The Alan Turing Institute

Cloud-first data science at the Turing

Martin O’Reilly
Principal Research Software Engineer
The Alan Turing Institute

turing.ac.uk @martinoreilly | @turinginst

The Alan Turing Institute is the national centre for data science, headquartered at the British Library.

UNIVERSITY OF CAMBRIDGE
THE UNIVERSITY OF EDINBURGH
UCL
WARWICK
THE UNIVERSITY OF WARWICK
UNIVERSITY OF OXFORD
EPSRC
Pioneering research and skills

Research engineering group

A permanent team of research software engineers and data scientists who work with researchers to increase the impact of their work.

https://www.turing.ac.uk/research-engineering/

The Alan Turing Institute

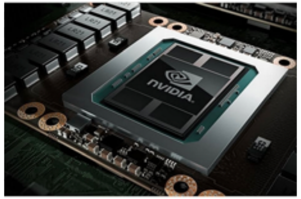
Challenges

Martin O'Reilly
Principal Research Software Engineer
The Alan Turing Institute

Challenge 1: How much compute can we afford?

Tier 1 national cluster

Tier 2 regional clusters



cirrus
Powered by epcc

Tier 3 institutional clusters

Turing Cray and Intel clusters

Partner university institutional clusters

Institutional cloud compute

Microsoft Azure

Challenge 2: The cloud learning curve

Model A: Workstation++

- Runs a handful of big VMs
- Gets more personal compute resource than can fit under a desk
- Manages VMs and workload manually

Model B: High throughput "cluster"

- Generally running large or urgent jobs
- Generally existing cluster users
- Need to deploy lots of workers

Challenges to self-sufficiency

- Unfamiliar compute / data abstractions
- Deploying cloud resources on demand
- Can't simply "lift and shift" existing code
- Usage / cost management

The Alan Turing Institute

Outcomes

Martin O'Reilly
Principal Research Software Engineer
The Alan Turing Institute

Outcome 1: How much cloud compute does $1,000,000 buy?

Azure core-years if spent on one compute type

CPU node type	On demand cores	Reserved cores (3yr reservation)	Low priority batch cores
Low memory HTC (2 GB / core)	2,667	7,280	11,356
Low memory HTC (4 GB / core)	2,351	6,003	11,284
Medium memory HTC (8 GB / core)	1,697	4,513	8,477
High memory HTC (14 GB / core)	745	1,908	3,723
Medium memory HPC (7 GB /core)	678	2,325	4,537
High memory HPC (14 GB / core)	908	1,735	3,388
GPU node type	**On demand GPUs**	**Reserved GPUs (3yr reservation)**	**Low priority batch GPUs**
Kepler K80 HTC	125	282	627
Kepler K80 HPC	114	257	570
Pascal P100 HTC	55	123	273
Pascal P100 HPC	50	112	248
Pascal P40 HTC	55	123	273
Pascal P40 HPC	50	112	248
Volta V100 HTC	37	-	369
Volta V100 HPC	334	-	335

Figures ignore storage costs and are based on Azure Linux VM pricing for East US region as at 11/04/2018 with 1 USD = 0.74 GBP. HTC = RDMA interconnect.

On-premise cluster comparisons

National (Tier 1):

- ARCHER: 118,080 cores with 0.75 or 1.5 GB / core*

Regional (Tier 2):

- CIRRUS: 10,080 cores with 4 GB / core
- JADE: 176 Pascal P100 GPUs
- Emerald: 372 Fermi M2090 GPUs [retired]

Institutional (Tier 3):

- UCL Grace: 10,944 cores with 5.3 GB / core
- QMUL Apocrita HTC: 2,280 cores with 2 or 10 GB / core**
- QMUL Apocrita HPC: 1,808 cores with 1.3 or 8 GB / core***

* 9,024 (7.6%) of ARCHER cores have 1.5 GB RAM / core

** 480 (21%) of Apocrita HTC cores have 10 GB RAM / core

*** 1,088 (60%) of Apocrita HPC cores have 8 GB RAM / core

Outcome 1: How much cloud compute does $1,000,000 buy?

Azure core-years if spent on one compute type

CPU node type	On demand cores	Reserved cores (3yr reservation)	Low priority batch cores
Low memory HTC (2 GB / core)	2,667	7,280	11,356
Low memory HTC (4 GB / core)	2,351	6,003	11,284
Medium memory HTC (8 GB / core)	1,697	4,513	8,477
High memory HTC (14 GB / core)	745	1,908	3,723
Medium memory HPC (7 GB /core)	678	2,325	4,537
High memory HPC (14 GB / core)	908	1,735	3,388
GPU node type	**On demand GPUs**	**Reserved GPUs (3yr reservation)**	**Low priority batch GPUs**
Kepler K80 HTC	125	282	627
Kepler K80 HPC	114	257	570
Pascal P100 HTC	55	123	273
Pascal P100 HPC	50	112	248
Pascal P40 HTC	55	123	273
Pascal P40 HPC	50	112	248
Volta V100 HTC	37	-	369
Volta V100 HPC	334	-	335

Figures ignore storage costs and are based on Azure Linux VM pricing for East US region as at 11/04/2018 with 1 USD = 0.74 GBP. HTC = RDMA interconnect.

On-premise cluster comparisons

National (Tier 1):

- ARCHER: 118,080 cores with 0.75 or 1.5 GB / core*

Regional (Tier 2):

- CIRRUS: 10,080 cores with 4 GB / core
- JADE: 176 Pascal P100 GPUs
- Emerald: 372 Fermi M2090 GPUs [retired]

Institutional (Tier 3):

- UCL Grace: 10,944 cores with 5.3 GB / core
- QMUL Apocrita HTC: 2,280 cores with 2 or 10 GB / core**
- QMUL Apocrita HPC: 1,808 cores with 1.3 or 8 GB / core***

* 9,024 (7.6%) of ARCHER cores have 1.5 GB RAM / core

** 480 (21%) of Apocrita HTC cores have 10 GB RAM / core

*** 1,088 (60%) of Apocrita HPC cores have 8 GB RAM / core

Outcome 1: How much cloud compute does $1,000,000 buy?

Azure core-years if spent on one compute type

CPU node type	On demand cores	Reserved cores (3yr reservation)	Low priority batch cores
Low memory HTC (2 GB / core)	2,667	7,280	11,356
Low memory HTC (4 GB / core)	2,351	6,003	11,284
Medium memory HTC (8 GB / core)	1,697	4,513	8,477
High memory HTC (14 GB / core)	745	1,908	3,723
Medium memory HPC (7 GB /core)	678	2,325	4,537
High memory HPC (14 GB / core)	908	1,735	3,388
GPU node type	**On demand GPUs**	**Reserved GPUs (3yr reservation)**	**Low priority batch GPUs**
Kepler K80 HTC	125	282	627
Kepler K80 HPC	114	257	570
Pascal P100 HTC	55	123	273
Pascal P100 HPC	50	112	248
Pascal P40 HTC	55	123	273
Pascal P40 HPC	50	112	248
Volta V100 HTC	37	-	369
Volta V100 HPC	334	-	335

Figures ignore storage costs and are based on Azure Linux VM pricing for East US region as at 11/04/2018 with 1 USD = 0.74 GBP. HTC = RDMA interconnect.

On-premise cluster comparisons

National (Tier 1):

- ARCHER: 118,080 cores with 0.75 or 1.5 GB / core*

Regional (Tier 2):

- CIRRUS: 10,080 cores with 4 GB / core
- JADE: 176 Pascal P100 GPUs
- Emerald: 372 Fermi M2090 GPUs [retired]

Institutional (Tier 3):

- UCL Grace: 10,944 cores with 5.3 GB / core
- QMUL Apocrita HTC: 2,280 cores with 2 or 10 GB / core**
- QMUL Apocrita HPC: 1,808 cores with 1.3 or 8 GB / core***

* 9,024 (7.6%) of ARCHER cores have 1.5 GB RAM / core

** 480 (21%) of Apocrita HTC cores have 10 GB RAM / core

*** 1,088 (60%) of Apocrita HPC cores have 8 GB RAM / core

Outcome 1: How much cloud compute does $1,000,000 buy?

Azure core-years if spent on one compute type

CPU node type	On demand cores	Reserved cores (3yr reservation)	Low priority batch cores
Low memory HTC (2 GB / core)	2,667	7,280	11,356
Low memory HTC (4 GB / core)	2,351	6,003	11,284
Medium memory HTC (8 GB / core)	1,697	4,513	8,477
High memory HTC (14 GB / core)	745	1,908	3,723
Medium memory HPC (7 GB /core)	678	2,325	4,537
High memory HPC (14 GB / core)	908	1,735	3,388
GPU node type	**On demand GPUs**	**Reserved GPUs (3yr reservation)**	**Low priority batch GPUs**
Kepler K80 HTC	125	282	627
Kepler K80 HPC	114	257	570
Pascal P100 HTC	55	123	273
Pascal P100 HPC	50	112	248
Pascal P40 HTC	55	123	273
Pascal P40 HPC	50	112	248
Volta V100 HTC	37	-	369
Volta V100 HPC	334	-	335

Figures ignore storage costs and are based on Azure Linux VM pricing for East US region as at 11/04/2018 with 1 USD = 0.74 GBP. HTC = RDMA interconnect.

On-premise cluster comparisons

National (Tier 1):

- ARCHER: 118,080 cores with 0.75 or 1.5 GB / core*

Regional (Tier 2):

- CIRRUS: 10,080 cores with 4 GB / core
- JADE: 176 Pascal P100 GPUs
- Emerald: 372 Fermi M2090 GPUs [retired]

Institutional (Tier 3):

- UCL Grace: 10,944 cores with 5.3 GB / core
- QMUL Apocrita HTC: 2,280 cores with 2 or 10 GB / core**
- QMUL Apocrita HPC: 1,808 cores with 1.3 or 8 GB / core***

* 9,024 (7.6%) of ARCHER cores have 1.5 GB RAM / core

** 480 (21%) of Apocrita HTC cores have 10 GB RAM / core

*** 1,088 (60%) of Apocrita HPC cores have 8 GB RAM / core

Outcome 1: How much cloud compute does $1,000,000 buy?

Azure core-years if spent on one compute type

CPU node type	On demand cores	Reserved cores (3yr reservation)	Low priority batch cores
Low memory HTC (2 GB / core)	2,667	7,280	11,356
Low memory HTC (4 GB / core)	2,351	6,003	11,284
Medium memory HTC (8 GB / core)	1,697	4,513	8,477
High memory HTC (14 GB / core)	745	1,908	3,723
Medium memory HPC (7 GB /core)	678	2,325	4,537
High memory HPC (14 GB / core)	908	1,735	3,388
GPU node type	**On demand GPUs**	**Reserved GPUs (3yr reservation)**	**Low priority batch GPUs**
Kepler K80 HTC	125	282	627
Kepler K80 HPC	114	257	570
Pascal P100 HTC	55	123	273
Pascal P100 HPC	50	112	248
Pascal P40 HTC	55	123	273
Pascal P40 HPC	50	112	248
Volta V100 HTC	37	-	369
Volta V100 HPC	334	-	335

Figures ignore storage costs and are based on Azure Linux VM pricing for East US region as at 11/04/2018 with 1 USD = 0.74 GBP. HTC = RDMA interconnect.

On-premise cluster comparisons

National (Tier 1):

- ARCHER: 118,080 cores with 0.75 or 1.5 GB / core*

Regional (Tier 2):

- CIRRUS: 10,080 cores with 4 GB / core
- JADE: 176 Pascal P100 GPUs
- Emerald: 372 Fermi M2090 GPUs [retired]

Institutional (Tier 3):

- UCL Grace: 10,944 cores with 5.3 GB / core
- QMUL Apocrita HTC: 2,280 cores with 2 or 10 GB / core**
- QMUL Apocrita HPC: 1,808 cores with 1.3 or 8 GB / core***

* 9,024 (7.6%) of ARCHER cores have 1.5 GB RAM / core

** 480 (21%) of Apocrita HTC cores have 10 GB RAM / core

*** 1,088 (60%) of Apocrita HPC cores have 8 GB RAM / core

Outcome 1: How much cloud compute does $1,000,000 buy?

Azure core-years if spent on one compute type

CPU node type	On demand cores	Reserved cores (3yr reservation)	Low priority batch cores
Low memory HTC (2 GB / core)	2,667	7,280	11,356
Low memory HTC (4 GB / core)	2,351	6,003	11,284
Medium memory HTC (8 GB / core)	1,697	4,513	8,477
High memory HTC (14 GB / core)	745	1,908	3,723
Medium memory HPC (7 GB /core)	678	2,325	4,537
High memory HPC (14 GB / core)	908	1,735	3,388

GPU node type	On demand GPUs	Reserved GPUs (3yr reservation)	Low priority batch GPUs
Kepler K80 HTC	125	282	627
Kepler K80 HPC	114	257	570
Pascal P100 HTC	55	123	273
Pascal P100 HPC	50	112	248
Pascal P40 HTC	55	123	273
Pascal P40 HPC	50	112	248
Volta V100 HTC	37	-	369
Volta V100 HPC	34	-	335

Figures ignore storage costs and are based on Azure Linux VM pricing for East US region as at 11/04/2018 with 1 USD = 0.74 GBP. HTC = RDMA interconnect.

On-premise cluster comparisons

National (Tier 1):

- ARCHER: 118,080 cores with 0.75 or 1.5 GB / core*

Regional (Tier 2):

- CIRRUS: 10,080 cores with 4 GB / core
- JADE: 176 Pascal P100 GPUs
- Emerald: 372 Fermi M2090 GPUs [retired]

Institutional (Tier 3):

- UCL Grace: 10,944 cores with 5.3 GB / core
- QMUL Apocrita HTC: 2,280 cores with 2 or 10 GB / core**
- QMUL Apocrita HPC: 1,808 cores with 1.3 or 8 GB / core***

* 9,024 (7.6%) of ARCHER cores have 1.5 GB RAM / core

** 480 (21%) of Apocrita HTC cores have 10 GB RAM / core

*** 1,088 (60%) of Apocrita HPC cores have 8 GB RAM / core

Outcome 1: How much cloud compute does $1,000,000 buy?

Azure core-years if spent on one compute type

CPU node type	On demand cores	Reserved cores (3yr reservation)	Low priority batch cores
Low memory HTC (2 GB / core)	2,667	7,280	11,356
Low memory HTC (4 GB / core)	2,351	6,003	11,284
Medium memory HTC (8 GB / core)	1,697	4,513	8,477
High memory HTC (14 GB / core)	745	1,908	3,723
Medium memory HPC (7 GB /core)	678	2,325	4,537
High memory HPC (14 GB / core)	908	1,735	3,388
GPU node type	**On demand GPUs**	**Reserved GPUs (3yr reservation)**	**Low priority batch GPUs**
Kepler K80 HTC	125	282	627
Kepler K80 HPC	114	257	570
Pascal P100 HTC	55	123	273
Pascal P100 HPC	50	112	248
Pascal P40 HTC	55	123	273
Pascal P40 HPC	50	112	248
Volta V100 HTC	37	-	369
Volta V100 HPC	34	-	335

Figures ignore storage costs and are based on Azure Linux VM pricing for East US region as at 11/04/2018 with 1 USD = 0.74 GBP. HTC = RDMA interconnect.

On-premise cluster comparisons

National (Tier 1):

- ARCHER: 118,080 cores with 0.75 or 1.5 GB / core*

Regional (Tier 2):

- CIRRUS: 10,080 cores with 4 GB / core
- JADE: 176 Pascal P100 GPUs
- Emerald: 372 Fermi M2090 GPUs [retired]

Institutional (Tier 3):

- UCL Grace: 10,944 cores with 5.3 GB / core
- QMUL Apocrita HTC: 2,280 cores with 2 or 10 GB / core**
- QMUL Apocrita HPC: 1,808 cores with 1.3 or 8 GB / core***

* 9,024 (7.6%) of ARCHER cores have 1.5 GB RAM / core

** 480 (21%) of Apocrita HTC cores have 10 GB RAM / core

*** 1,088 (60%) of Apocrita HPC cores have 8 GB RAM / core

Outcome 1: How much cloud compute does $1,000,000 buy?

Storage costs as data size grows

		Download £ / year (for N annual downloads of full dataset)			
Dataset size	**Storage £ / year**	**1**	**10**	**100**	**1000**
1 GB	£0.19	£0.00	£0.00	£0.00	£0.00
10 GB	£1.87	£0.03	£0.27	£2.71	£27.08
100 GB	£18.72	£0.51	£5.15	£51.46	£514.58
1 TB	£191.69	£5.52	£55.20	£551.96	£5,519.58
10 TB	£1,916.93	£55.44	£554.40	£5,543.96	£55,439.58
100 TB	£19,169.28	£493.23	£4,932.27	£49,322.67	£493,226.67
1 PB	£196,293.43	£3,489.45	£34,894.51	£348,945.07	£3,489,450.67
10 PB	£1,962,934.27	£33,373.87	£333,738.67	£3,337,386.67	£33,373,866.67
100 PB	£19,629,342.72	£332,218.03	£3,322,180.27	£33,221,802.67	£332,218,026.67

$1,000,000 = £737,000

Figures are based on Azure "Hot" tier Blob storage Locally Redundant Storage (LRS) pricing for East US region as at 11/04/2018. Conversion rate used is 1 USD = 0.74 GBP. Note: Download cost per GB for 150-500 TB is about 60% that for 5 GB -10 TB with the price for > 500 TB listed as "Contact us". For these calculations, downloads above 500 TB have used the 150-500 TB cost per GB.

Outcome 1: How much cloud compute does $1,000,000 buy?

Storage costs as data size grows

		Download £ / year (for N annual downloads of full dataset)			
Dataset size	**Storage £ / year**	**1**	**10**	**100**	**1000**
1 GB	£0.19	£0.00	£0.00	£0.00	£0.00
10 GB	£1.87	£0.03	£0.27	£2.71	£27.08
100 GB	£18.72	£0.51	£5.15	£51.46	£514.58
1 TB	£191.69	£5.52	£55.20	£551.96	£5,519.58
10 TB	£1,916.93	£55.44	£554.40	£5,543.96	£55,439.58
100 TB	£19,169.28	£493.23	£4,932.27	£49,322.67	£493,226.67
1 PB	£196,293.43	£3,489.45	£34,894.51	£348,945.07	£3,489,450.67
10 PB	£1,962,934.27	£33,373.87	£333,738.67	£3,337,386.67	£33,373,866.67
100 PB	£19,629,342.72	£332,218.03	£3,322,180.27	£33,221,802.67	£332,218,026.67

$1,000,000 = £737,000

Figures are based on Azure "Hot" tier Blob storage Locally Redundant Storage (LRS) pricing for East US region as at 11/04/2018. Conversion rate used is 1 USD = 0.74 GBP.
Note: Download cost per GB for 150-500 TB is about 60% that for 5 GB -10 TB with the price for > 500 TB listed as "Contact us". For these calculations, downloads above 500 TB have used the 150-500 TB cost per GB.

Outcome 1: How much cloud compute does $1,000,000 buy?

Storage costs as data size grows

		Download £ / year (for N annual downloads of full dataset)			
Dataset size	**Storage £ / year**	**1**	**10**	**100**	**1000**
1 GB	£0.19	£0.00	£0.00	£0.00	£0.00
10 GB	£1.87	£0.03	£0.27	£2.71	£27.08
100 GB	£18.72	£0.51	£5.15	£51.46	£514.58
1 TB	£191.69	£5.52	£55.20	£551.96	£5,519.58
10 TB	£1,916.93	£55.44	£554.40	£5,543.96	£55,439.58
100 TB	£19,169.28	£493.23	£4,932.27	£49,322.67	£493,226.67
1 PB	£196,293.43	£3,489.45	£34,894.51	£348,945.07	£3,489,450.67
10 PB	£1,962,934.27	£33,373.87	£333,738.67	£3,337,386.67	£33,373,866.67
100 PB	£19,629,342.72	£332,218.03	£3,322,180.27	£33,221,802.67	£332,218,026.67

$1,000,000 = £737,000

Figures are based on Azure "Hot" tier Blob storage Locally Redundant Storage (LRS) pricing for East US region as at 11/04/2018. Conversion rate used is 1 USD = 0.74 GBP. Note: Download cost per GB for 150-500 TB is about 60% that for 5 GB -10 TB with the price for > 500 TB listed as "Contact us". For these calculations, downloads above 500 TB have used the 150-500 TB cost per GB.

Outcome 1: How much cloud compute does $1,000,000 buy?

Storage costs as data size grows

		Download £ / year (for N annual downloads of full dataset)			
Dataset size	**Storage £ / year**	**1**	**10**	**100**	**1000**
1 GB	£0.19	£0.00	£0.00	£0.00	£0.00
10 GB	£1.87	£0.03	£0.27	£2.71	£27.08
100 GB	£18.72	£0.51	£5.15	£51.46	£514.58
1 TB	£191.69	£5.52	£55.20	£551.96	£5,519.58
10 TB	£1,916.93	£55.44	£554.40	£5,543.96	£55,439.58
100 TB	£19,169.28	£493.23	£4,932.27	£49,322.67	£493,226.67
1 PB	£196,293.43	£3,489.45	£34,894.51	£348,945.07	£3,489,450.67
10 PB	£1,962,934.27	£33,373.87	£333,738.67	£3,337,386.67	£33,373,866.67
100 PB	£19,629,342.72	£332,218.03	£3,322,180.27	£33,221,802.67	£332,218,026.67

$1,000,000 = £737,000

Figures are based on Azure "Hot" tier Blob storage Locally Redundant Storage (LRS) pricing for East US region as at 11/04/2018. Conversion rate used is 1 USD = 0.74 GBP. Note: Download cost per GB for 150-500 TB is about 60% that for 5 GB -10 TB with the price for > 500 TB listed as "Contact us". For these calculations, downloads above 500 TB have used the 150-500 TB cost per GB.

Outcome 2: The cloud learning curve

Current support

- Training sessions
- “Azure club” support sessions
- Support from research engineering
- Some orchestration scripts
- Azure tooling such as Batch Shipyard

Current adoption

- Better than last year but a way to go
- A few big “high throughput” users
- Most use “workstation++” model
- Most deploy via portal
- Most don’t use low priority
- Even cluster users struggle

Some Azure highlights

Web archive analysis

Data Study Groups

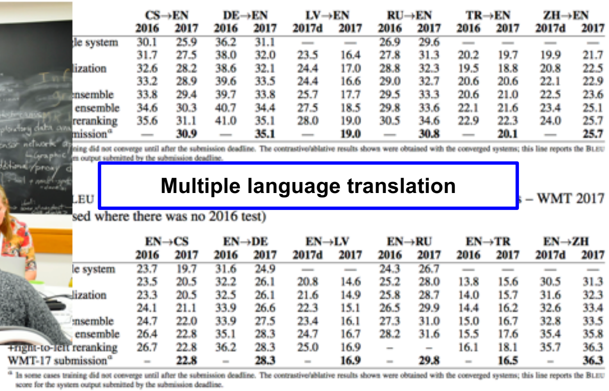


Table 2: BLEU scores for translating news *into* English (WMT 2016 and 2017 test sets – WMT 2017 dev set is used where there was no 2016 test)

	CS→EN 2016	CS→EN 2017	DE→EN 2016	DE→EN 2017	LV→EN 2017d	LV→EN 2017	RU→EN 2016	RU→EN 2017	TR→EN 2016	TR→EN 2017	ZH→EN 2017d	ZH→EN 2017
le system	30.1	25.9	36.2	31.1	—	—	26.9	29.6	—	—	—	—
	31.7	27.5	38.0	32.0	23.5	16.4	27.8	31.3	20.2	19.7	19.9	21.7
lization	32.6	28.2	38.6	32.1	24.4	17.0	28.8	32.3	19.5	18.8	20.8	22.5
	33.2	28.9	39.6	33.5	24.4	16.6	29.0	32.7	20.6	20.6	22.1	22.9
nsemble	33.8	29.4	39.7	33.8	25.7	17.7	29.5	33.3	20.6	21.0	22.5	23.6
ensemble	34.6	30.3	40.7	34.4	27.5	18.5	29.8	33.6	22.1	21.6	23.4	25.1
reranking	35.6	31.1	41.0	35.1	28.0	19.0	30.5	34.6	22.9	22.3	24.0	25.7
mission[a]	—	**30.9**	—	**35.1**	—	**19.0**	—	**30.8**	—	**20.1**	—	**25.7**

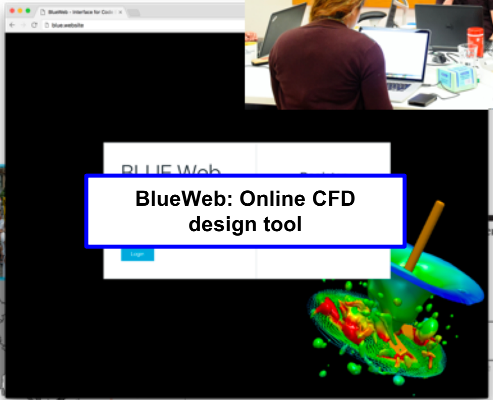
ning did not converge until after the submission deadline. The contrastive/ablative results shown were obtained with the converged systems; this line reports the BLEU m output submitted by the submission deadline.

LEU … – WMT 2017 sed where there was no 2016 test)

	EN→CS 2016	EN→CS 2017	EN→DE 2016	EN→DE 2017	EN→LV 2017d	EN→LV 2017	EN→RU 2016	EN→RU 2017	EN→TR 2016	EN→TR 2017	EN→ZH 2017d	EN→ZH 2017
le system	23.7	19.7	31.6	24.9	—	—	24.3	26.7	—	—	—	—
	23.5	20.5	32.2	26.1	20.8	14.6	25.2	28.0	13.8	15.6	30.5	31.3
lization	23.3	20.5	32.5	26.1	21.6	14.9	25.8	28.7	14.0	15.7	31.6	32.3
	24.1	21.1	33.9	26.6	22.3	15.1	26.5	29.9	14.4	16.2	32.6	33.4
nsemble	24.7	22.0	33.9	27.5	23.4	16.1	27.3	31.0	15.0	16.7	32.8	33.5
ensemble	26.4	22.8	35.1	28.3	24.7	16.7	28.2	31.6	15.5	17.6	35.4	35.8
+right-to-left reranking	26.7	22.8	36.2	28.3	25.0	16.9	–	–	16.1	18.1	35.7	36.3
WMT-17 submission[a]	–	**22.8**	–	**28.3**	–	**16.9**	–	**29.8**	–	**16.5**	–	**36.3**

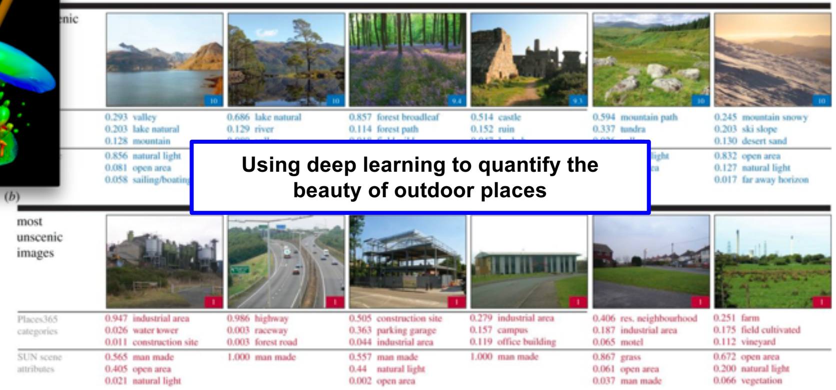
[a] In some cases training did not converge until after the submission deadline. The contrastive/ablative results shown were obtained with the converged systems; this line reports the BLEU score for the system output submitted by the submission deadline.

Multiple language translation



BlueWeb: Online CFD design tool

Quantifying Human Behaviour Using Online Photographs



Using deep learning to quantify the beauty of outdoor places

The Alan Turing Institute

Future

Martin O'Reilly
Principal Research Software Engineer
The Alan Turing Institute

Future

More compute for our $$$

- More use of low priority compute
- Compute directly on data stored outside of Azure
- Both require use of more sophisticated workflows, so need an easier learning curve
- More proactive monitoring of spending

Making the learning curve easier

- Make the Azure compute / data abstractions more transparent
- Provide easy to use workflows for a range of "standard" Turing use cases

turing.ac.uk
@turinginst

moreilly@turing.ac.uk
@martinoreilly