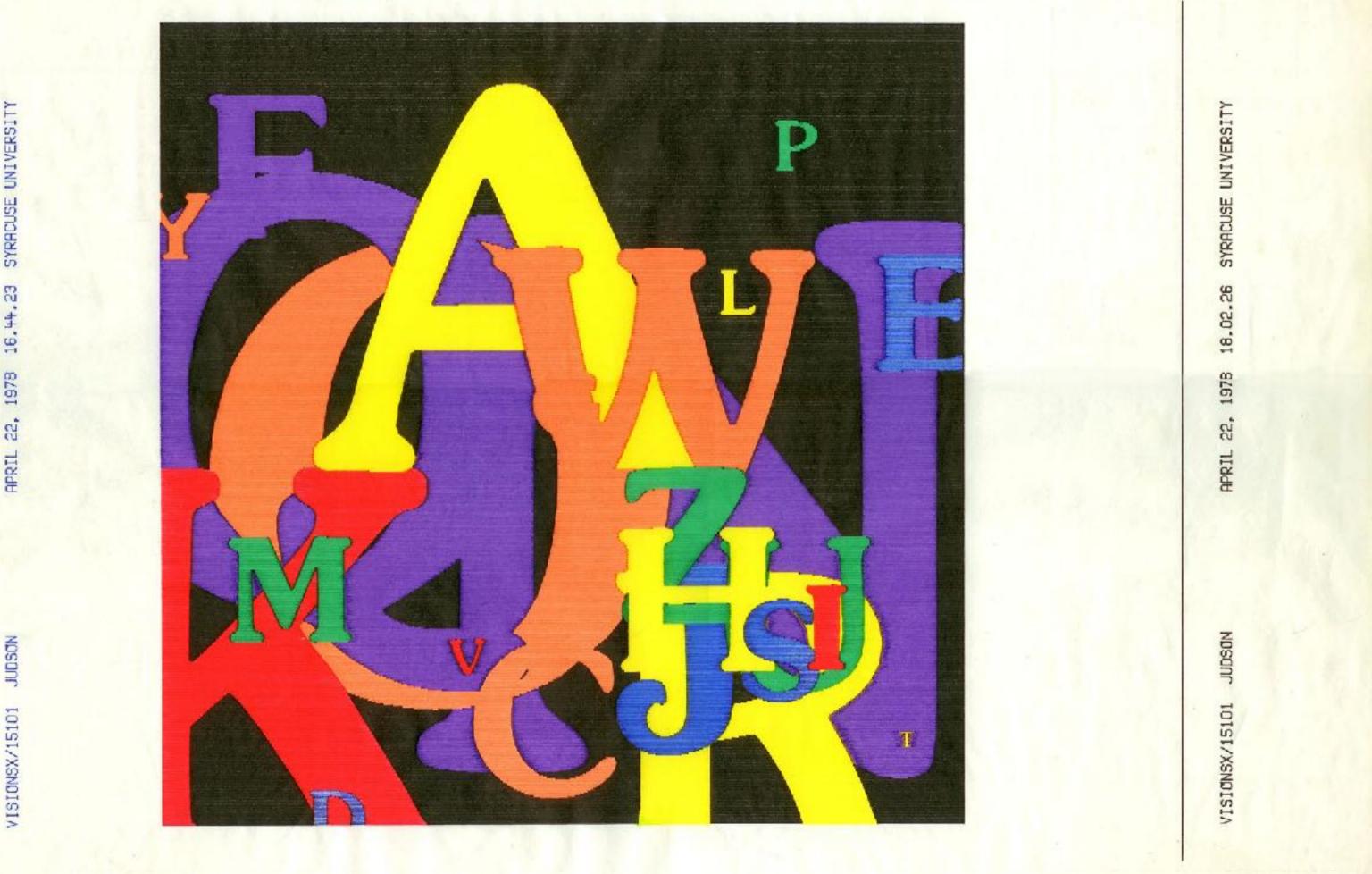
From Design to Search in High-Dimensional Spaces

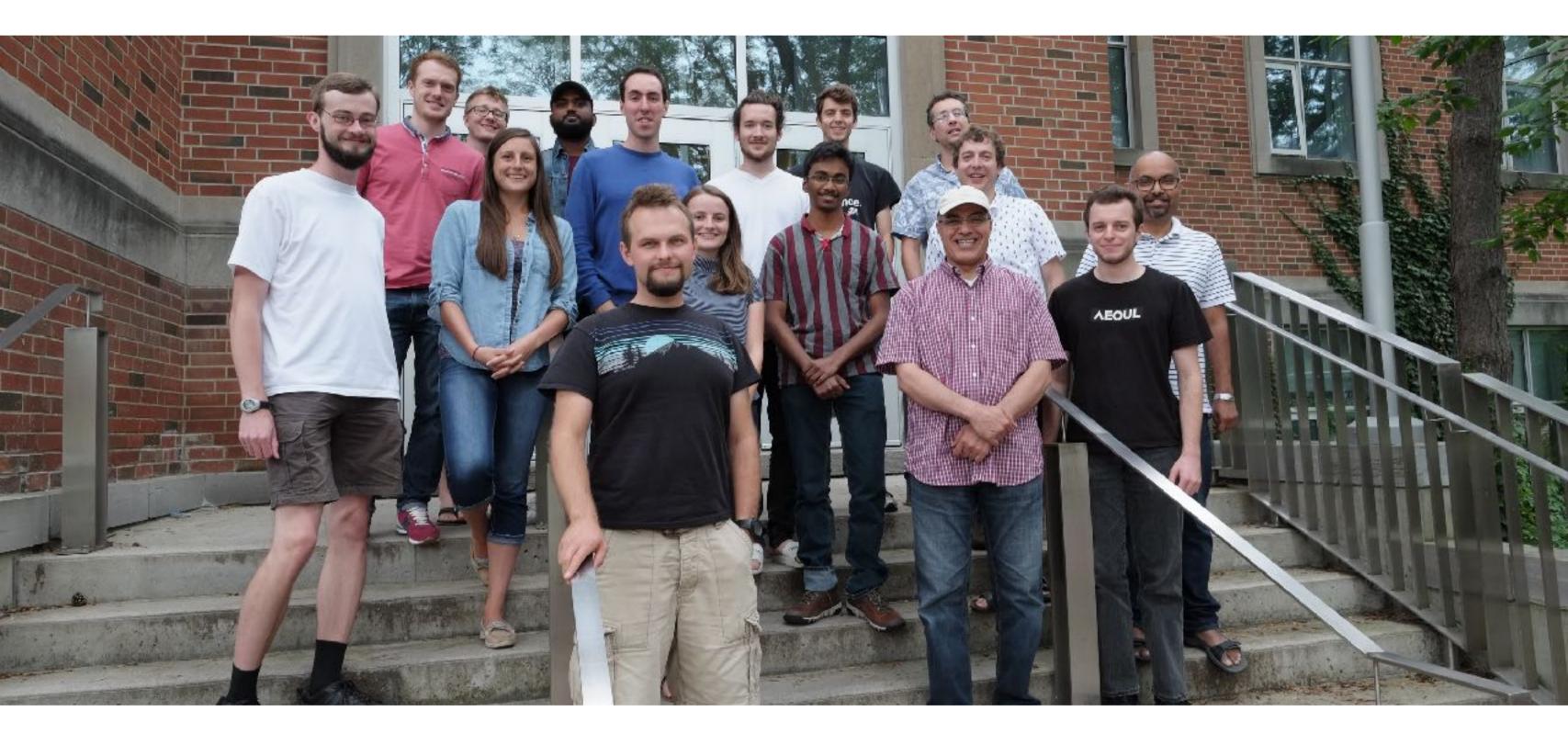
Graham Taylor

University of Guelph, Vector Institute for Artificial Intelligence and Canadian Institute for Advanced Research

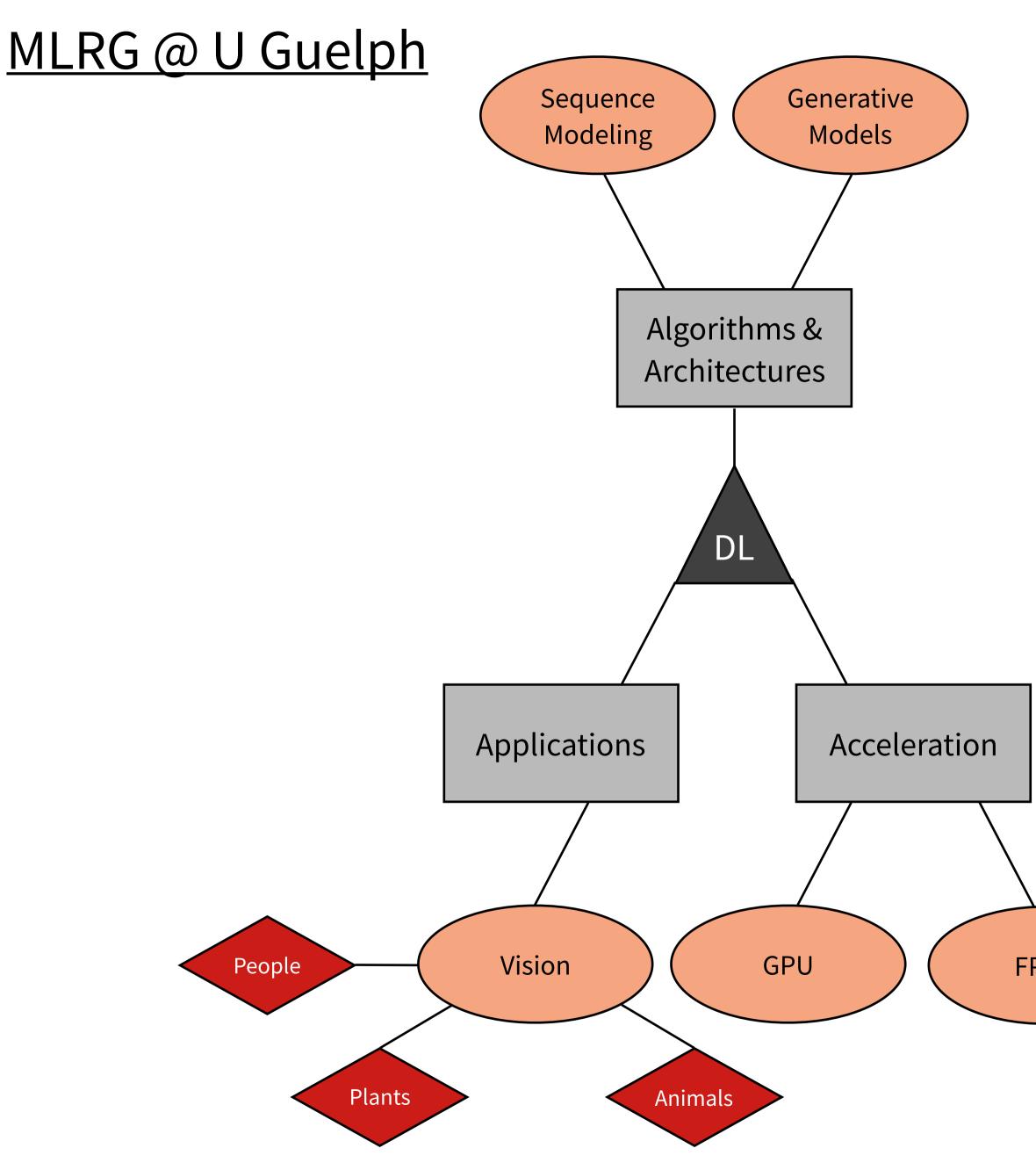


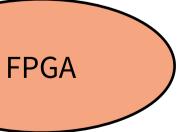
Letter Field by Judson Rosebush, 1978

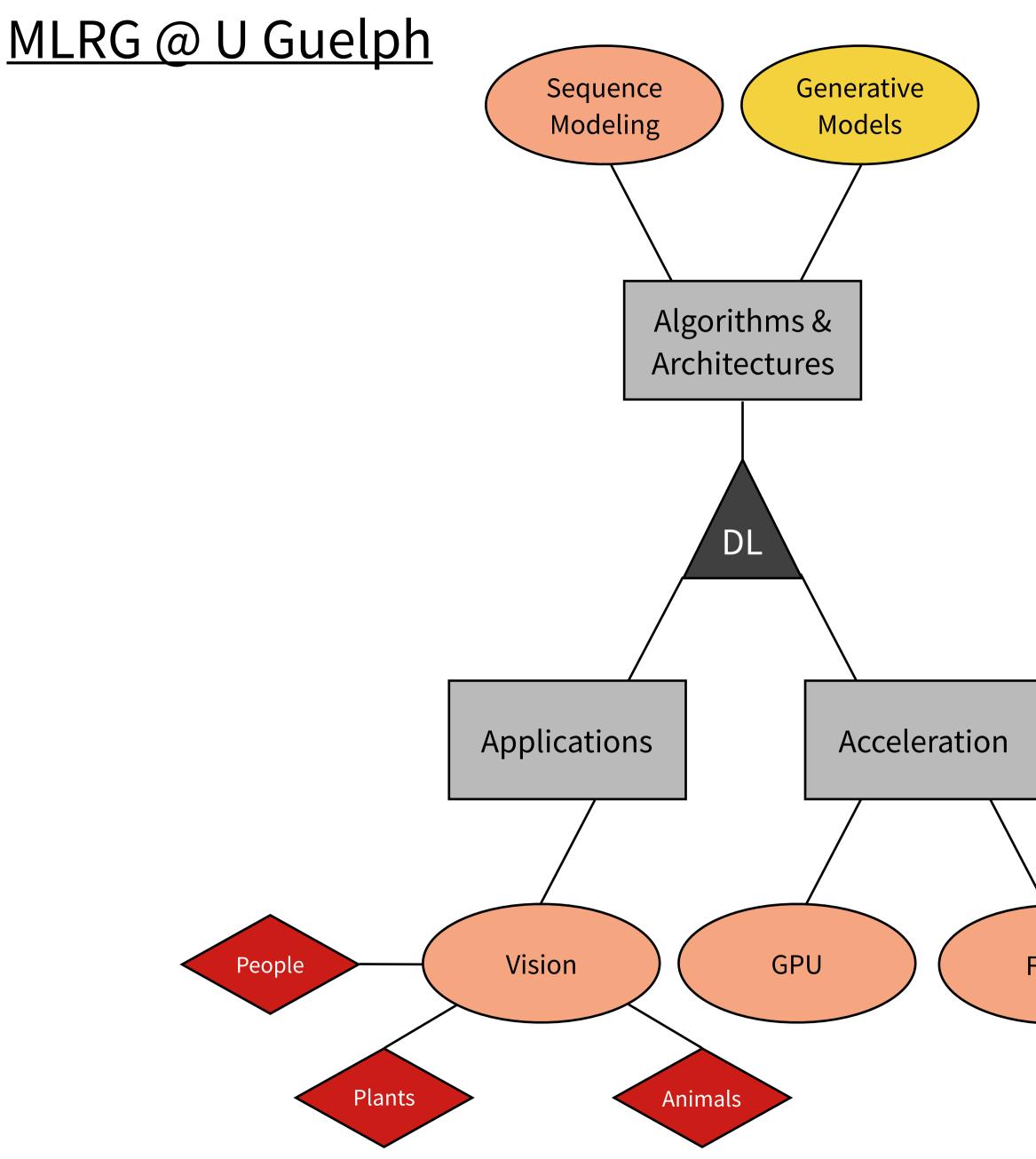
Research group: MLRG @ U of G

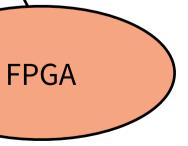


10 July 2018 / 2 Canada-France-Iceland · Design to Search / G Taylor









smartgeometry 2018 Machine Minds

Workshops

9am - 9pm, Monday May 7 - Thursday May 10, Daniels Graduate Studio

Al strategies for space frame d	lesign
behavioural enviro[NN]ments	Sam Joyce, Kate Jeffery, Jonathan Irawan, Verina
data mining the city	Danil Nagy, Jim Stoddart, Lorenz
fibrous timber joints	Dominga Garufi, Hans Wagner, Tobias Schwinn, Dyl
fresh eyes	Kyle Steinfeld, Kat Park, Adam Menges, Samatha
inside the black box	Mark Cichy. Ult
materials as probes	Billie Faircloth. Christopher Connock. Ryan Welch. Patric
mind ex machina	Panagiotis Michalatos. Nono Martinez Alonso. Jose Luis Garcia del
soft office	Giulio Brugnaro, Brian Ringley, Maria Y
sound and signal	John Granzow, Catie Newell, K

Conference

9am - 6pm, Friday May 11, Daniels Principal Hall

David Benjamin	
Graham Taylor	Vector Institute / Next Al
Dan Price	adidas Future Team
Alex Tessier	Autodesk Research
Dimitrie Stefanescu	Speckle.Works
Raffaello D'Andrea	ETH Zurich
Kate Jeffery	Jeffery Lab UCL
Jeremy Hadall Manufa	cturing Technology Centre
Workshop Presentations	SG Cluster Champions

Exhibition Opening

D

ĀN

IELS

6pm - 10pm, Friday May 11, Daniels Grad Studio

9am - 6pm, Saturday May 12, Daniels Principal Hall

Christian Derix	SUPERSPACE Wood
	Kara
Elissa Ross	MESH Ge
Daniel Davis	
Christopher Connock	KieranTim
Matt Jezyk	Αι
	Philip Beesley A
Mark Cichy	
Alex Josephson	Pa
Eric Burry	Ever
Nick Puckett	
Molly Steenson	Carnegie
Mickey McManus	

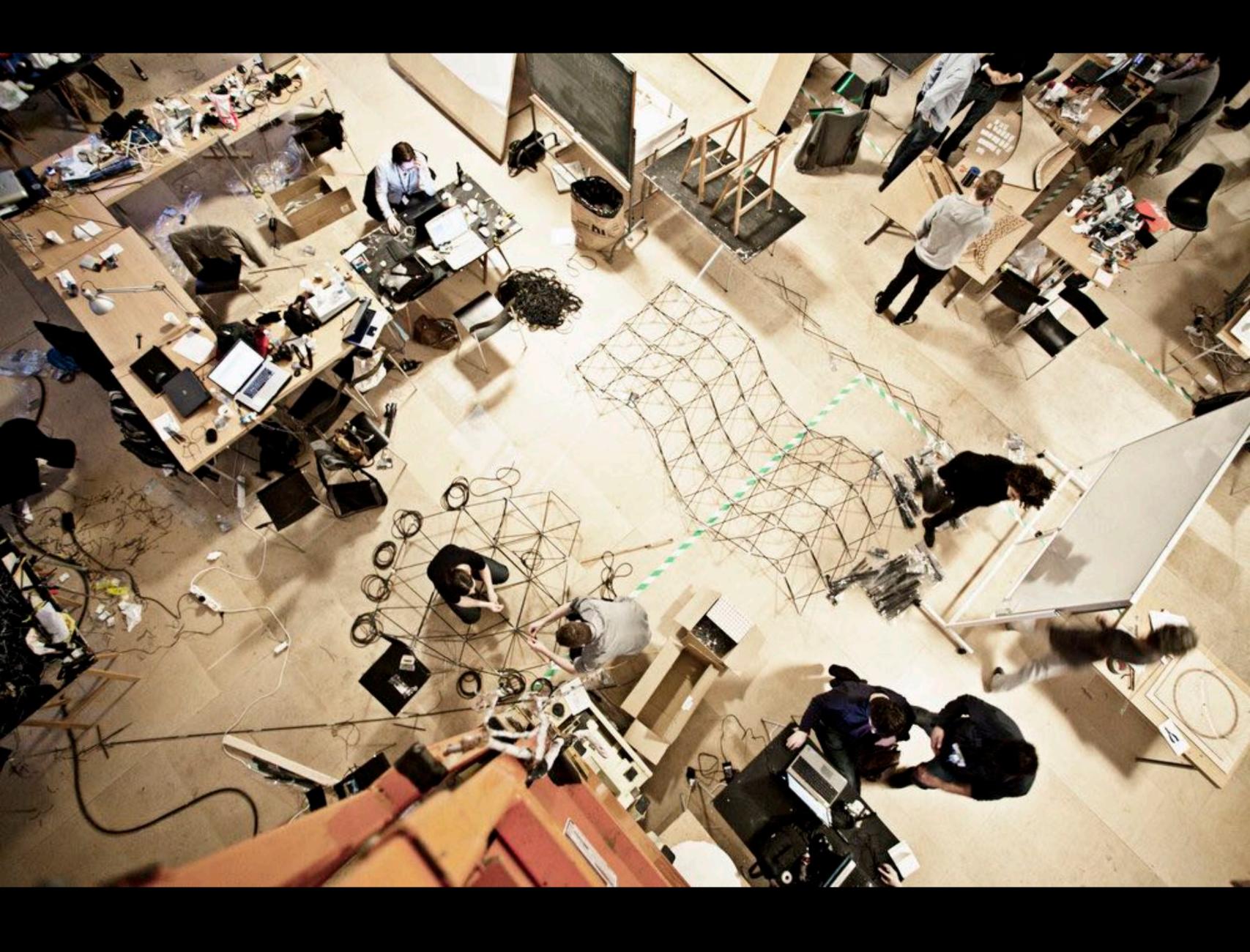
www.smartgeometry.org





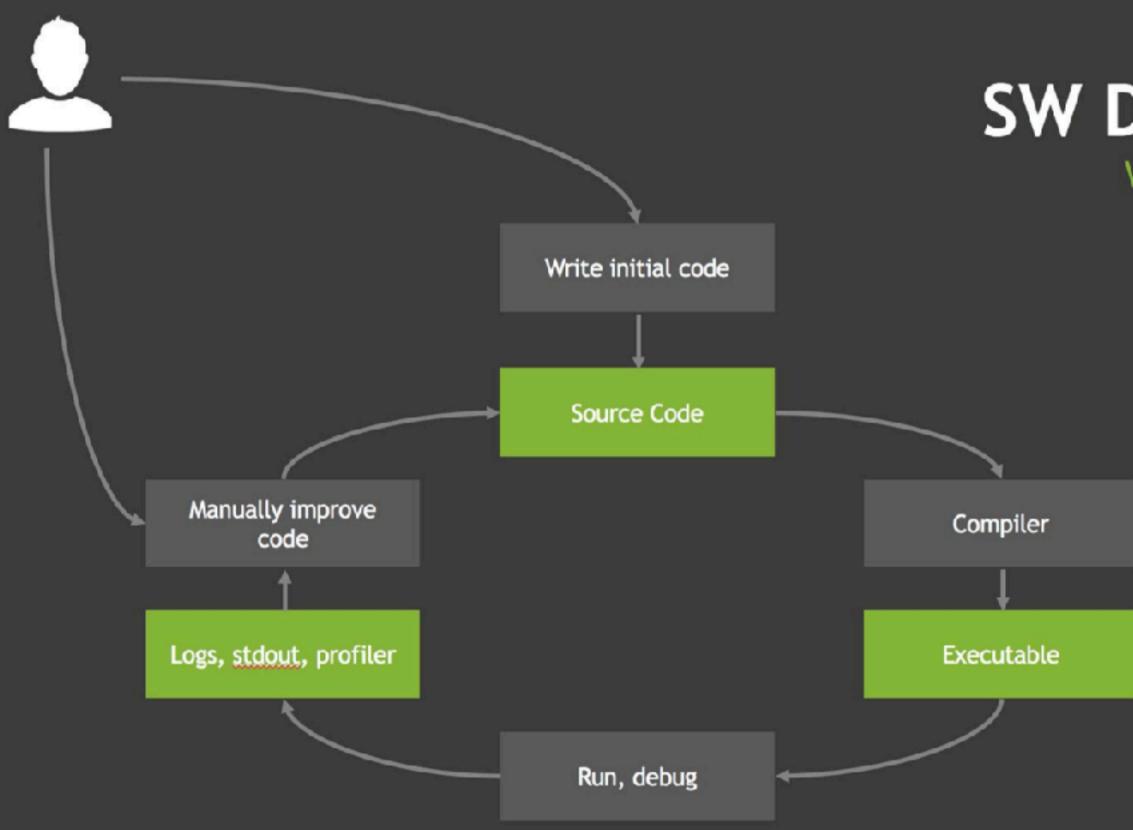
ep Aksoz a Christie zo Villagi lan Wood na Walker tan Byrne ick Weiss el Castillo Yablonina Kim Harty

is Bagot mba 3D eometry WeWork nberlake utodesk rchitect Dialog artisans ntscape OCAD e Mellon Design









SW DEV PROCESS Write code, compile, test, debug, repeat...

via Clement Farabet



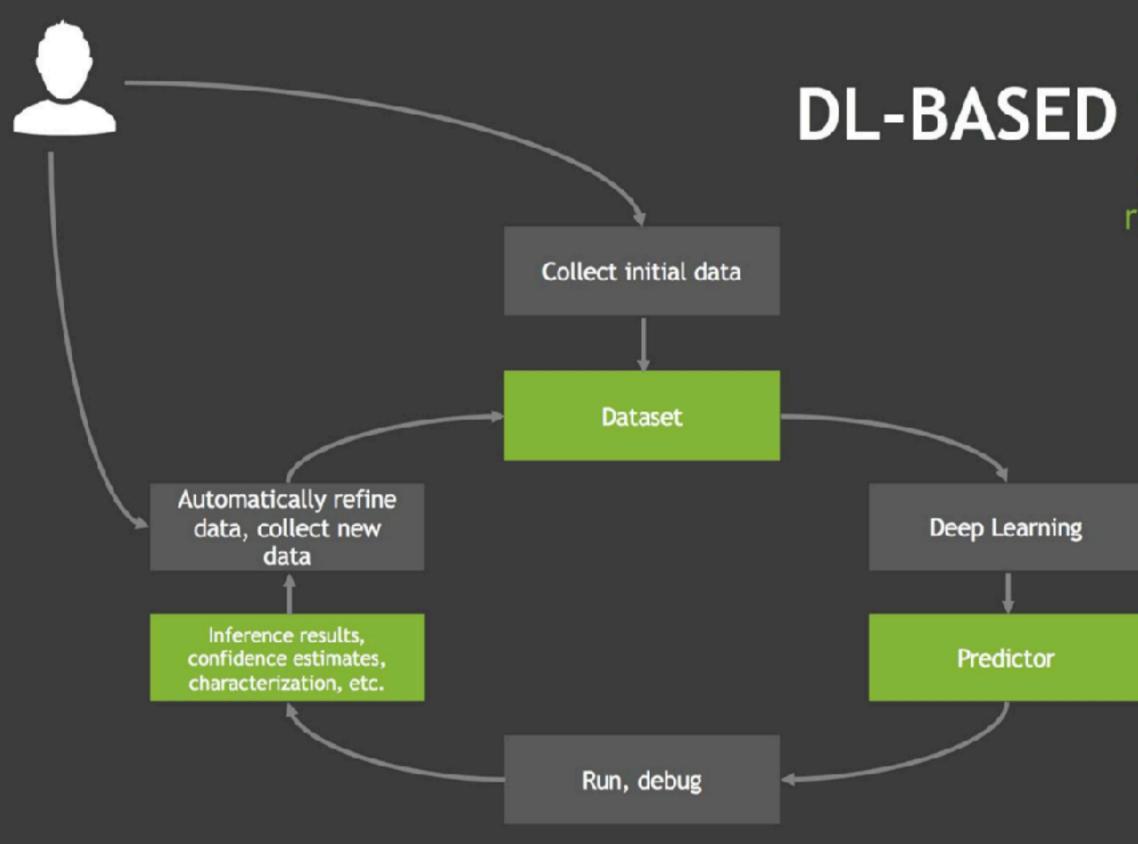
Source code	 Data
Compiler	 Dee
Executable	 Prec

а

ep Learning

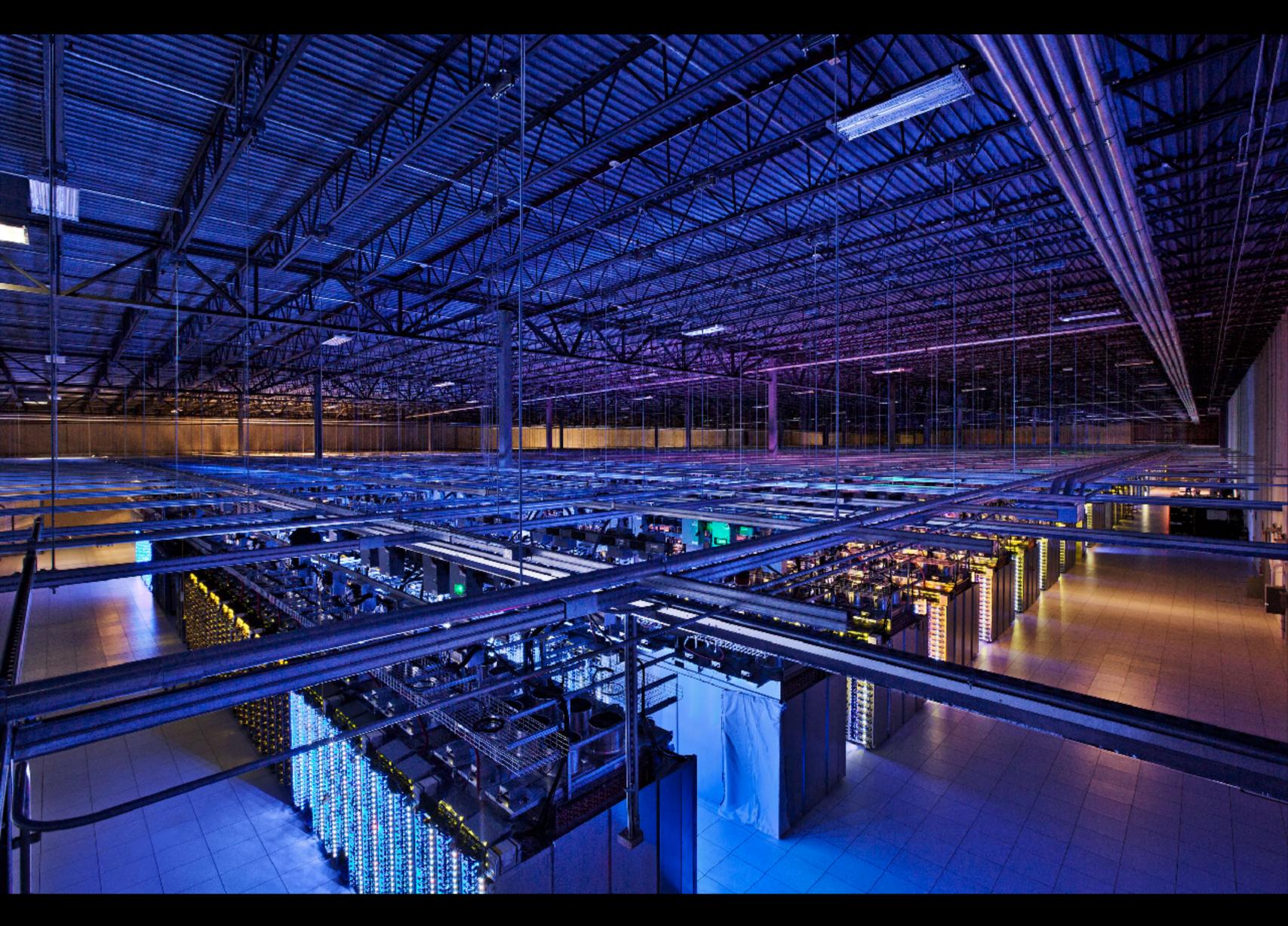
dictor

via Clement Farabet

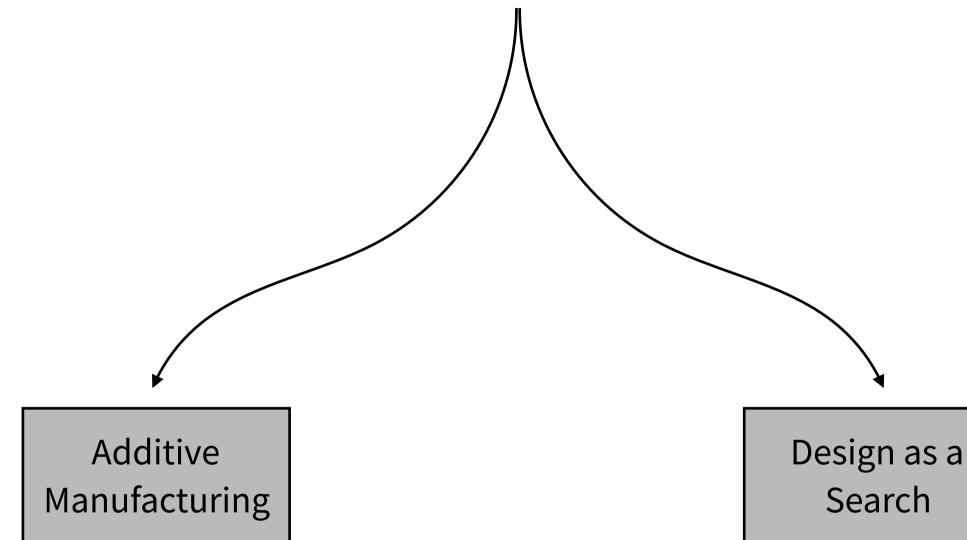


DL-BASED SW PROCESS Collect initial data, train, run/debug, mine new data

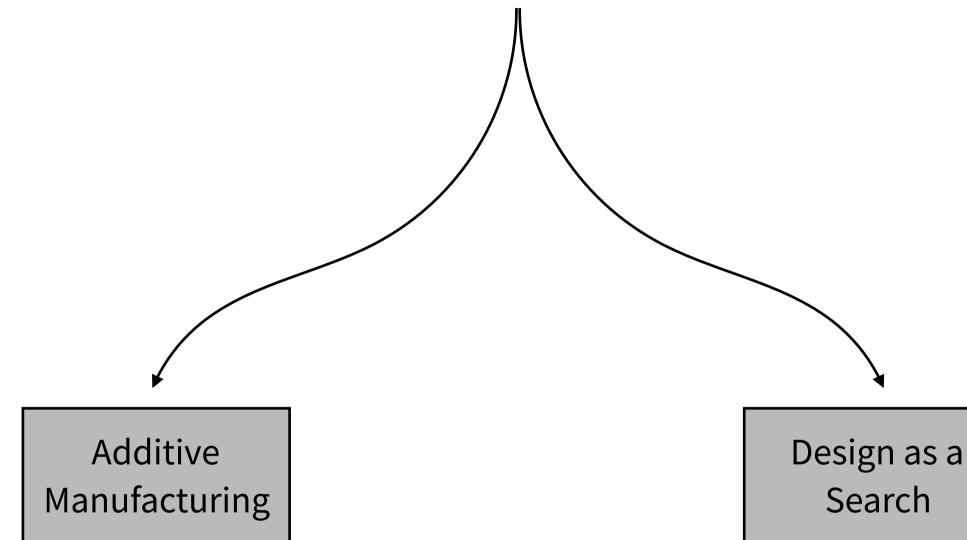
via Clement Farabet



via <u>wired.com</u>

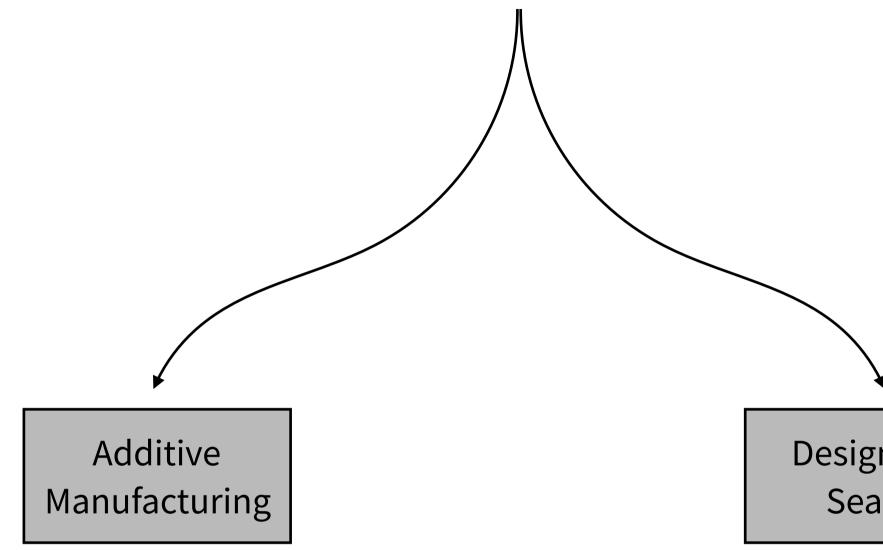


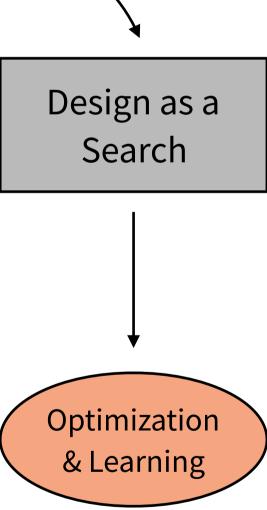
Search



Search







Medium



Andrej Karpathy Follow Director of AI at Tesla. Previously Research Scientist at OpenAI and PhD student at Stanford. I like to train deep neural nets on large datasets. Nov 11, 2017 · 8 min read

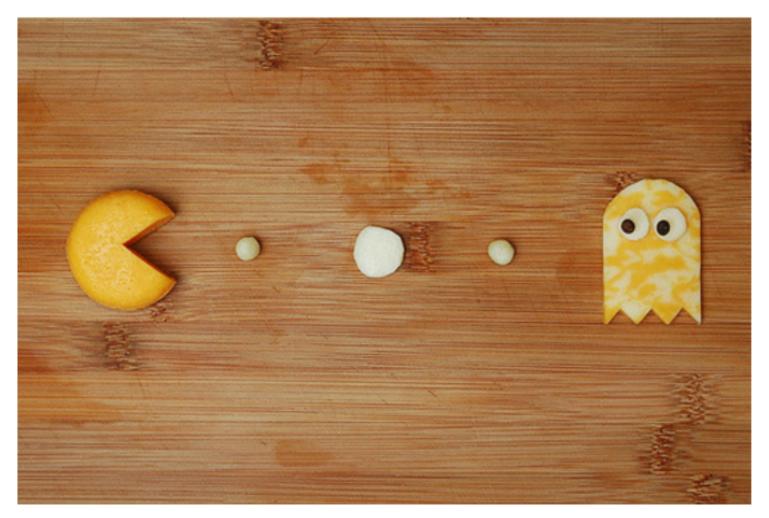
Software 2.0

I sometimes see people refer to neural networks as just "another tool in your machine learning toolbox". They have some pros and cons, they work here or there, and sometimes you can use them to win Kaggle competitions. Unfortunately, this interpretation completely misses the forest for the trees. Neural networks are not just another classifier, they represent the beginning of a fundamental shift in how we write software. They are Software 2.0.

The "classical stack" of **Software 1.0** is what we're all familiar with—it is written in languages such as Python, C++, etc. It consists of explicit instructions to the computer written by a programmer. By writing each line of code, the programmer is identifying a specific point in program space with some desirable behavior.

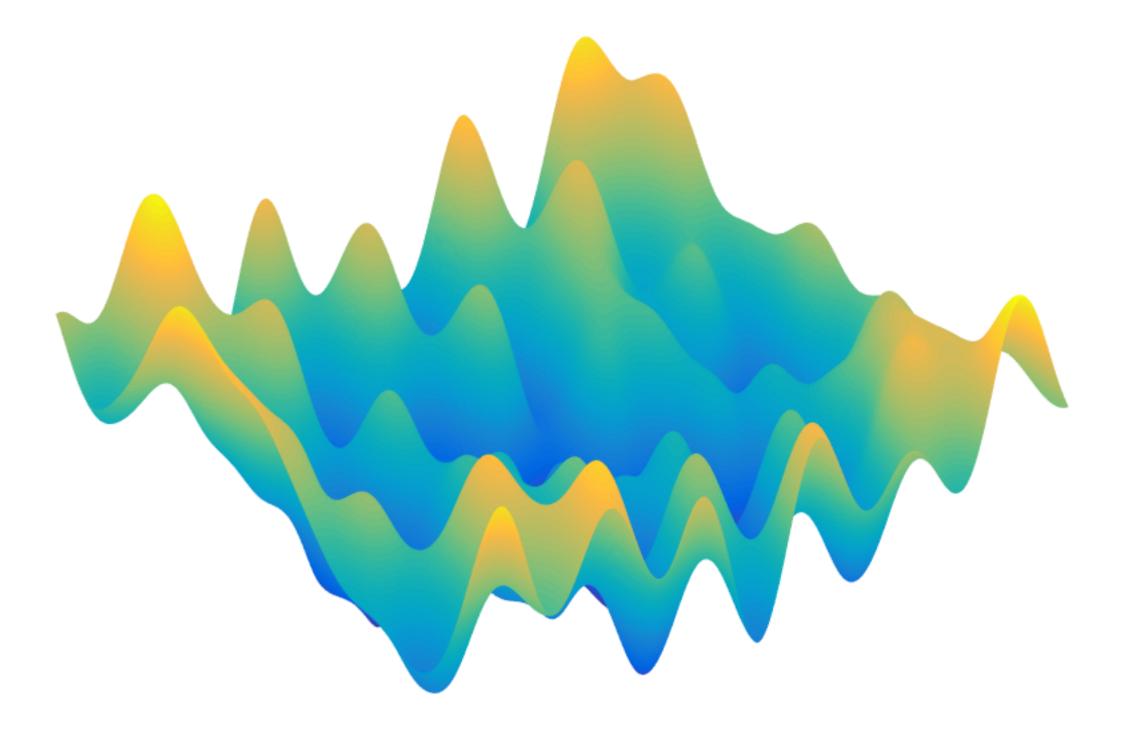
Deep Learning is Eating Software

NOVEMBER 13, 2017 By Pete Warden in UNCATEGORIZED 22 COMMENTS



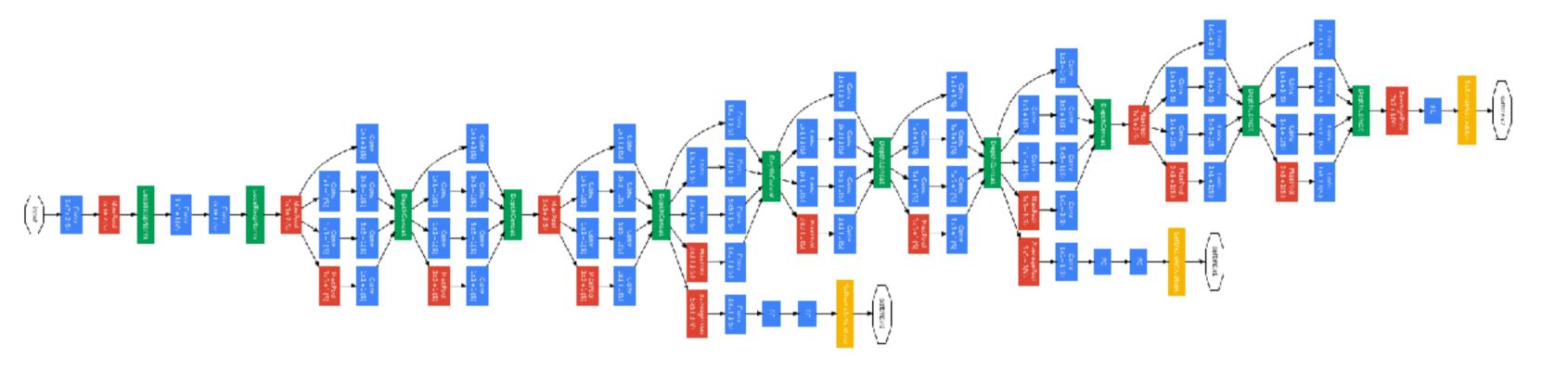


Optimization vs. Learning



via Yuejie Chi

Optimization in Machine Learning



Parameter search:

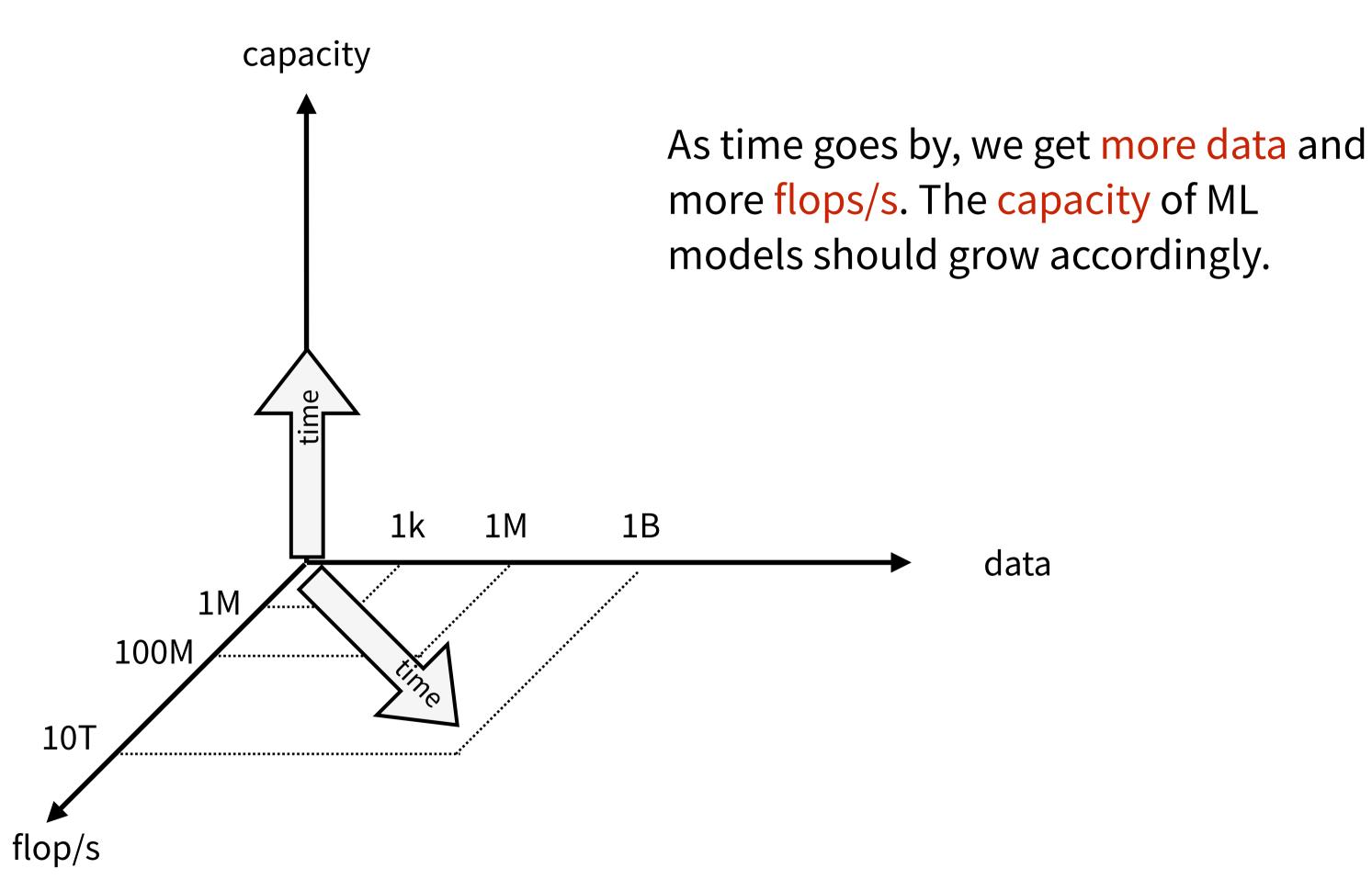
- Knobs inside the boxes lacksquare
- 100M 1B parameters
- Core of ML

Model search:

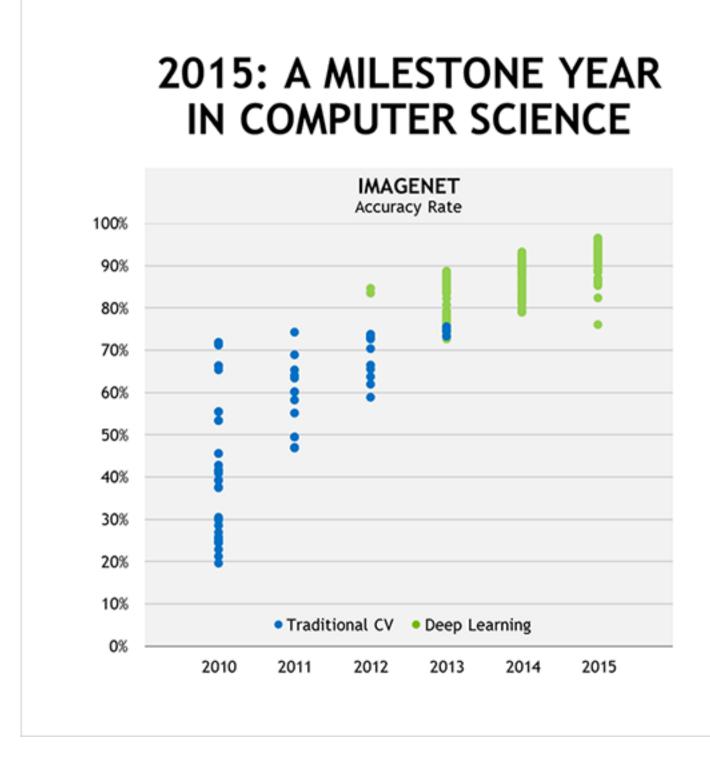
- •
- ullet

Typically fixed architecture 10 - 100* *hyperparameters* Traditionally a human task

GoogLeNet (circa 2014)

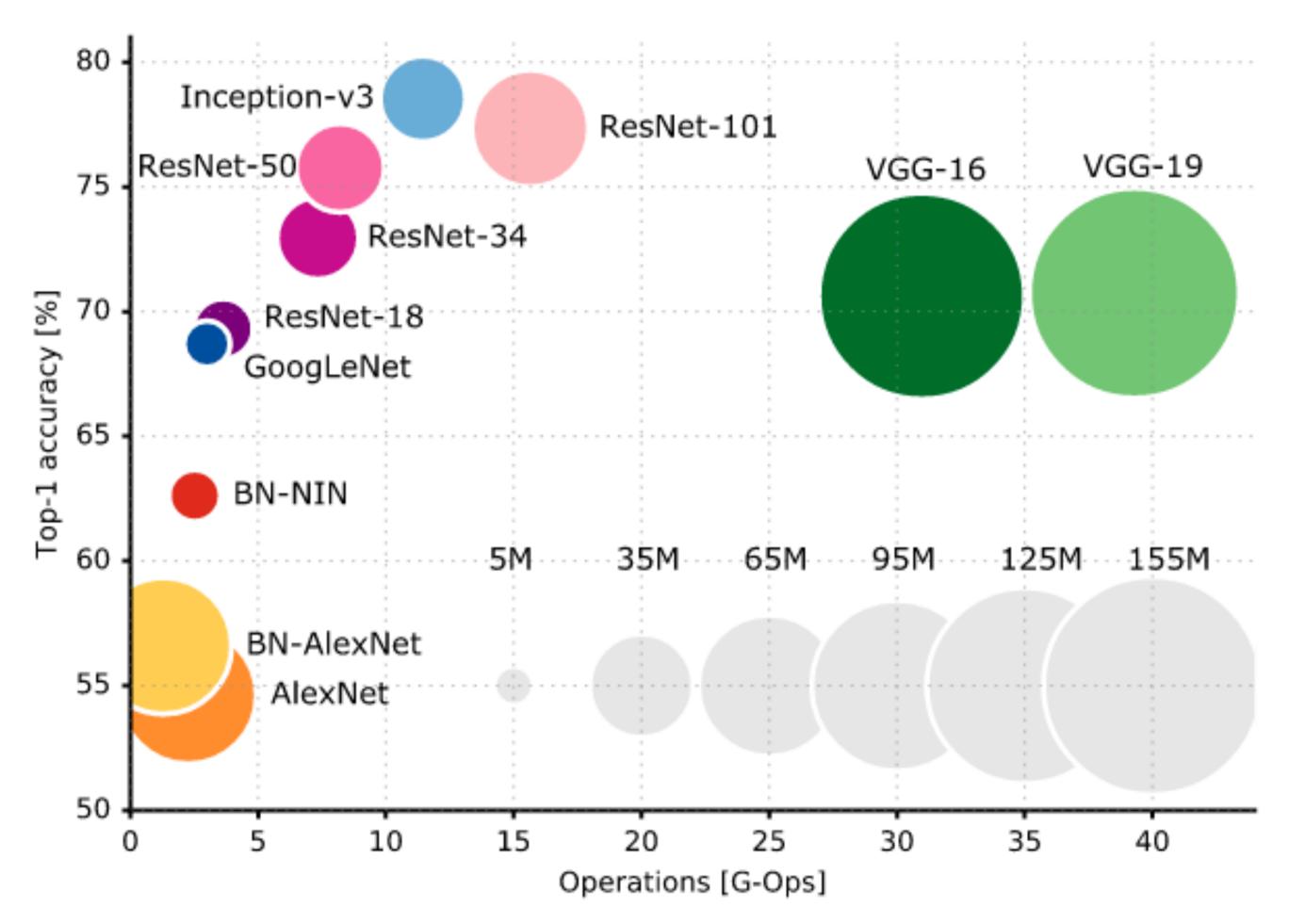


inspired by Marc'Aurelio Ranzato



via NVIDIA

Saturation?



Eugenio Culurciello's blog: https://culurciello.github.io/tech/2016/06/04/nets.html

Learning architectures

« Smerity.com

- Romanticized notion of DL - end of feature engineering
- Feature engineering • has decreased
- Architectures have become more complex

10 July 2018 / 24 Canada-France-Iceland · Design to Search / G Taylor the new feature engineering

Two of the most important aspects of machine learning models are feature extraction and feature engineering. Those features are what supply relevant information to the machine learning models.

Representing the word **overfitting** using various feature representations:

Morphological = [(prefix, over-), (root, fit), (suffix=imperfect tense, -ing)] **#** Unigrams = ['o', 'v', 'e', 'r', 'f', 'i', 't', 't', 'i', 'n', 'g'] **#** Bigrams = ['ov', 've', 'er', 'rf', 'fi', 'it', 'tt', 'ti', 'in', 'ng'] # Trigrams = ['ove', 'ver', 'erf', 'rfi', 'fit', 'itt', 'tti', 'tin', 'ing'] **Word vector** = [-0.26, 0.34, 0.48, -0.06, 0.16, 0.11, 0.13, -0.15, 0.47, -0.49, 0.07, -0.39, -0.13, -0.15, 0.06, 0.09] *****

If the features are few or irrelevant, your model may have a hard time making any useful predictions. If there are too many features, your model will be slow and likely overfit.

Humans don't necessarily know what feature representation are best for a given task. Even if they do, relying on feature engineering means that a human is always in the loop. This is a far cry from the future we might want, where you can throw any dataset at a

http://smerity.com/articles/2016/architectures_are_the_new_feature_engineering.html

In deep learning, architecture engineering is

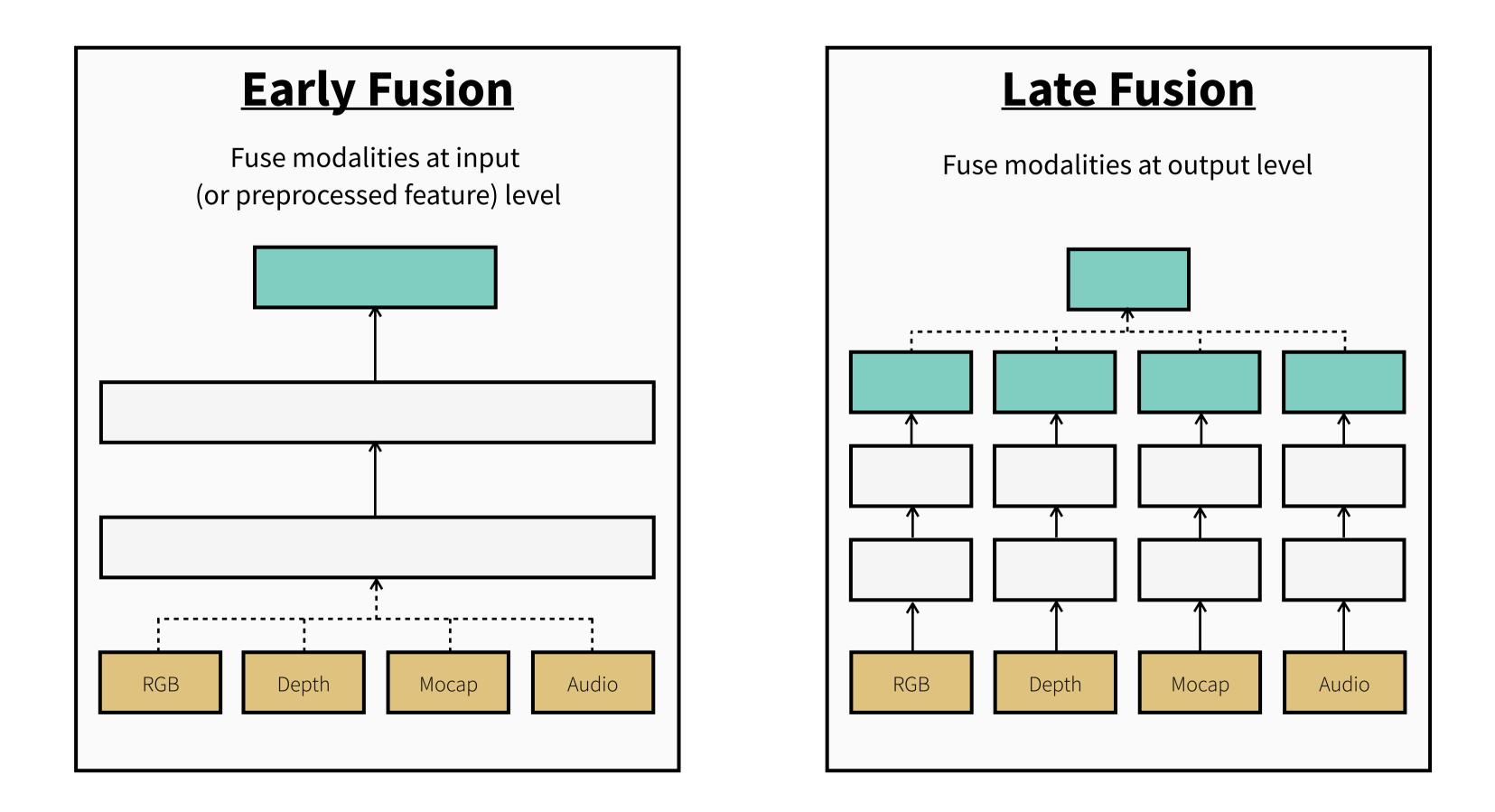
June 11, 2016

Example: Learning Multimodal Fusion for Gesture Recognition



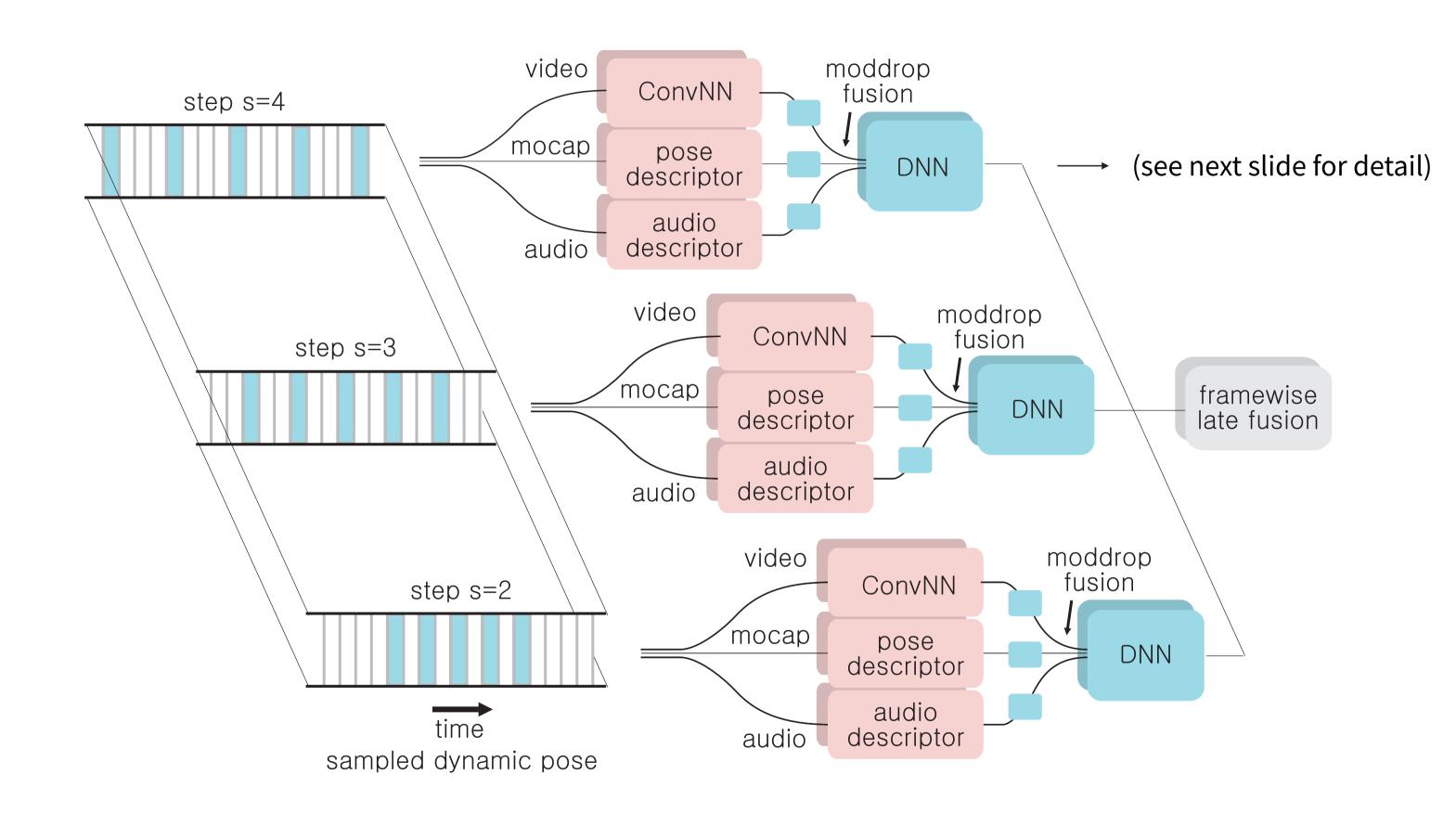
10 July 2018 / 25 Canada-France-Iceland · Design to Search / G Taylor Neverova, Wolf, Taylor (2016)

Early vs. late fusion





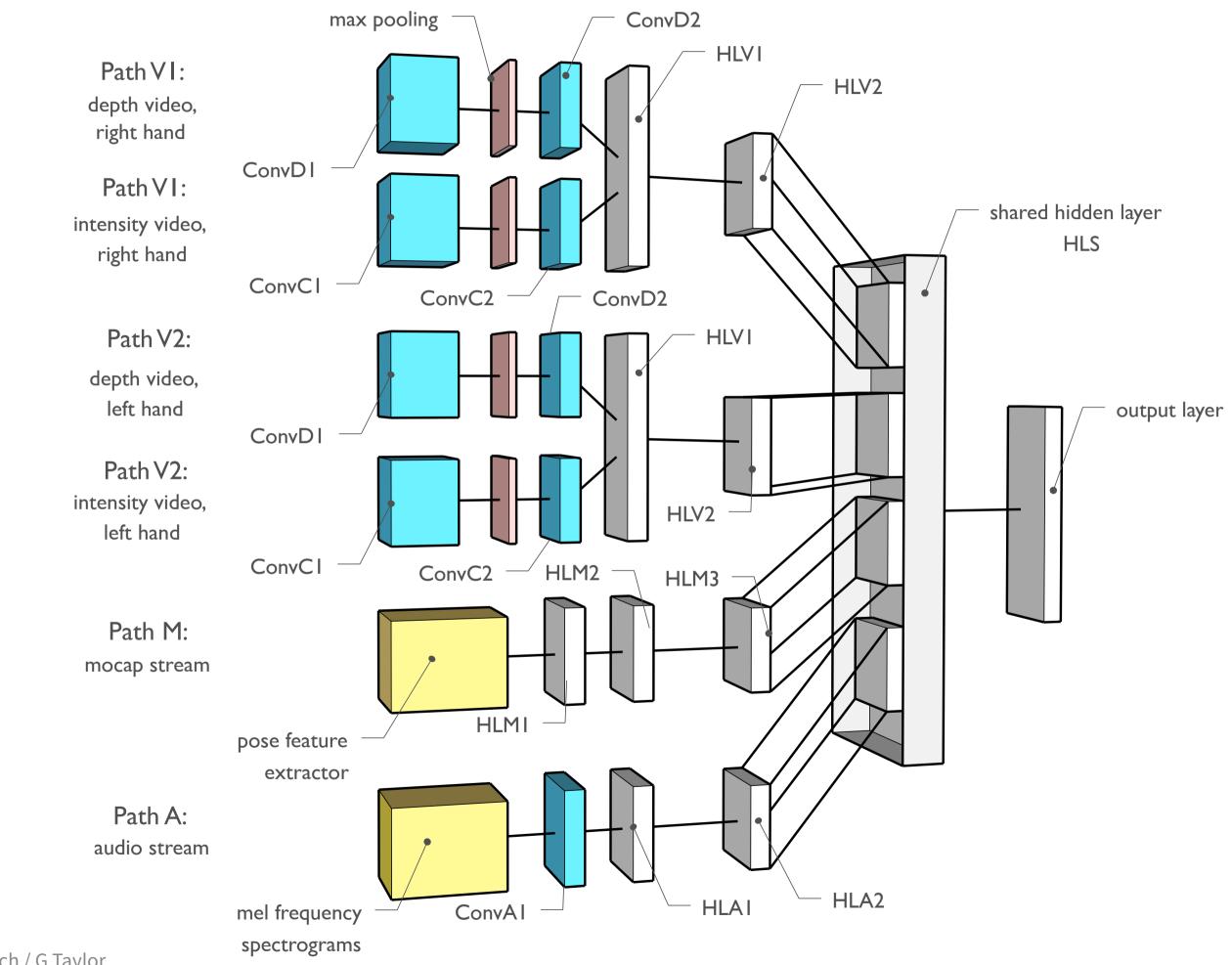
A multi-scale architecture



Operates at 3 temporal scales corresponding to dynamic poses of 3 different durations

10 July 2018 / 27 Canada-France-Iceland · Design to Search / G Taylor

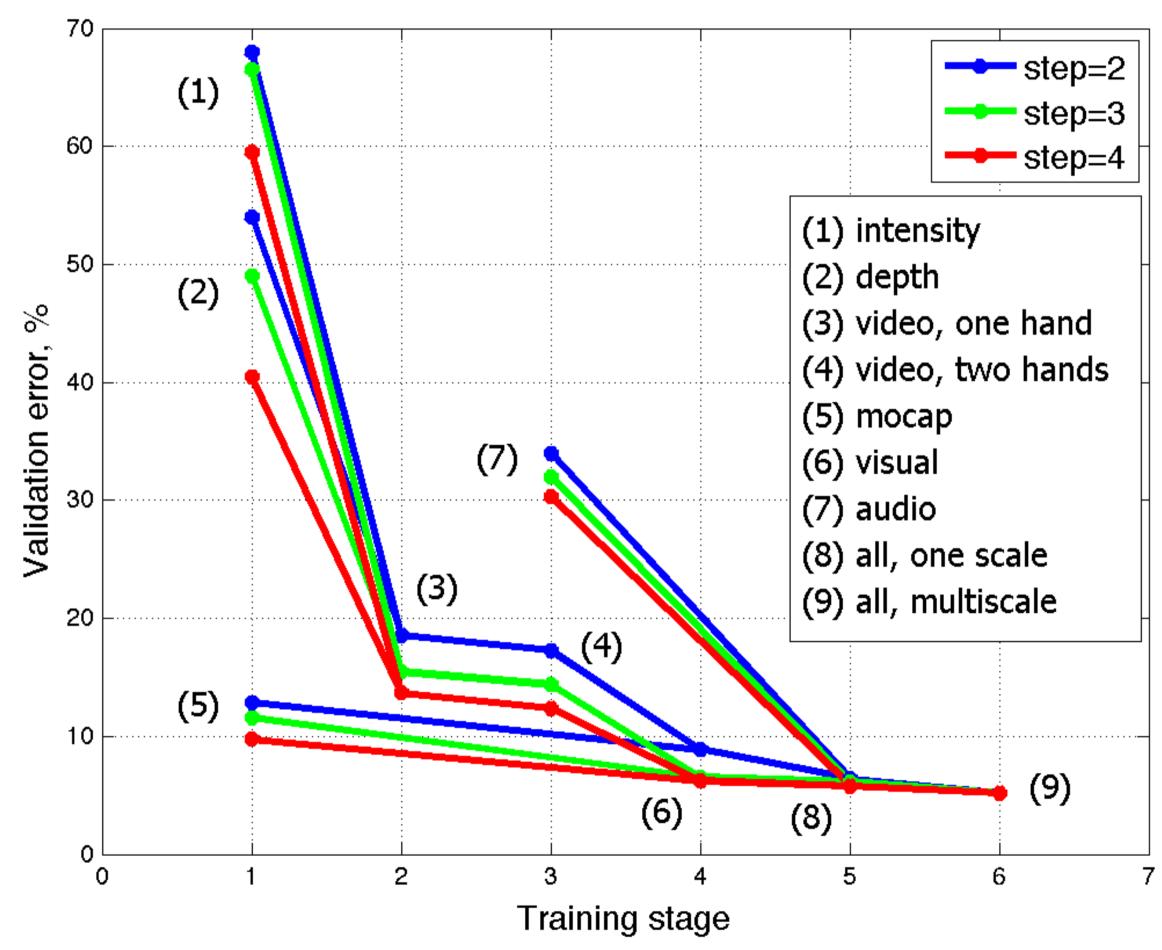
Single-scale deep architecture



10 July 2018 / 28 Canada-France-Iceland · Design to Search / G Taylor

Neverova, Wolf, Taylor (2016)

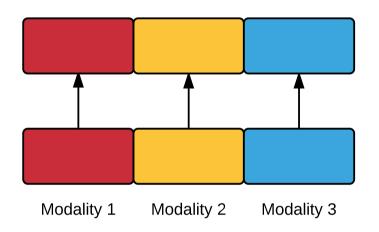
Error evolution during iterative training

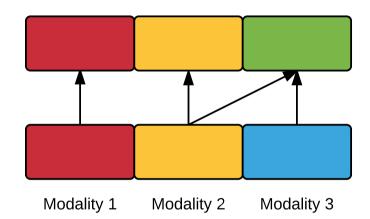


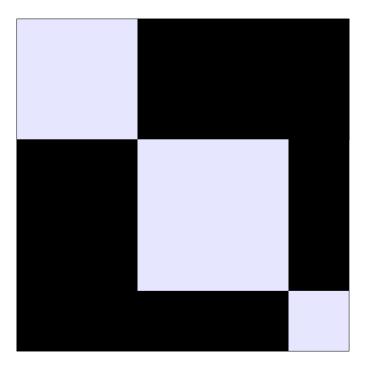
10 July 2018 / 29 Canada-France-Iceland · Design to Search / G Taylor

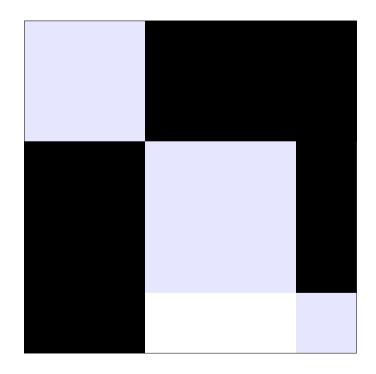
Learning Fusion Architectures (1) Three typical fusion architectures achievable by *Modout*, with

Three typical fusion architectures achievab corresponding weight masks:





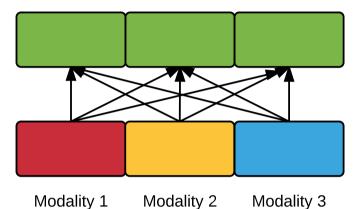


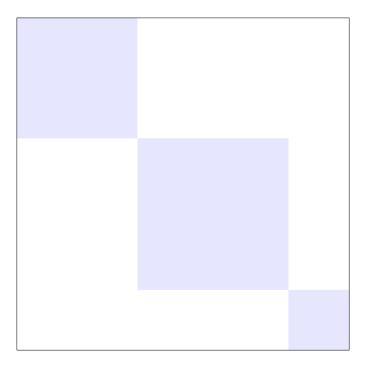


Fused

<u>Independent</u>





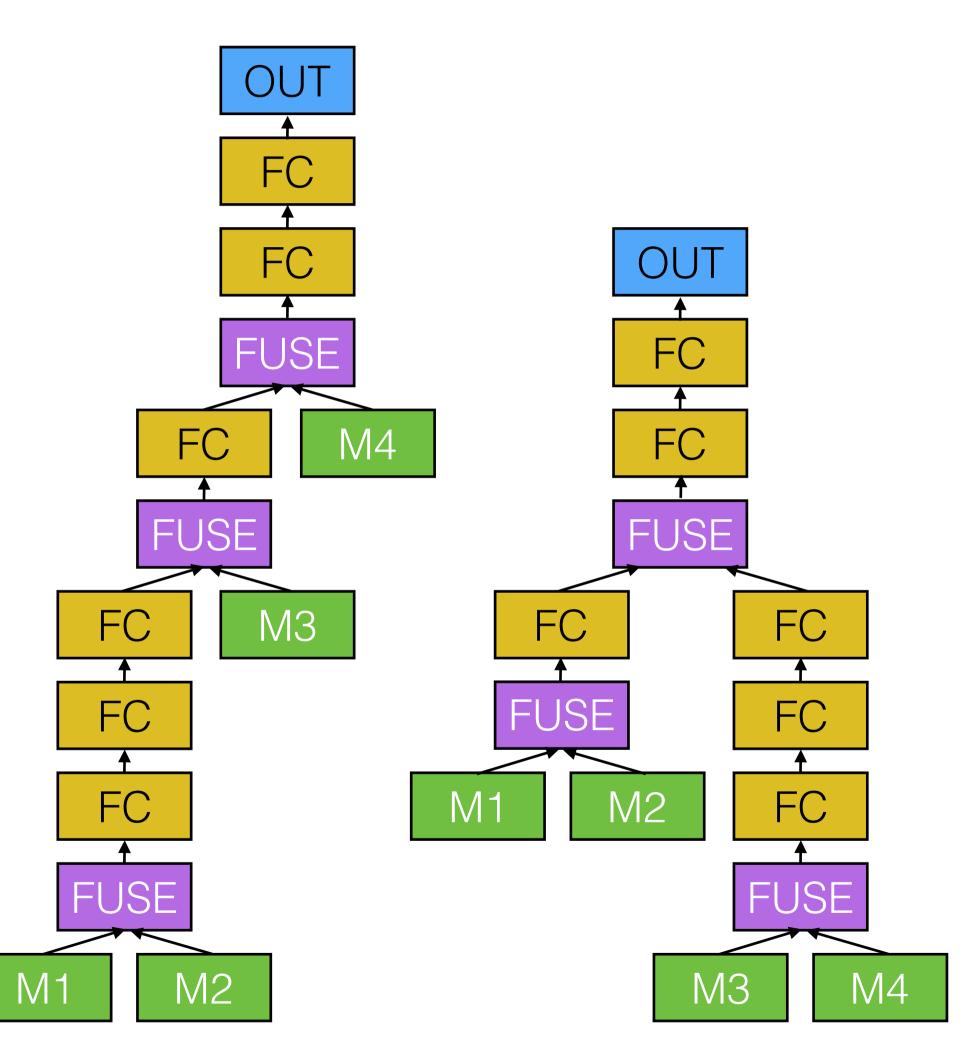


Fully-connected

Li, Neverova, Wolf, Taylor (2017)

Learning Fusion Architectures (2)

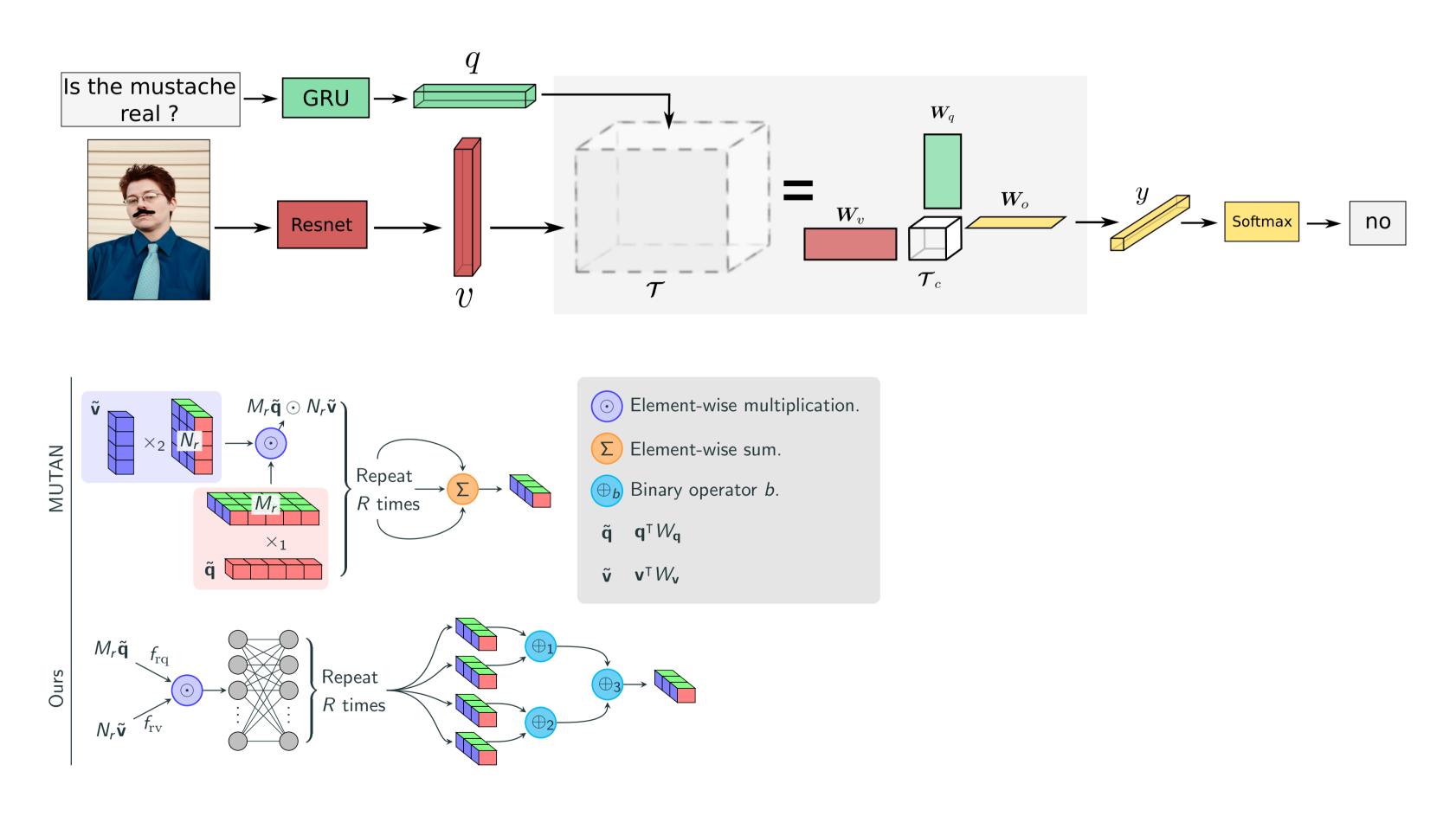
Applying *Bayesian Optimization* to Search Space Over Graphs



10 July 2018 / 31 Canada-France-Iceland · Design to Search / G Taylor

Ramachandram, Lisicki, Shields, Amer, Taylor (2018)

Learning Fusion Operators for VQA



¹⁰ July 2018 / 32 Canada-France-Iceland · Design to Search / G Taylor

<u>Duke</u>, Taylor (2018)

The unit is environmentally friendly. Equipment that contacts water is chemically inert and releases no harmful lubricants.

The top of the screw connects to a generator installed in a small service room above the waterline. Depending on the site and the screw, the amount of electricity produced cars range-from one to 300 ki owatts.

> he custom-designed screw is installed at the head alongside the existing channel

The Archimedes screw was invented as a simple, reliable pumping technology to life water in the third century BC and has been used this way for 2,000 years. But run it in everse --- positioned at a site with a water drop so that the weight of the flowing water turns the screw -- and it can power a generator to make electricity.

> Once the water exits the screw it is immediately directed back into the original channel, ensuring 100 percent of the original flow continues downstream.

¢.

The screw is designed to ensure safe passage for fish and other aquatic species. It turns at low speed and has compressible rubber bumpers at entry. In carrying the water, the screw acts more like an elevator than a blender, transferring a relatively stationary "bucket" of water down the column with each rotation. Not only is there enough space in the chamber for large fish (up to 15 centimetres wide and of any length --- in cluding eels), there is no added pressure or shear stress as exists with more conventional turbing technology.

CreenBug's screw generator works with flows ranging from 103 to 10,000 litres per second and "heads" of water (height of the drop) from one to 10 metres, with about five metres being optimal.

The power can be used by the property owner locally or, where regulations permit, sold back into the regional power grid. GreenBug's business model encourages site owners to participate as investors to share in this payback.

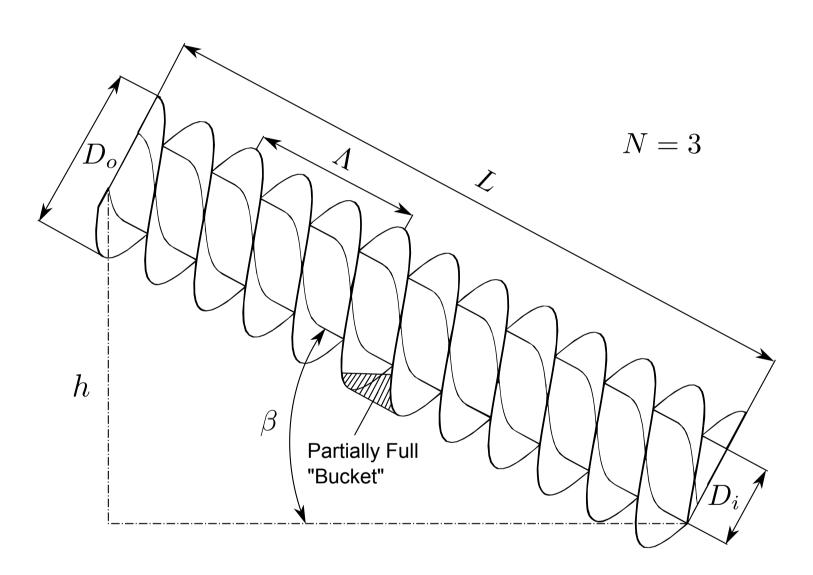
Water is diverted to the screw through a grated intake above the weir, with nopooling or flow interruption. Although flow rates vary naturally, the screw maintains its efficiency even when the volume is low.

via Canadian Geographic

Bayesian Optimization of Archimedes Screw Turbine

<u>Design parameters:</u>

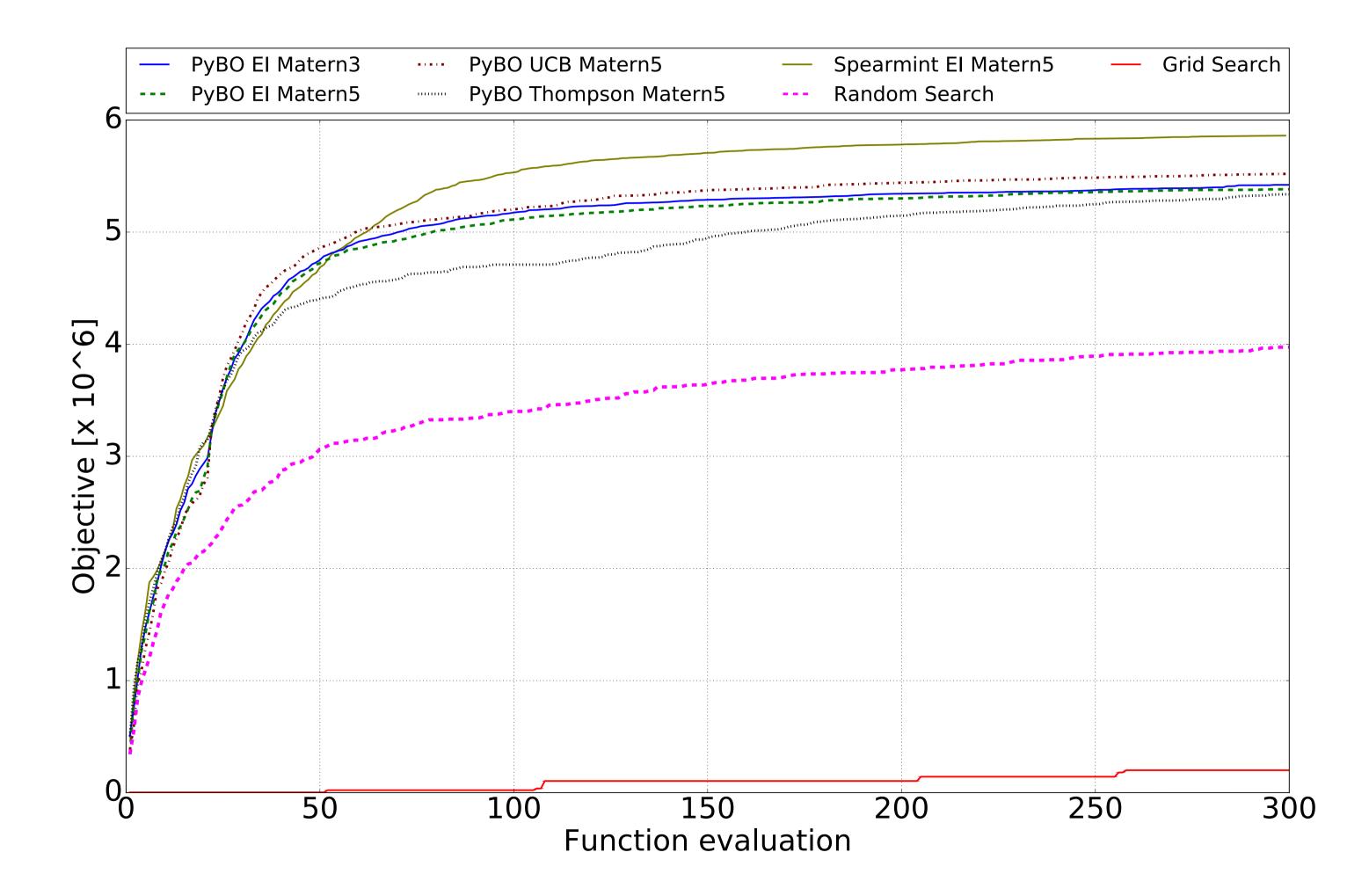
- D_i inner cylinder diameter
- D_o outer cylinder diameter
- *h* working head
- N number of flights
- β tilt angle of the screw
- Λ pitch
- $ullet f_{\max}$ maximum fill of the bucket



Optimized for power output / mass of steel



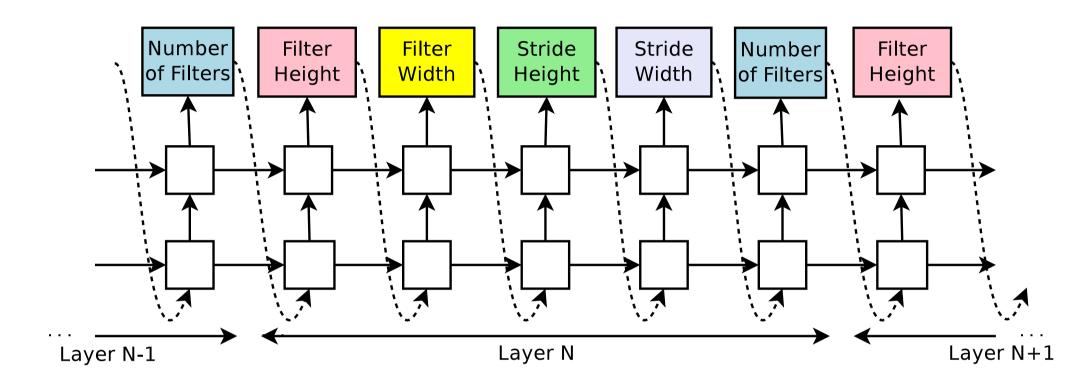
Lisicki, Lubitz, Taylor (2016)



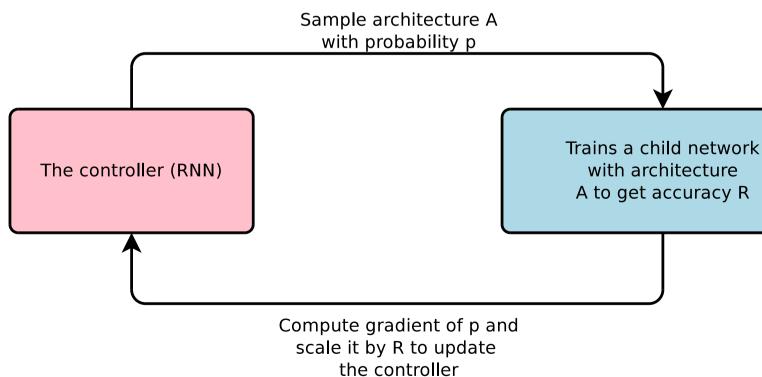
Lisicki, Lubitz, Taylor (2016)

Meta-learning: Architecture

Recurrent neural network (RNN) controller outputs architecture as string



RNN controller trained with reinforcement learning



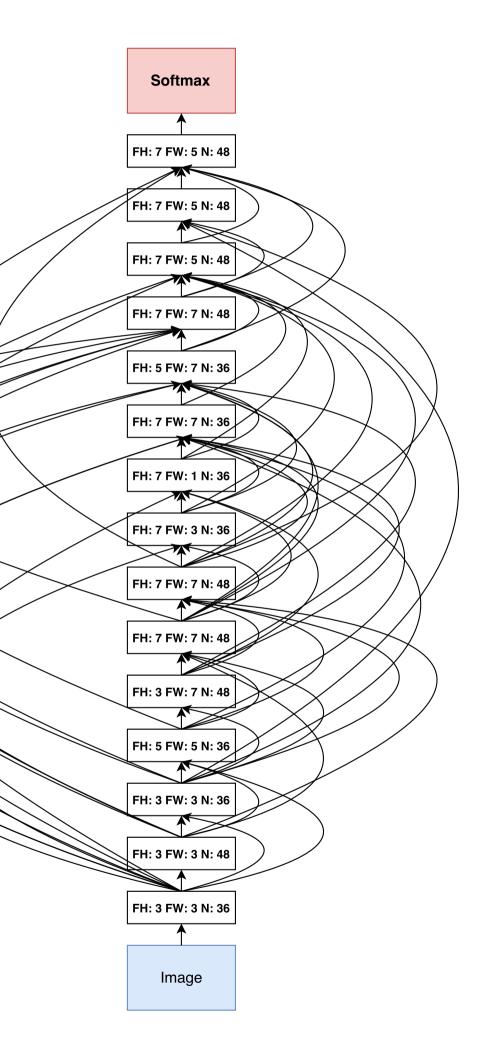
10 July 2018 / 36 Canada-France-Iceland · Design to Search / G Taylor

via Zoph and Le (2017)

Discovered CNN Architecture

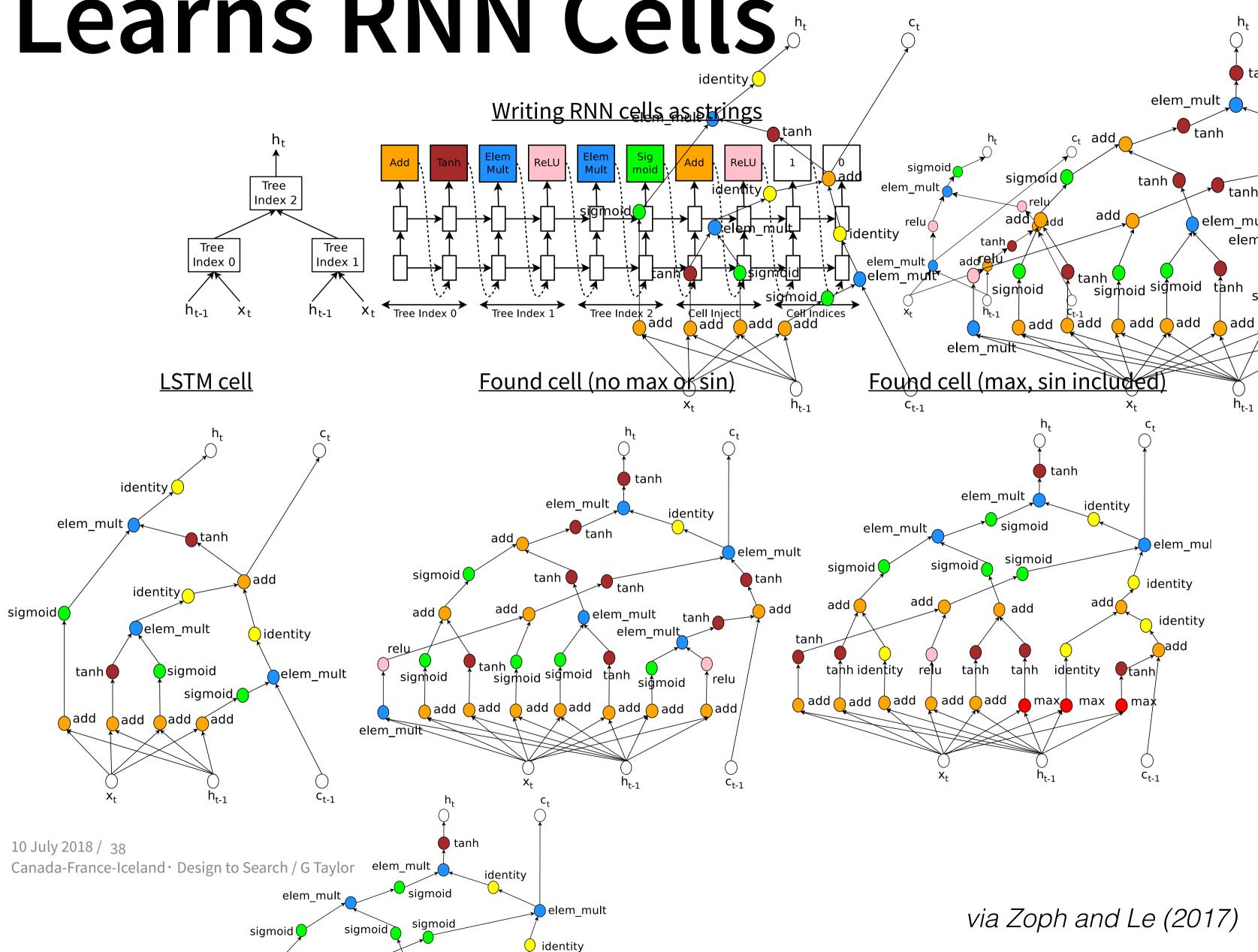
Error rate on CIFAR

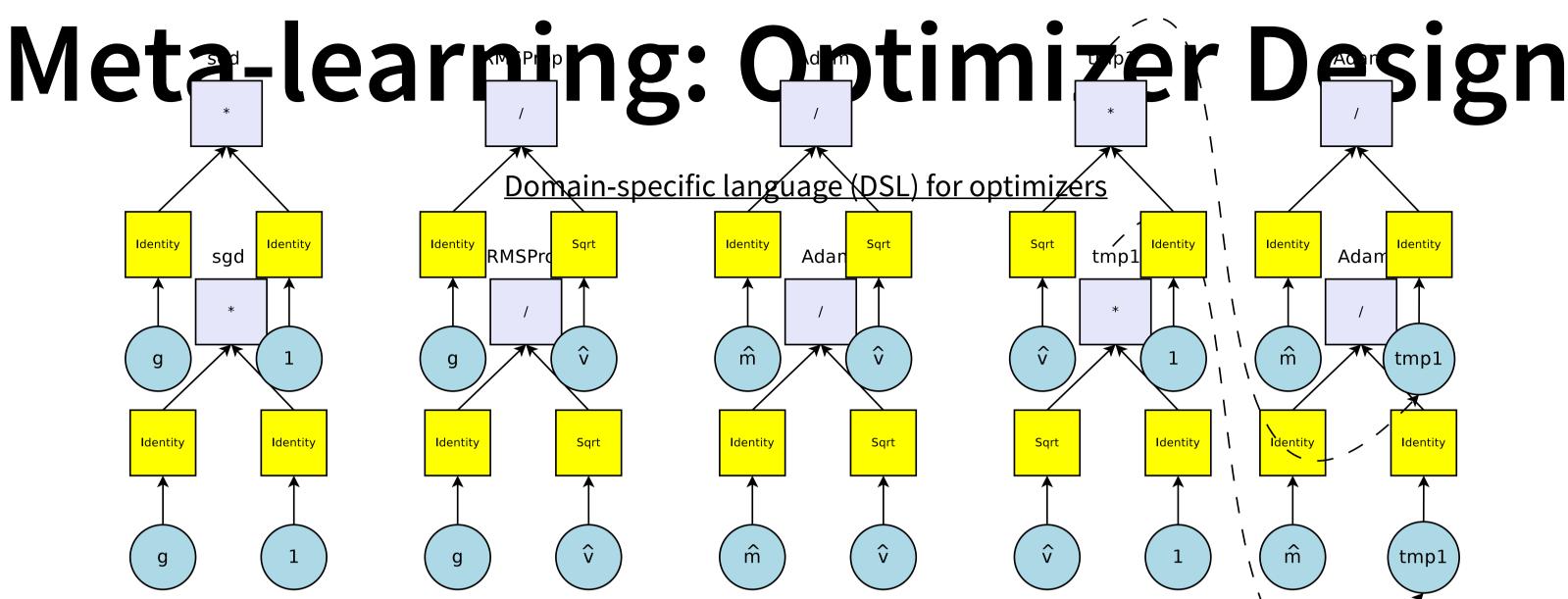
Model	Depth	Parameters	Error rate (%)
Network in Network (Lin et al., 2013)	-	_	8.81
All-CNN (Springenberg et al., 2014)	-	-	7.25
Deeply Supervised Net (Lee et al., 2015)	-	-	7.97
Highway Network (Srivastava et al., 2015)	-	-	7.72
Scalable Bayesian Optimization (Snoek et al., 2015)	-	-	6.37
FractalNet (Larsson et al., 2016)	21	38.6M	5.22
with Dropout/Drop-path	21	38.6M	4.60
ResNet (He et al., 2016a)	110	1.7M	6.61
ResNet (reported by Huang et al. (2016c))	110	1.7M	6.41
ResNet with Stochastic Depth (Huang et al., 2016c)	110	1.7M	5.23
	1202	10.2M	4.91
Wide ResNet (Zagoruyko & Komodakis, 2016)	16	11.0M	4.81
	28	36.5M	4.17
ResNet (pre-activation) (He et al., 2016b)	164	1.7M	5.46
	1001	10.2M	4.62
DenseNet $(L = 40, k = 12)$ Huang et al. (2016a)	40	1.0M	5.24
DenseNet(L = 100, k = 12) Huang et al. (2016a)	100	7.0M	4.10
DenseNet $(L = 100, k = 24)$ Huang et al. (2016a)	100	27.2M	3.74
DenseNet-BC ($L = 100, k = 40$) Huang et al. (2016b)	190	25.6M	3.46
Neural Architecture Search v1 no stride or pooling	15	4.2M	5.50
Neural Architecture Search v2 predicting strides	20	2.5M	6.01
Neural Architecture Search v3 max pooling	39	7.1M	4.47
Neural Architecture Search v3 max pooling + more filters	39	37.4M	3.65



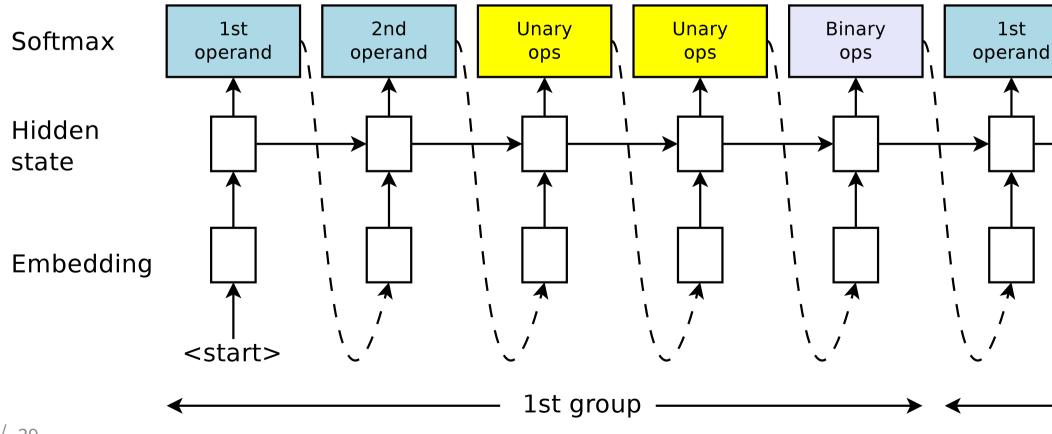
via Zoph and Le (2017)

Learns RNN Cells

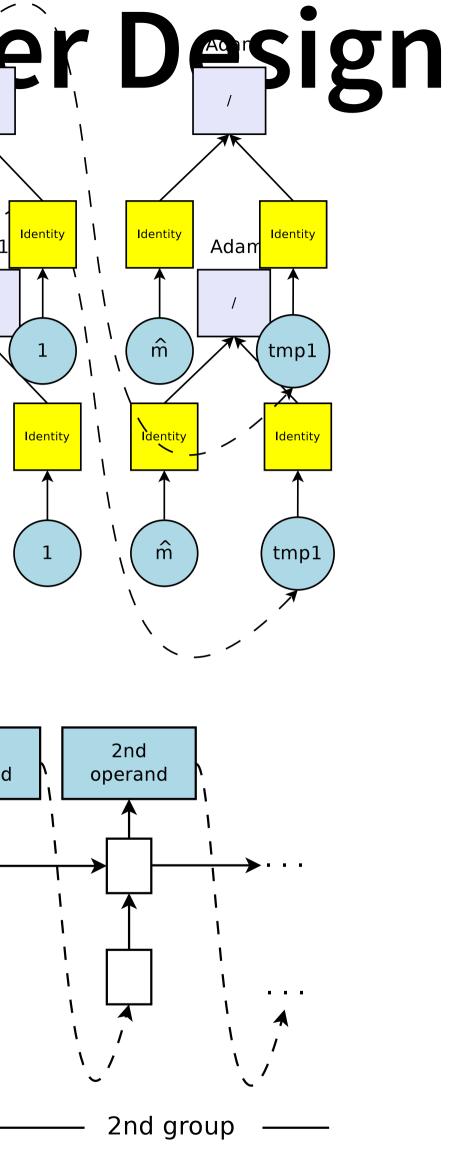




RNN controller outputs optimizer as string



10 July 2018 / 39 Canada-France-Iceland · Design to Search / G Taylor



via Zoph and Le (2017)

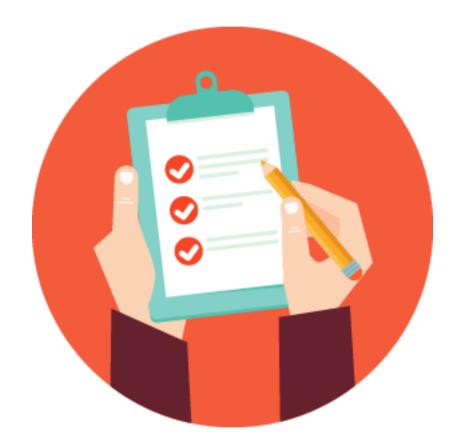


via medium.com



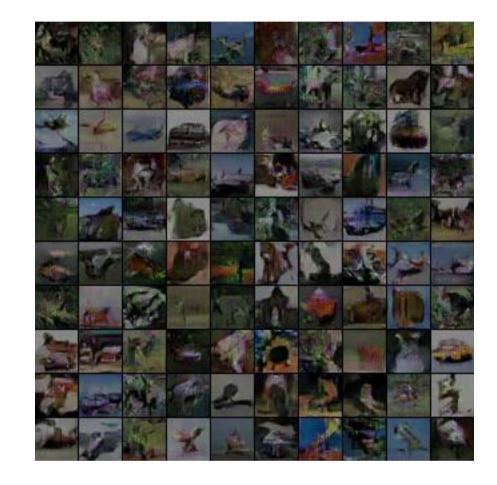


Evaluation?



10 July 2018 / 42 Canada-France-Iceland · Design to Search / G Taylor

Inception Score?





IS = 6.45

(Higher is better)

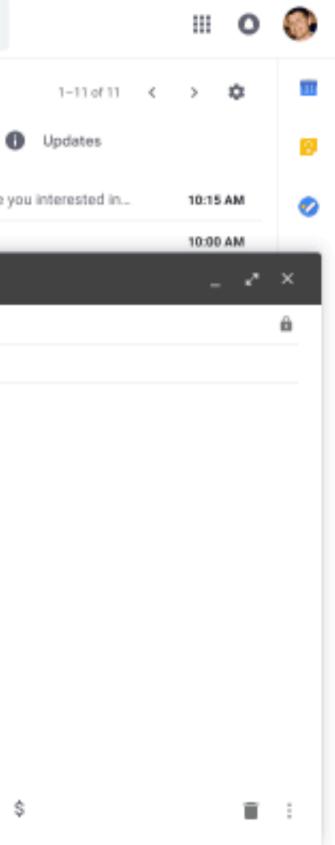
10 July 2018 / 43 Canada-France-Iceland · Design to Search / G Taylor

IS = 6.31

Im, Ma, Taylor, Branson (2018)

Gmail: Smart Compose

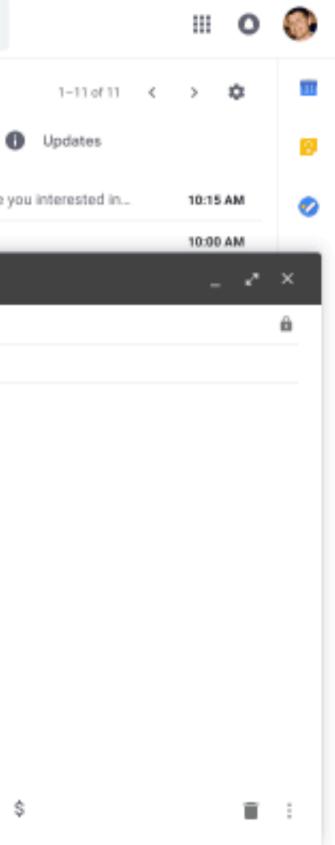
≡	M Gmail	Q, Search mail		Ŧ
ł	Compose	🗆 - C' : 🖵 Primary	Social Promotions 2 new	
	Inbox 3		Think with Google	
\star	Starred	📄 🔶 Salit Kulla	Trip to Cairngorms National Park - Planning for a trip in	n July. Are
0	Snoozed	📩 📩 Brianna, John 2	Surf Sunday?- Great. Let's meet at Jack's at 8am, then?	
	Important	📄 🚖 Luis, me, Anastasia 3	Best Japane Taco Tuesday	
≻	Sent			
	Work	🖄 Daniel Vickery		
Ŷ	More	📩 📩 Nick Kortendick	Work Pres Taco Tuesday	
		🗌 📩 Tim Greer	Work Bus	
		🔄 📩 Karen, Meredith, James 5	Hiking this v	
		🔄 🚖 Anissa, Meredith, James 3	Mike's surpr	
		🗌 📩 Song Chi	Cooking cla	
		📄 🚖 Cameron, Tyler, Dylan 6	Pictures fro	
			IMG_0	
		🗌 📩 Mizra Sato	My roadtrip	
		0.33 GB (2%) of 15 GB used Manage	Send A () 00 😂 🛆	. 0



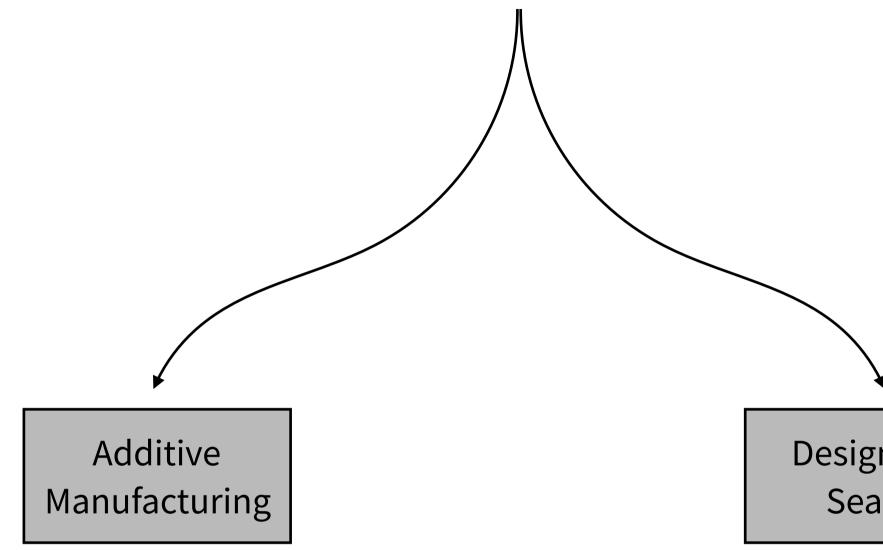
via Google

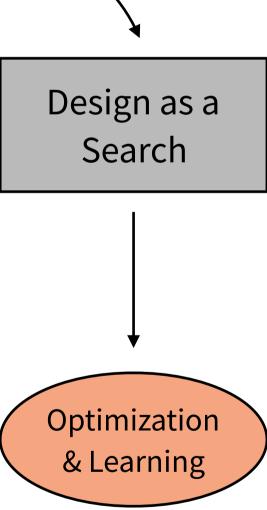
Gmail: Smart Compose

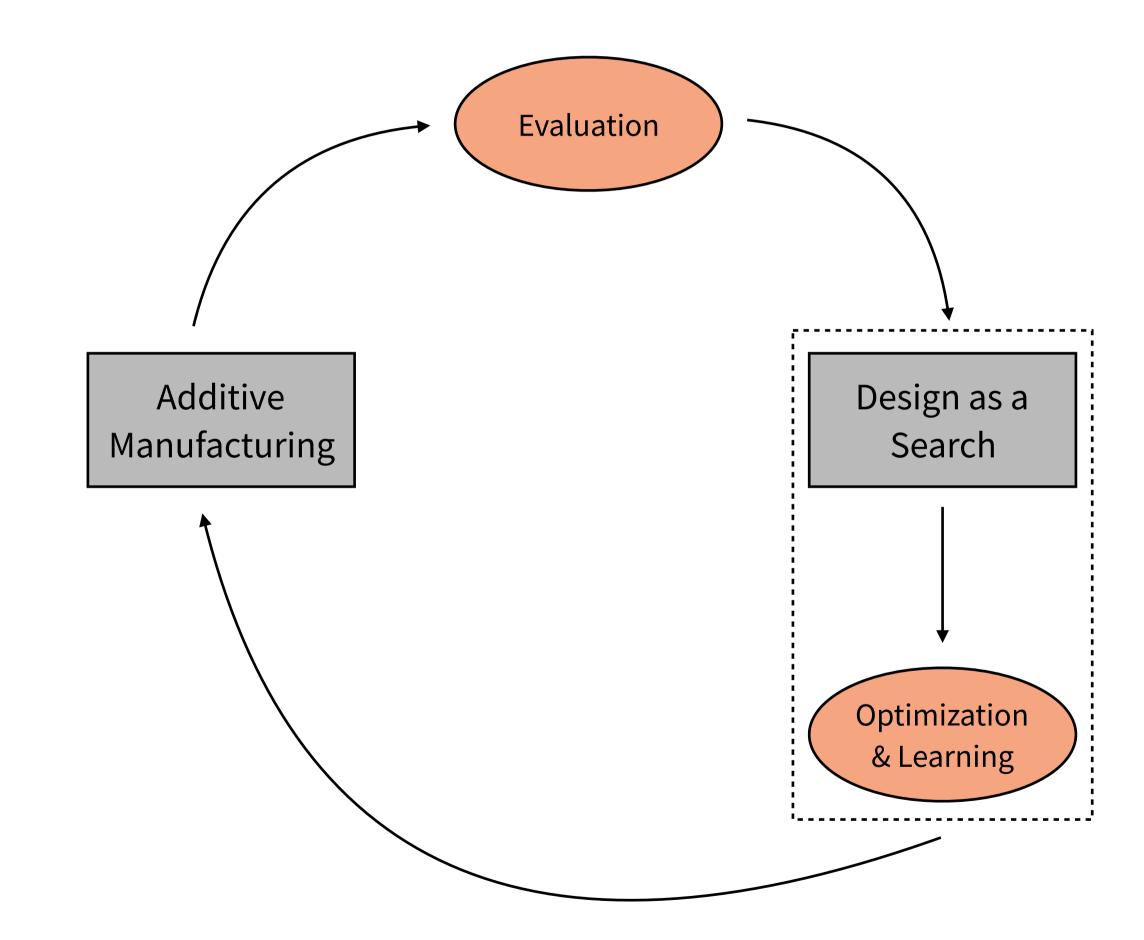
≡	M Gmail	Q, Search mail		Ŧ
ł	Compose	🗆 - C' : 🖵 Primary	Social Promotions 2 new	
	Inbox 3		Think with Google	
\star	Starred	📄 🔶 Salit Kulla	Trip to Cairngorms National Park - Planning for a trip in	n July. Are
0	Snoozed	📩 📩 Brianna, John 2	Surf Sunday?- Great. Let's meet at Jack's at 8am, then?	
	Important	📄 🚖 Luis, me, Anastasia 3	Best Japane Taco Tuesday	
≻	Sent			
	Work	🖄 Daniel Vickery		
Ŷ	More	📩 📩 Nick Kortendick	Work Pres Taco Tuesday	
		🗌 📩 Tim Greer	Work Bus	
		🔄 📩 Karen, Meredith, James 5	Hiking this v	
		🔄 🚖 Anissa, Meredith, James 3	Mike's surpr	
		🗌 📩 Song Chi	Cooking cla	
		📄 🚖 Cameron, Tyler, Dylan 6	Pictures fro	
			IMG_0	
		🗌 📩 Mizra Sato	My roadtrip	
		0.33 GB (2%) of 15 GB used Manage	Send A () 00 😂 🛆	. 0

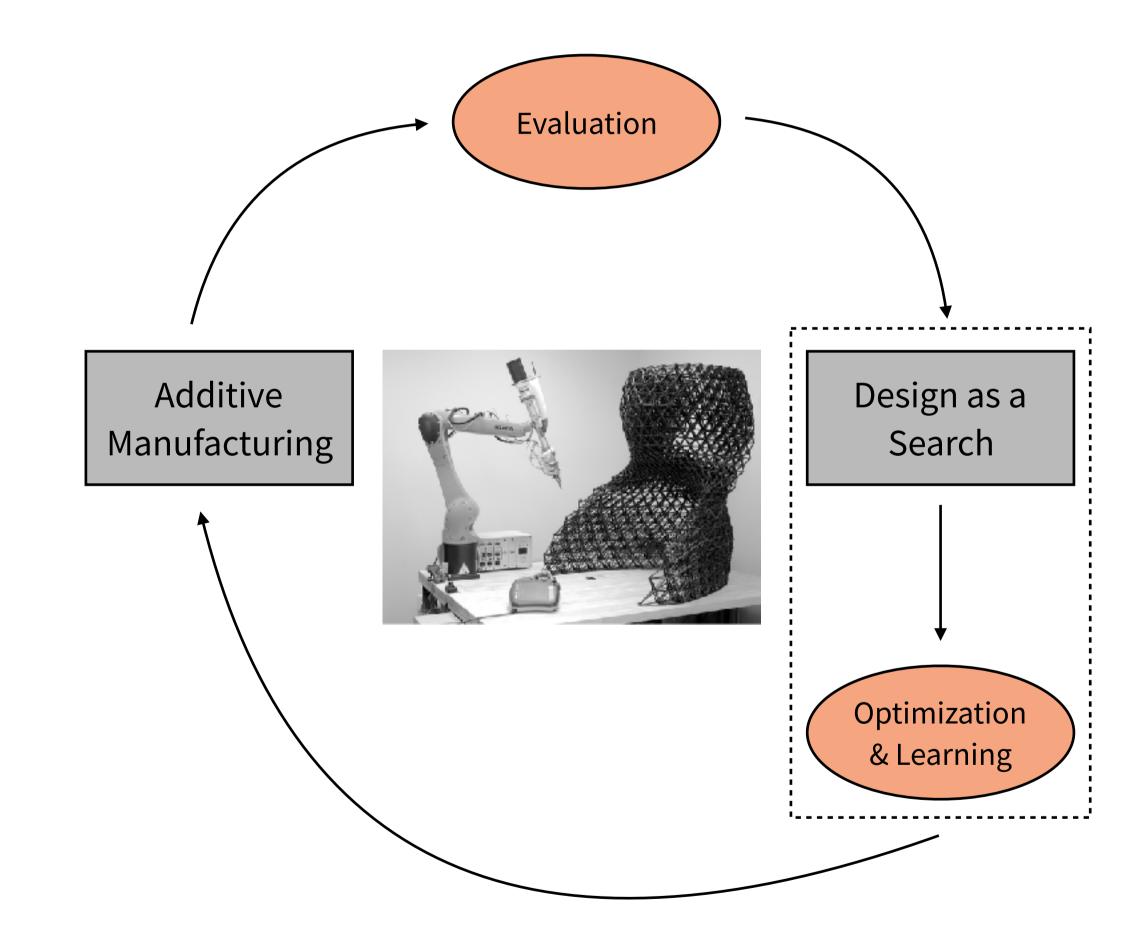


via Google









via nbww.com

