

It's The Algorithm, Stupid

The UK's coronavirus underperformance compared to Asia is one stark example of a bigger problem; a national operating system missing vital 2020s system updates. We need to re-programme UK education for the cross-disciplinary algorithm economy. Dr Alex Connock

In 1992, James Carville, then presidential campaigner Bill Clinton's political strategist, famously put up a sign in the Arkansas campaign HQ, which read:

- 1. Change vs more of the same*
- 2. The economy, stupid*
- 3. Don't forget health care*

All three dictums remain hugely applicable to politics today – in fact also to civic life, business, the military, education, what programmes to schedule on your TV channel, how to commission a popular news site, and how to manage a global pandemic and concomitant economic shutdown.

But thirty years on, in these uncertain times, those dictums need an upgrade.

With the arrival of the internet, hyper-targeted digital advertising, echo-chamber social distribution channels, the *YouTube* 'rabbit hole', *Netflix* and *Spotify* recommendation engines, predictive analytics in general, *Cambridge Analytica*

specifically, cyber warfare and deliberate misinformation, the second dictum could be expanded a little. Add in a point:

4. *The Algorithm, stupid*

This chapter is just a quick primer on the topic: but the subject is one that demands a fundamental re-mapping of our education system, our universities, our lifelong learning patterns, how we manage data, how we plan for future crises, and quite possibly the structure of our government and public affairs.

That doesn't mean the *Dirty Dozen* approach of Dominic Cummings' "mistfits and weirdos." It does mean that everyone is going to have to recalibrate their formulae for 21st century success, and learn a lot more maths. We're going to be on this topic for a while.

What actually is an algorithm ?

One of the harder words in the English language to spell - anyone for 'algorhythm' ? - an algorithm, in maths and computer science, is 'a precisely defined set of mathematical or logical operations for the performance of a particular task.' (*Oxford English Dictionary*).

Algorithms are used in data processing, automated reasoning and myriad other dimensions of almost every waking minute of your stylish, modern life: every time you are even looking at your iPhone, for instance. A useful Institute of Mathematics piece¹ describes how algorithms can be *proclaimed*, such as the tax code where for a given set of circumstances if your income is X, then you pay Y in tax. Or they can be *inferred*, such as the wind chill factor which was

¹ <https://ima.org.uk/6910/the-debate-about-algorithms/>

experimentally determined by a multinational team of scientists. Algorithms can be secret even when we know they are being used – like with insurance calculations, where we know our past scrapes in the Asda car park and rowdy street are being used against us, but not exactly how.

Algorithms produced in machine learning, where a computer can produce a formula based on a large lump of ‘training data’ may not be transparent or understood in the first place. *Advisory* algorithms give a human some advice, a recommendation. *Determinative* algorithms make a firm decision – like the Duckworth-Lewis in cricket, and (it seems) the famous A level results of August 2020 (see below). In that case they are the subject of GDPR regulation article 22.1 around ‘a decision based solely on automated processing’; these affect lives, by discriminating in personal finance decisions for instance.

Starting from an initial state and an initial input, the instructions for an algorithm work through a sequence of defined states, eventually producing an output. That will happen fast: an *Apple* A14 chip, found in the newest iPhone 12, can execute 11 trillion operations per second. Algorithms work you through your tax return, say, or at a most basic level, your road tax payment.

Algorithms were found in Babylonian and Greek maths, used by Arab mathematicians for code breaking, and titled after a Latinised version of the name of 9th century Muhammad ibn Musa al-Khwarizmi.

Forwarding fast to the modern era (and skipping a lot of relevant history) algorithms were developed in the 1920s to solve intractable, large-scale maths problems, then morphed into Alan Turing’s machines of the late 1930s and World War Two code-breaking. Turing then wrote a famous, foundational paper in the journal *Mind on Artificial Intelligence* in October 1950, using the phrase *The Imitation Game*. Even today, an algorithm can be considered to be

a sequence of operations that can be simulated by a Turing-complete system. If you want any more than that brief introduction, do please talk to someone who is an actual maths expert.

How are algorithms used ?

Algorithms are used everywhere in society; whether and how to build a new aircraft carrier, whether to give you a loan, whether to give you the rifle you want on level 3 of *SniperElite*, whether to give your parents cancer treatment, or whether *Waitrose* offers you a discount on low calorie crème fraîche. But as the Institute of Mathematics piece puts it: ‘nowhere is commerce using opaque machine learning algorithms more than in the internet itself.’

Algorithms drive almost everything you see on the internet unless you directly enter the address of the site you want to visit (which few people do) – and even then, at hundreds of telecom waypoints they will be critically involved in routing your computer there. Algorithms determine the results you are served by *Google* to be more relevant to your geography - so that when you’re looking for a Chinese takeaway in Cheshire you’re not offered websites for the Peking Garden, Acapulco. They drive the upsell you are offered (“make that a large fries?”) at an automated US McDonalds drive through, based on the weather, your number plate and previous orders, the time of day, other people’s real-time orders around the nation, what McDonalds promoted on the local radio station just now, and other factors. They work out the percentage likelihood of rain in your postcode next Thursday. Algorithms determine the choice of dresses distributed to each branch of some forward-looking high street fashion stores, based on usage patterns, and the films you are recommended on *Netflix*. (Interestingly, a clever National Film and Television School student of mine, Adrian Gordon, demonstrated that *Netflix* massively prioritises its own produced content in the feed, showing its own content over 50% of the time

even though that is barely 3% of the library. He did that by creating 48 distinct user profiles, and giving them all unique viewing tastes. He may have had a rather slow Summer.)

You find your next retro 1970s acoustic guitar/singer track on *Spotify* because of algorithms, and they are clever enough to deduce your mood on Tuesday afternoon from which particular James Taylor songs you are repeating.

Algorithms drive the movies you are offered on *Amazon Prime* - sadly I seem to be locked in a never-ending tunnel of non-English-language dystopian thrillers - and that's just a microcosm of *Amazon*'s preternatural knowledge of your disposition towards everything from Scotch tape price points to Scotch whisky preferences for work gifts (i.e. yourself).

Top of the news in 2020 is the algorithmic heft behind conspiracy theory *QAnon* and its wide-ranging, delusional, Instagrammed nonsense about child trafficking from pizza restaurants, coronavirus vaccine threat, 'the global kabal' which organised the epidemic as cover for a political takeover, and all the rest of it. *The New York Times*' brilliant *Rabbit Hole* series² (a must-listen, if you haven't) demonstrates the compulsive power of the *YouTube* recommendation engine in serving relentless conspiracy material to users, and driving commercial value (the pre-roll advertising, on a cost-per-thousand views model) from the repeat traffic. The switching power of social algorithms means a sucker for anti-vacc misinformation will likely be served material about 5G towers too, even if they have never expressed a direct interest, because the system knows how to recognise a vulnerable conspiracy theorist. Under scrutiny, *Facebook* and *Youtube* pivoted to more aggressive interventions against such content in late 2020, particularly around aggressive and targeted

² <https://www.nytimes.com/column/rabbit-hole>

election advertising – and they did that intervention, of course, with the algorithm.

So what went wrong in the A level algorithm ?

British society excels at computer science in all sorts of areas – but not always. In the UK our celebrity algorithm of 2020 was the exams “mutant algorithm” (as Boris Johnson described it) debacle, which got the A level results wrong. *Ofqual* never fully released its actual algorithm, the Direct Centre Performance Model, though it did publish a 319 page technical review. But the basic outline is out there on *Wikipedia* (NB: because of that, be vigilant this *just might* not be perfect information.) It was basically a three-stage process. (1) Teachers were asked to take account in their predictive grades factors including prior attainment of the students and mock exam results. (2) Schools were asked to rank their students within each grade. (3) The School’s predictions were then calibrated for the past performance of that school. Underpinning the third round of adjustment was that the government was keen to avoid coronavirus being an opportunity for national grade inflation.

So, it worked something like this:

For large schools with more than 15 students taking the A level

$$P_{kj} = (1-r_j) C_{kj} + r_j(C_{kj} + q_{kj} - p_{kj})$$

And for smaller schools with fewer than 15 students

$$P_{kj} = CAG$$

Where....

N is the number of pupils in the subject being assessed

k is a specific grade

j indicates the school

C_{kj} is the historical grade distribution of grade at the school (centre) over the last three years, 2017-19

q_{kj} is the predicted grade distribution based on the class's prior attainment at GCSEs.

p_{kj} is the predicted grade distribution of the previous years, based on their GCSEs.

r_j is the fraction of pupils in the class where historical data is available.

CAG is the centre assessed grade.

P_{kj} is the result, which is the grade distribution for each grade at each school

What went wrong was that the algorithm had in-built (though most likely accidental) prejudices. When the A levels were announced on 13 August, some 36% were below teachers' assessments (the centre-assessed grade). Students at small schools, or those taking minority subjects, got higher grades than teacher predictions. The algorithm did not have the right weightings for students at larger state schools and sixth form colleges, who saw their results downgraded -

sometimes so that they fitted past results, as if every classroom was really a facsimile of the kids who sat there the previous year, a pretty reductive view of human talents. UCAS said in mid-August that 15,000 students were rejected by their first-choice university because they did not get their required grades from the algorithm. *The Royal Statistical Society* said they had offered to help with the algorithm's construction, but not when they saw the NDA they had to sign.

It didn't get good press. Lord Falconer said the algorithm failed under the overarching objectives of *Ofqual's* 2009 establishment to give reliable grades to each student, and Labour as a whole argued that it was unlawful and discriminatory. The government U-turned on 17 August.

But there was something more than technical in play as well. Nick Maughan, founder of an education charity, wrote in *City AM*: "To use philosophical terminology, the algorithm had a fundamental category mistake built into it. It treated the efforts, abilities, and corresponding hopes and expectations of millions of young people as inputs that could not only be quantified but also qualified across an arbitrary set of figures and functions."

What happened with Coronavirus ?

Covid 19 isn't an algorithm, but combatting it has required algorithms in myriad ways, from mapping viral spread models to vaccine design. Britain had its victories. The National Health Service performed well, and the world-leading vaccine is from Oxford University, but the point remains: the UK performed poorly against international benchmarking. As Tom McTague in *The Atlantic* put it: "If almost all countries failed, then Britain failed more than most."

He describes a picture of “a country whose systemic weaknesses were exposed with appalling brutality,” and locates those weaknesses in two decades of national policy calamities including the 2011 Libyan intervention, the ‘near miss’ (watch this space) of Scottish independence in 2014, Brexit 2016 and more.

British coronavirus experience, with 46,000 deaths (and counting) had a higher death rate than almost any other country on the European continent, some 8 times higher than Germany. In the second wave of coronavirus, Britain’s advertised ‘world-class test and trace system’ was really not world class, and at certain points did not test or trace contacts in over half of new cases of coronavirus.

“The British government as a whole,” says McTague, “made poorer decisions, based on poorer advice, founded on poorer evidence, supplied by poorer testing, with the inevitable consequence that it achieved poorer results than almost any of its peers. It failed in its preparation, its diagnosis, and its treatment.”³

Why was that?

Finding the fault in the UK operating system

It all starts with how we organise education on faculty-specific grounds. Around the UK today, academics are off to work (in person, by *Zoom* or *Teams*) in universities with distinct faculties, buildings, canteens (or more likely an on-campus *Subway*, inexplicably beloved by students) for each different discipline

³ <https://www.theatlantic.com/international/archive/2020/08/why-britain-failed-coronavirus-pandemic/615166/>

critical to the UK's globally competitive future - say Maths or Media, Medicine, Business, Engineering or Computer Science.

Those academics are diligently teaching their distinct fields to hundreds of thousands of students, on focussed degree courses, publishing research in separate peer-reviewed journals at separate conferences, and vigorously hoarding the intellectual and organisational airspace around their subjects with the laudable aim of professional excellence.

Take an example: media matters to Britain as an economic driver. According to the Government's own website, the UK Creative industries contribute almost £13m to the UK economy *every hour*. But most media faculties today deliver course profiles segmented familiarly to anyone who read a David Lodge novel from the 1970s – Broadcast Journalism, Radio, Film Studies.

Business faculties matter too. In industry-facing, former polytechnic, 'skills-focussed' universities, many teach purist Organisational Behaviour, Accounting and Finance, not Artificial Intelligence and Analytics, as their core courses.

But outside the pageantry of degrees days, neither Business nor Media faculty will have much to do with Computer Science colleagues across campus. There will be few, if any, shared classes or research projects. '*Artificial intelligence*' will be a catch phrase that gets a mention over lunch – possibly confused with *Augmented Reality* - but which features rarely in comfortably retro courses on Documentary Filmmaking or Radio Journalism, that prepare students for a production and employment ecology which began vertiginous decline 30 years ago. In a classic instance of national industrial failure, the UK missed a chance to lead the streaming revolution by not backing its computer scientists and broadcasters to lead the world together. The BBC launched the world-beating,

video streaming *iPlayer* in 2007, the same year as *Netflix*, but was held back when the Competition Commission in 2009 blocked an idea (‘*Project Kangaroo*’) to globalise the *iPlayer*’s content to include *ITV* and *Channel 4*. *Netflix*, meanwhile, kicked off at the same time with a million-dollar prize - for a recommendation algorithm of course - and went on to build a business with an enterprise value, in October 2020, of \$488 billion. We could have done with another business like that on these shores right now.

And this is the point about algorithms.

As the UK’s academics leave work each day, alongside all the students they teach - enough people, in the UK alone, to fill all the stadia in the *Premiership* to capacity and more, if only social distancing would allow – they coalesce into a single, recognisable market segment; the digital consumer. They set off down the street listening to personalised song lists on *Spotify*, sit down on the couch to their own unique entry-point on *Netflix*, *Amazon Prime*, *Apple TV+* and *YouTube*, *Facebook*, *Instagram* and *TikTok*. Each one of their consumption choices is unique and algorithm-customised, on a platform which perfectly combines the cutting-edge attributes not of one university discipline, but three or four different faculties - Computer Science, Maths, Media and Business. Then the next morning they go back to their individual faculties again.

The scale of the competition

On a single day in 2020 (15 October), Indonesia, Turkey and Brazil all announced that they were putting *Artificial Intelligence* at the heart of their national strategies. America’s *FAANG* tech/media companies (*Facebook*, *Apple*, *Google/YouTube*, *Netflix* etc) meanwhile define the terms not only of modern consumption, but also the levers of content and creativity, political life, security, society, information, geopolitics and finance. Their scale is such that

they dwarf what we conventionally know of as ‘media’. *Apple* has throughout 2020 skirted a market capitalisation greater than the entire FTSE 100 index, exposing the collectively weak future-proofing in the City of London’s increasingly motley collection of oil and mining companies as principal home for the nation’s 21st century pensions. The *daily* variance in Amazon’s market capitalisation can be up to ten times greater than the entire value of the UK’s premier commercial broadcaster, the once world-leading ITV, or the total annual available programming budget of the very pioneer of broadcasting, the BBC. The *additional* market capitalisation awarded to *WalMart* when there was even a hint of it joining a process to buy Chinese social media youth platform *TikTok* in Summer 2020, was around 10 times greater than the *entire* annual programme spend of the UK’s own similarly ‘youth-focussed’ *Channel 4*.

TikTok is famous for one thing above all: its’ algorithm. It has a strong ability to serve sequentially compelling videos to users based on its sophisticated, black-box, China-located user profiling technology. Even the on-pause negotiations to repatriate US *Tik Tok*_user data to the US doesn’t touch the location and composition of the recommendation algorithm, the true value. And the relative valuation demonstrates a massive disconnect. The versions of the world that we are not only teaching, but which many of our universities are architecturally *structured* to represent by the way their buildings are organised, and the media that we are consuming, are being played out in parallel universes.

The universe that really matters is the one that has the algorithm in it. Computer science matters today; not instinct, not who you know or who you went to university with, not some punchy Latin quotes from your Classics degree, not even necessarily your elite civil service training. The algorithm is the *Rosetta Stone* to understanding not just the media in the 2020s, but also the

business and politics of the future, the health system, drones, cyberwarfare, real warfare, climate, farming, design, retail, marketing and almost any subject you could conceive. And yet to many of our students – all except those who study science and maths - we are barely even teaching the algorithm at all.

Can we fix it ?

The answer is not just to develop, own and exploit the best algorithms we can in any given discipline, which we are good at. It goes much broader.

We need a new flexibility in the way we even think about universities and particularly courses within them. A segmented faculty structure which might have been right for the mid to late twentieth century is much less relevant to the mid 21st century. We need to cross faculties on a daily basis – Media and Computer Science, Medicine and Engineering, Maths and English. We might need to add coding to every degree we teach. Workers who don't know about the predictive analytics that lie behind the most successful distribution, retail and content serving operations will not be able to take on the competition in subjects from home delivery supermarkets to audio production to healthcare.

So, for a single mantra to redefine how we re-frame our education this decade, it's hard to do a lot better than this: **it really is the algorithm.**

About the author.

Dr Alex Connock is Fellow in Management Practice (Marketing) at Saïd Business School, University of Oxford and co-director of its 2021 Diploma in Artificial Intelligence for Business.