Journal of Law, Technology and Trust, Vol 1, Issue 1, 2020

Between the Devil and the Deep Blue Sea (of data): navigating the temptations of the post-COVID hybrid campus

Andrew Cormack

Andrew Cormack is Chief Regulatory Adviser, Jisc, UK.
For correspondence: <Andrew.Cormack@jisc.ac.uk>

Abstract

For several years, universities and colleges have been exploring the potential use of activity data – already gathered by their digital systems – to improve their processes. Learning processes were the first to adopt these techniques, with a wide range of "learning analytics" services already in use. Similar approaches to curriculum development and effective use of campus spaces are also being explored. The COVID-19 pandemic has raised the question of whether infection prevention and support processes might also be data-enhanced. The adoption of a hybrid teaching mode by many institutions – where students are present on campus, but some face-to-face activities move online – is likely to increase the amount of data available. However this comes at a time of high stress, for both students and staff, when many of the concerns that had already been raised about "analytics" – over-simplification, lack of contextual awareness, dataveillance and a possible reaction against technology – are particularly salient. Emergency laws may provide less guidance. This paper suggests four questions – Will it Help? Will it Work? Will it Comfort? and Will it Fly? – as a framework for discussing data-enhanced processes among campus communities. Five "concept cars" are used to illustrate how these questions can be used to explore ideas and identify those likely to be widely accepted. If a proposed use of data cannot achieve consensus and trust there is unlikely to be willing compliance with its data collection or advice.

Keywords: Analytics, COVID-19, Ethics, Smart cities

The Post-COVID campus: data, stress and opportunity

Universities and colleges already have a lot of information about how their staff and students use digital systems. Authentication systems log who is active; networks log where they are sending traffic to and receiving it from; platforms such as Virtual Learning Environments (VLEs) record what they do. Traditionally that information was used to debug problems, to investigate misuse, perhaps for capacity planning.

More recently, it has been realised that the same data can also be used to understand how students learn, what curriculum styles are effective, even how to make more effective use of physical and digital spaces. Data from VLEs may show, much faster than the traditional assessment cycle, if students are falling behind, having problems with particular concepts, or could benefit from more challenging material. Complementing this field of Learning Analytics is Curriculum Analytics, which uses data from several courses to manage workloads, compare styles and outcomes, or determine why particular courses or modules may be more or less popular.[1] More recently, while the Intelligent Campus concept often involves deploying new sensors, interesting work has also been done with existing data, for example for capacity planning and identifying opportunities to use, and equip, areas outside lecture theatres as immediate, informal learning spaces.[2,3]

Several authors have highlighted the need for care in this "datafication".[4] Sclater identifies eighty-six legal and ethical concerns in Learning Analytics literature.[5] Selwyn points out that reducing educational processes to data or statistics is likely to "grossly underestimate the social complexity of classrooms, schools, and the complicated lives that students lead".[6] Ferguson responds that researchers, developers, students and tutors must all have a "well-developed understanding" of the processes within which data are collected and responses proposed.[7] Prinsloo and Slade warn that excessive focus on metrics, or merely the process of measurement, could increase stress on students (and staff, who are also in an "asymmetrical … power relationship" with their

---

[1] Niall Sclater, *Learning Analytics Explained* (London: Routledge, 2017), ch 3-7.

[2] Jisc, "Guide to the Intelligent Campus: Using data to make smarter use of your university or college estate" (n.d), available at https://repository.jisc.ac.uk/6882/1/Intelligent_Campus_Guide.pdf (accessed 17 August 2020).

[3] Luke McCrone, "Learning in the Spaces In Between", (WONKHE, 22 August 2019) available at https://www.northumbriajournals.co.uk/index.php/jltt/announcement/view/17 (accessed 19 August 2020).

[4] Ben Williamson, "Ten Definitions of Datafication (in education)" (2018), available at https://codeactsineducation.wordpress.com/2018/03/17/10-definitions-datafication (accessed 17 August 2020).

[5] Niall Sclater, "Developing a Code of Practice for Learning Analytics" (2016) 3(1) *Journal of Learning Analytics* 16–42, p. 22.

[6] Neil Selwyn, "What's the Problem with Learning Analytics?" (2019) 6(3) *Journal of Learning Analytics*, 11–19, p.12.

[7] Rebecca Ferguson, "Ethical Challenges for Learning Analytics" (2019) 6(3) *Journal of Learning Analytics*, 25–30, p. 26.

employers).[8] How to ensure students have sufficient data literacy to use dashboards as a motivator rather than a self-fulfilling prophecy is an area of active research.[9,10] Edwards warns that observation and measurement can even be counter-productive: if students or staff feel they are under surveillance, or "dataveillance",[11] they may avoid the very spaces and services that institutions are trying to use more effectively.[12] Existing work on smart cities highlights the particular sensitivity of data gathered from the spaces where we live and work.[13]

Whereas the use of data to support learning, teaching and campus design has been developing gradually for several years, the challenge of reopening campuses during a global pandemic may require institutions to explore and develop new uses of data in a matter of weeks or months. Might existing data sources help to support or – as institutions move beyond emergency mode – replace physical, procedural or informational safety measures?

Changes made in response to the virus may already have had the side-effect of institutions collecting more data about their students and staff. Most universities have adopted a 'hybrid' model of teaching where students are expected to be present on campus but activities that are incompatible with social distancing, such as large lectures, are conducted remotely with students participating using video-conference technology from their rooms.[14] Restrictions on off-campus facilities may also result in students spending more time in their rooms or, provided social distancing can be achieved, libraries. Considering the two main sources of incidental data – networks/authentication and learning platforms – these changes increase both the amount of time that students will be spending within their data-gathering scope and the overlap between them. In a normal campus environment, many students make significant use of a variety of non-campus networks, whether home or café WiFi, mobile devices on public transport journeys or, through the eduroam™ service, at other institutions.[15] Jisc's 2019/20 Student Experience Survey found 80% of HE students and 63% of FE knew that institutional systems could be accessed from off-

---

[8] Paul Prinsloo and Sharon Slade, "Student Vulnerability, Agency and Learning Analytics: An Exploration" (2016) 3(1) *Journal of Learning Analytics*, 159–182, p. 160.
[9] Jennifer Pei-Ling Tan et al, "Learner Dashboards a Double-Edged Sword? Students' Sense-Making of a Collaborative Critical Reading and Learning Analytics Environment for Fostering 21st Century Literacies" (2017) 4(1) *Journal of Learning Analytics*, 117–140, pp. 132-134.
[10] Liz Bennett, "Students' learning responses to receiving dashboard data: Research Report Jan 2018" (2018), available at https://www.srhe.ac.uk/downloads/reports-2016/LizBennet-scoping2016.pdf (accessed 17 August 2020).
[11] Selwyn, "What's the Problem with Learning Analytics?" *supra* n. 6.
[12] Lilian Edwards, "Privacy, Security and Data Protection in Smart Cities" (2016) 2(1) *European Data Protection Law Review*, 28-58, p. 58.
[13] Andrew Cormack, "See No…, Hear No…, Track No…: Ethics and the Intelligent Campus" (2019) 3(1) *Journal of Information Rights, Policy and Practice*, 3(1), pp. 1-3.
[14] PA Media, "Most UK Universities Planning In-person Teaching from Autumn, Poll Finds" (Guardian, 17 June 2020), available at https://www.theguardian.com/education/2020/jun/17/most-uk-universities-planning-in-person-teaching-from-autumn-poll-finds (accessed 17 Aug 2020).
[15] GEANT, "About eduroam" (n.d), available at https://www.eduroam.org/about/ (accessed 17 August 2020).

campus: it seems likely that most of these were doing so.[16,17] Health restrictions applied by off-campus locations, as well as personal protection choices, seem likely to change those habits and result in students spending more of the day connected to their own institution's network. The move from face-to-face to on-line delivery, of lectures in particular, means that activities that were previously invisible to learning platforms and their logs will now be visible there. Furthermore both trends will increase the periods for which institutions have simultaneous data from both sources: studying on local networks rather than off-campus makes network data available alongside the existing learning platform records; using digital platforms for on-campus learning delivery adds platform data where the institution may only previously have had information from the network.

Claims that data and technology can reduce virus spread have been a prominent feature of the pandemic. We are all now familiar with terms such as "transmission rates", "hotspots", "contact tracing", "customer registration", "apps" and "bubbles". Universities and colleges might well be considering whether their data might be used in these, or other, ways. Data might have roles in informing individuals or institutions about the local situation, in supporting those who need to take additional precautions or to self-isolate, in planning and management from social distancing to provision of resources or, perhaps, in enforcement. Ideas might be borrowed from a very wide range of other contexts, given that campuses are likely to play multiple roles: as semi-public indoor spaces, places where long-distance travellers meet and socialise, workplaces, homes, even primary care facilities. It's plausible that appropriate use of data might help to make other protective measures more effective, or allow them to be relaxed while keeping risk at an acceptable level.

However, at the same time as the pandemic offers more data, and more potential uses of it, it also leads to much more stress. It was recognised many years ago that datafication might create "learner and educator unease" even in normal teaching and learning environments;[18] these risks will inevitably be higher in a situation where the base level of stress is already raised. Students who have had their education disrupted may well be concerned about their futures if education continues to be different from what they (and their future employers) expected;[19] staff may have demographically legitimate concerns whether returning to anything near normal may place them at

---

[16] Jisc, "Student digital experience insights survey 2020. UK Higher Education survey findings" (September 2020) available at https://www.jisc.ac.uk/sites/default/files/student-dei-he-report-2020.pdf (accessed 21 September 2020).
[17] Jisc, "Learner digital experience insights survey 2020. UK Further Education survey findings" (September 2020) available at https://www.jisc.ac.uk/sites/default/files/student-dei-fe-report-2020.pdf (accessed 21 September 2020).
[18] George Siemens, "Learning analytics: envisioning a research discipline and a domain of practice" (Association for Computing Machinery, 2012) *Proceedings of the 2nd International Conference on Learning Analytics and Knowledge (LAK '12)* 4-8, p.6.
[19] National Union of Students, "Coronavirus and Students Survey" (April 2020) available at https://www.nusconnect.org.uk/resources/covid-19-and-students-survey-report (accessed 18 August 2020).

serious health risk;[20] all are likely to be suffering from “techno-stress” from several months of digital replacing social contact,[21] and concerns for families and friends. Even before the virus outbreak, wellbeing of staff and students was a concern.[22,23,24] In these new circumstances campuses must be even more supportive and non-threatening if the reaction to them is not to make wellbeing, and possibly even virus transmission, worse.[25]

If staff and students are not comfortable with how information about them is being used, or do not trust that the data and processing are well protected, they may well respond by trying to make data gathering harder. Cormack describes how responses such as turning mobile devices off (or leaving them at home), swapping smartcards or using off-campus spaces could undermine attempts to provide effective equipment and spaces for learning, or reduce campus efficiency and security.[26] In a pandemic, those responses will prevent warnings being received, corrupt the location information needed to identify infection hotspots, and encourage transmission between the campus and its surroundings. Any measures we use to reduce infection must be trusted, or they will have the opposite effect.

In normal circumstances, compliance with legal safeguards would be the essential foundation for establishing trust in data processing. Edwards suggests that data collection and use in smart cities will often use the legitimate interests justification – now GDPR[27] Article 6(1)(f) – if informed consent is not appropriate.[28] This imposes an explicit requirement to consider all the rights and freedoms (not just privacy) of individuals and to ensure that these do not override any benefit delivered by the processing.[29] When working to reduce the spread of a pandemic, however, institutions

---

[20] Nicholas G Davies et al, “Age-dependent effects in the transmission and control of COVID-19 epidemics” (2020) 26 *Nat Med* 1205–1211. doi:10.1038/s41591-020-0962-9.
[21] Monica Molino et al, “Wellbeing Costs of Technology Use during Covid-19 Remote Working: An Investigation Using the Italian Translation of the Technostress Creators Scale” (2020) 12 *Sustainability* 5911. doi:10.3390/su12155911, p. 3.
[22] Susan Guthrie et al, “Understanding Mental Health in the Research Environment” (Royal Society, June 2017) available at https://royalsociety.org/~/media/policy/topics/diversity-in-science/understanding-mental-health-in-the-research-environment.pdf (accessed 17 August 2020).
[23] Craig Thorley, “Not by Degrees: Improving Student Mental Health in the UK’s Universities” (IPPR, September 2017) available at https://www.ippr.org/research/publications/not-by-degrees (accessed 17 August 2020).
[24] UniversitiesUK, “Stepchange: Mentally Healthy Universities” (2020) available at https://www.universitiesuk.ac.uk/policy-and-analysis/reports/Documents/2020/uuk-stepchange-mhu.pdf (accessed 17 August 2020).
[25] Gurbaaz Gill, “What Students Need to Know is that Someone is Looking Out for Them” (WONKHE, 30 July 2020) available at https://wonkhe.com/blogs/what-students-need-is-to-know-that-someone-is-looking-out-for-them/ (accessed 17 August 2020).
[26] Cormack, “See No…, Hear No…, Track No…”, *supra* n 13, pp. 8,10.
[27] Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation).
[28] Edwards, “Smart Cities”, *supra* n 12, p. 43.
[29] Cormack, A.N. (2016). Downstream Consent: A Better Legal Framework for Big Data. *Journal of Information Rights, Policy and Practice, 1*(1), p.None. DOI: http://doi.org/10.21039/irpandp.v1i1.9

are much more likely to use Article 6(1)(e) public interest – Article 9(i) specifically identifies a public interest in "protecting against serious cross-border threats to health" – or even Article 6(1)(c) legal obligation, neither of which requires the data controller to consider individual rights or to obtain the individual's permission. Although coronavirus regulations are likely to be relied upon as the underlying "member state law", those regulations generally lack the safeguards that Article 6(3) expects such an enabling law to provide. In a situation where there is more data, more incentive to use it, more need for trust and fewer legal safeguards, the processes by which institutions choose which applications they will implement and which data they will use bear a heavy burden.

This paper suggests a four-step approach to discuss and assess ideas for data reuse, to increase the likelihood of choosing those that will be trustworthy, trusted and make a positive contribution, both in the circumstances of a pandemic, and beyond. First we introduce five "concept applications" that will be used to illustrate how this approach can be used to explore risks, benefits and dependencies.

"Concept Cars" for Thinking About Data Reuse

In the same way as the motor industry uses concept cars to explore possible future developments for technical feasibility and social acceptability,[30] this paper uses five concept applications. All are based on translations of existing ideas: either those from the pre-COVID world that might be used for new purposes on a post-pandemic campus; or taking ideas that have been mooted for such a campus and speculating how data might be used to support them. All use the same data source: location data from wireless access points (such information is subject to the GDPR and Information Commissioner's guidance on WiFi Analytics[31] but, since educational institutions do not generally offer publicly-available telecommunications services, not the ePrivacy Directive[32]). They have been chosen to represent four of the five levels in Cormack's Intelligent Campus intrusiveness model: counting, identifying, recording and processing.[33] None is roadworthy in its current form.

Capacity Warnings (counting)

Data from access points is already used to give an estimate of space utilisation.[34] Inverting this idea: by counting the number of wireless transmitters – mobile phones, tablets, laptops, etc. – in a lecture theatre and applying an appropriate devices per

[30] Wikipedia, "Concept Car" (n.d) available at https://en.wikipedia.org/wiki/Concept_car (accessed 17 August 2020).

[31] Information Commissioner, "Wi-Fi location Analytics" (16 February 2016) available at https://ico.org.uk/media/for-organisations/documents/1560691/wi-fi-location-analytics-guidance.pdf (accessed 21 September 2020).

[32] Directive 2002/58/EC of the European Parliament and of the Council of 12 July 2002 concerning the processing of personal data and the protection of privacy in the electronic communications sector (Directive on privacy and electronic communications), Article 2(c).

[33] Cormack, "See No…, Hear No…, Track No…", *supra* n 13, pp. 5-6.

[34] Lone Rooftop, "Understand Behaviour to Maximise the Economic Potential of the Estate" (n.d.) available at https://www.lonerooftop.com/solutions/highereducation (accessed 17 August 2020).

person factor, it might be possible to detect when the space is approaching safe capacity and use digital displays outside to direct latecomers to alternative locations.

Bottleneck Detection (counting)
Travellers through airports were familiar with queue-length monitoring systems, indeed these were clearly prominent in the minds of the drafters of the ePrivacy Regulation.[35] These use the speed of movement of wireless transmitting devices to determine where queues are forming (for example at security checks) and how long it is likely to take to pass through them. On a post-virus campus, any enforced dwell time (for example when leaving a lecture theatre) is likely to indicate a place where social distancing may break down. Bottleneck detection might be used to change signed routes around campus or even, if that is not possible, to alter timetables to reduce the number of people scheduled to pass through them in a short time interval.[36]

Location-specific Information (identifying)
Many mobile device apps use location as a trigger to display information relevant to where the user currently is.[37] Push notifications might be particularly helpful, for example to indicate where hand sanitiser can be found when someone enters a building, to provide directions around one-way systems or to discourage people from taking short-cuts through crowded areas.

Contact Recording (recording)
There has been a great deal of interest in technical systems to support contact tracing in case of virus outbreaks.[38] Since data from access points can record the location of devices, these records might be a way to work out, when an infection is reported, who else might have had significant exposure to it.

Bubble-Bursting (analysing)
Some institutions have discussed establishing "study bubbles", where students in a tutorial or lecture group would also be allocated shared accommodation.[39] Provided they stay together in these bubbles, students might be able to live more normally, with

---

[35] Proposal for a Regulation of the European Parliament and of the Council concerning the respect for private life and the protection of personal data in electronic communications and repealing Directive 2002/58/EC (Regulation on Privacy and Electronic Communications) (European Commission, 10 January 2017) COM(2017) 10 final 2017/003 (COD), Recital 25.
[36] E.g. Fraser Whieldon, "College drops plans for 'alternate week' rota after parent complaints" (FE Week, 14 September 2020) available at https://feweek.co.uk/2020/09/14/college-drops-plans-for-alternate-week-rota-after-parent-complaints/ (accessed 21 September 2020).
[37] Techopedia, "Location-Aware Application" (Techopedia, 2 December 2014) available at https://www.techopedia.com/definition/30836/location-aware-application (accessed 17 August 2020).
[38] Sarah Boseley, "Coronavirus: contact tracing explained" (Guardian, 17 April 2020) available at https://www.theguardian.com/world/2020/apr/17/coronavirus-contact-tracing-explained (accessed 11 August 2020).
[39] Rachel Hall, "UK universities plan to create 'social bubbles' when campuses reopen" (Guardian, 3 June 2020) available at https://www.theguardian.com/education/2020/jun/03/uk-universities-create-social-bubbles-campus-reopen-students-coronavirus (accessed 11 August 2020).

fewer restrictions needed to contain transmission of any infection. Wireless data might be used in real-time to detect "foreign" devices entering a space currently reserved for a bubble and to raise an alert.

Exploring Ideas for Data-enhanced Processes

Cardullo and Kitchin's studies of smart cities projects found that the most successful uses of data in shared living spaces were those where occupants of the space had meaningful involvement from the start of the project as Leaders, Decision-makers and Co-creators.[40] The epitome of this approach is Barcelona where the council identifies a problem, and citizens help to choose among a number of ways to address it.[41,42] The success in Barcelona contrasts strongly with the data/technology-led waterfront development in Toronto, which has recently failed after three years of planning.[43] Both experiences seem especially relevant where there is an agreement on the challenge – reopening campuses after a pandemic – but where different groups of occupants have significantly different assessments of the risks and benefits. Technology-led smart city projects result in, at best, indifference or, at worst, active resistance to data gathering. Particularly where compliance with any guidance, even participation in data collection, is likely to be voluntary, consensus and trust are essential.

Identifying uses of data that can support that consensus will require inter-disciplinary discussions, among – at least – staff, students, facilities managers, technologists and data scientists.[44] StudentMinds refer to this as "Co-production".[45] Rather than offering a single proposal – which may result in participants feeling they have only a binary choice – it may be better to offer a range from which they can select, prioritise, or even develop new ideas. The "concept applications" from the previous section might be used to fill in any gaps in that range, or at least as illustrations of the level of description likely to be helpful for this stage of discussion. Once one or two ideas have been selected as worth pursuing, more detailed investigation – including of processes, technologies, ethical and legal issues – can take place.

[40] Paolo Cardullo and Rob Kitchin, "Being a 'citizen' in the smart city: up and down the scaffold of smart citizen participation in Dublin, Ireland" *GeoJournal*, 12 Jan 2018 doi:10.1007/s10708-018-9845-8.

[41] Ross Tieman, "Barcelona: smart city revolution in progress" (Financial Times, 26 October 2017) available at https://www.ft.com/content/6d2fe2a8-722c-11e7-93ff-99f383b09ff9 (accessed 11 August 2020).

[42] Urban Hub, "Smart City 3.0 – Ask Barcelona about the next generation of smart cities" (13 February 2018) available at https://www.urban-hub.com/cities/smart-city-3-0-ask-barcelona-about-the-next-generation-of-smart-cities/ (accessed 17 August 2020).

[43] Gerrit De Vynck and Natalie Wong, "Alphabet's Dream of a Smart City in Toronto is Over" (Bloomberg, 7 May 2020) available at https://www.bloomberg.com/news/articles/2020-05-07/alphabet-s-dream-of-a-smart-city-in-toronto-is-over (accessed 11th Aug 2020).

[44] Jackie Labbe, "Planning for September Means Listening to the Experts" (WONKHE, 20 July 2020) available at https://wonkhe.com/blogs/planning-for-september-means-listening-to-the-experts/ (accessed 17 August 2020).

[45] StudentMinds, "Planning for a Sustainable Future: the Importance of University Mental Health in Uncertain Times" (n.d.) available at https://www.studentminds.org.uk/uploads/3/7/8/4/3784584/2005_planning_for_a_sustainable_future_a5.pdf (accessed 17 August 2020), p. 12.

To frame the initial discussions, and achieve a common understanding of context, this paper suggests four questions, to be answered in sequence: Would it Help? Would it Work? Would it Comfort? and Would it Fly? The “concept applications” from the previous section are used to illustrate the kind of explorations these questions might prompt.

Would it Help?

Technology and data alone can rarely solve real-world problems. They may help us to identify problems, to implement solutions, or to indicate when those solutions aren’t working. But they may also make problems worse, result in solutions being misunderstood, or break otherwise viable solutions by overloading them. This question explores the processes around any data/technology proposal:

- Do we know what those processes are, or need to be?
- Do we have the resources and skills – both internal and external – that they will need?
- What will happen to the processes if the data/technology proposal works?
- What if it fails, or is too successful?

Looking at the “concept applications” from this perspective, both Capacity Warning and Bottleneck Detection require an alternative location or route available to direct students to, so that social distancing can be restored. This might be a nearby backup lecture theatre with a video link; but if the alternative is to send students back to their rooms then travel time is likely to mean they miss some or all of the live lecture, so there needs to be equivalent provision for them to catch up. Whatever the alternative is, students must move on quickly, otherwise a warning of limited capacity inside the lecture theatre could simply create a bottleneck outside. Either of these approaches will work best as part of a capacity planning approach that tries to design out the problems – by creating one-way systems, reducing class sizes, or staggering start times – then uses the technology to identify where those plans need to be adjusted: for example, whether movement routes need to allow for denser traffic at the end of a lecture or the start. Ideally, if we get our space design right, we would have the warning technology but never need to use it.

While Capacity Warning and Bottleneck Detection might have some value as standalone signage systems, Contact Recording and Bubble Bursting are essentially useless without a much broader policy and support framework. What action do we need those identified by the data to take, how will we inform them, and how will we support that altered behaviour? This may well involve external partners: from infection testing to delivery of food and other requirements to those self-isolating. Do those services have sufficient capacity? Internally, if our space planning relies on the use of study-bubbles, how will that be disrupted if some or all members of a bubble need to social distance or self-isolate? Sensitivity is critical to both these applications: when will we use the data to trigger an alert? The Ethics Advisory Board for the UK Government’s contact tracing app identified the tension between using self-diagnosis for speed, but at the cost of many false alarms, as against waiting for a virological test.[46] An unknown device appearing in a bubble might simply mean that a student has upgraded their phone. What percentage of false alarms should we design processes and messaging for, bearing in mind that false alarms will mean people being told to change behaviour, perhaps very significantly, when it was not necessary?[47]

Would it Work?

Now we know what processes the data/technology will need to support, we can examine whether it can deliver what is needed, in both technical and practical terms. What proportion does it need to cover – of campus occupants, of spaces, or of those occupants' daily lives? If the processes need individuals to change their behaviour, will they be willing and able to make those changes? If they depend on individuals *not* changing behaviour, in response either to the process or to external factors, how reliable is that? If changes – to either behaviour or technology – are needed, are those going to be permanent, or something we can set a fixed end, or at least review and renewal, date for?

All of the concept applications use location data from WiFi access points, so an obvious question is whether that is sufficiently accurate for the purpose. Most devices will connect to the access point with the strongest signal:[48] that will usually be the closest one but radio propagation or compatibility may result in choosing one further away, perhaps as far as 100m depending on technology.[49] Simply looking at which access point was used will therefore give only a coarse estimate of position. This may be sufficient for Capacity Warning, where the walls of a lecture theatre may give sufficient isolation to distinguish "in" from "out", but almost certainly not for Contact Tracing or Bubble Bursting, where we need an accuracy similar to the likely infection distance. For Location-Specific Information, access point data might be sufficient to identify when a device has entered a building or lecture theatre, but more accurate location information could be obtained within the device itself: either by an app using all the location sources available to it (which typically include GPS, mobile phone and wifi signals[50]), or simply by having QR codes at entrances for visitors to scan.

---

[46] Ethics Advisory Board (CV19 App), "Letter from Ethics Board to Secretary of State (24 April 2020)" available at https://nhsbsa-socialtracking.powerappsportals.com/EAB%20Letter%20to%20NHSx.pdf (accessed 17 August 2020) p. 3.

[47] BBC, "COVID in Scotland: Hundreds of Dundee students told to isolate" (23 September 2020) available at https://www.bbc.co.uk/news/uk-scotland-tayside-central-54261668 accessed 23 September 2020).

[48] Cisco, "Meraki Documentation: Roaming Technologies" (2020), available at https://documentation.meraki.com/MR/WiFi_Basics_and_Best_Practices/Roaming_Technologies (accessed 17 August 2020).

[49] Utilize Windows, "Overview of 802.11 Wireless Standards" (2011) available at https://www.utilizewindows.com/overview-of-802-11-wireless-standards/ (accessed 17 August 2020).

[50] Stephen Lawson, "Ten Ways Your Smartphone Knows Where You Are" (IDG News Service, 6 April 2012) available at https://www.pcworld.com/article/253354/ten_ways_your_smartphone_knows_where_you_are.html (accessed 14 August 2020).

Many wireless management platforms have options to calculate the location of devices much more accurately, by comparing the strength of the device's signal as seen from different access points (a process known as 'trilateration'[51]). If enabled – usually at additional cost – this might be sufficiently accurate for Contact Tracing and Bubble Bursting, but this would depend on the technology and the locations of access points in the area of concern. Even accurate figures need to be placed in context: data might well show two people spending several hours less than two metres apart, but if they are in adjacent bedrooms, there is little risk of infection.[52] Bottleneck detection is typically done using trilateration, but an initial estimate might be possible using simple connection data and looking for peaks in connected device counts that move between adjacent access points. If there are two hundred devices inside a lecture theatre and none outside, a minute later two hundred outside and none inside, and only one door, we might wonder whether social distancing was being maintained.

Some of the concepts require additional information. If we are using numbers of transmitting devices to estimate numbers of people (as in Capacity Warnings) we need to know what the conversion factor is, and whether it varies significantly between groups. For Contact Recording we need to be able to link devices to individuals – though ideally we would only do this at the stage of using the data to trace an infection risk; for Bubble Bursting we also need to know which devices belong in a bubble.

For the concepts that try to predict individuals' infection risk, we also need to consider what proportion of that risk will be covered by our data, and what a useful proportion would be. WiFi data is likely to cover most students and staff when they are on campus: but how much of their infection risk is incurred off-campus, or from people not carrying or not connecting wireless devices? For Contact Recording to suppress infection rates, it has been suggested that 56% of interactions need to be covered.[53] An individual wishing to know when to seek a test might still be interested at a lower percentage. But if students are relying on a bubble to safely relax social distancing behaviour, a higher rate of coverage might be needed.

Finally, if we plan to advise students or staff to change behaviour, can they afford – both financially and educationally – to do that? If there are strong counter-incentives,

---

[51] Scott Bell, Wook Rak Jung, and Vishwa Krishnakumar, "WiFi-based enhanced positioning systems: accuracy through mapping, calibration, and classification" (Association for Computing Machinery, 2010) in *Proceedings of the 2nd ACM SIGSPATIAL International Workshop on Indoor Spatial Awareness (ISA '10)* 3–9. doi:10.1145/1865885.1865888.

[52] Ross Anderson, "Contact Tracing in the Real World" (LightBlueTouchpaper, 12 April 2020) available at https://www.lightbluetouchpaper.org/2020/04/12/contact-tracing-in-the-real-world/ (accessed 17 August 2020).

[53] Leo Kelion, "Coronavirus: NHS Contact Tracing App to Target 80% of Smartphone Users" (BBC, 16 April 2020) available at https://www.bbc.co.uk/news/technology-52294896 (accessed 17 August 2020).

such as loss of income or disruption to study, then voluntary compliance with that advice is much less likely.[54]

Would it Comfort?

Perhaps most important of all in current circumstances, will the data-enhanced processes we have now sketched be seen as reassuring by already stressed students and staff? Consensus may be particularly difficult when staff and students have different stresses: students worried about their learning, qualifications and finances if the campus does not return to near-normal, staff worried about their health if it does. But the risk of failure – even of adverse consequences – is high if any group is more worried as a result of the proposed solution than they were before. Might these respond by refusing to participate: by time-consuming legal means of objecting or withdrawing consent, or by counter-productive practical ones such as avoiding the spaces and services we have designed to be safe or turning off the devices we need to use for emergency communications? Avoiding these outcomes is likely to require discussion both of what safeguards can be applied (trustworthy systems), and of how to communicate them (trusted systems).

Technological safeguards are somewhat limited by using data that have already been collected: extreme technical data minimisation, such as can be built into de-centralised contact-tracing apps, is unlikely to be possible.[55] However there should be considerable scope for process and policy safeguards: for example data disposal periods related to infection processes will be much shorter than those of the security and fault-finding processes for which the data are currently collected (Harvard retain data for a maximum of 28 days for their contact recording system).[56] For communications, a helpful starting point may be the choice of an appropriate legal basis for any processing of personal data: actions that are clearly linked to narrowly-defined health and safety or anti-virus provisions are likely to be more reassuring than those that serve a vague – and perhaps less clearly time-constrained – public interest.

Three of the concepts – Capacity Warnings, Bottleneck Detection and Location-Specific Information – should be immediately and visibly reassuring to individuals that their health is being considered in campus planning. Contact Recording and Bubble-Bursting are less visible, and require individuals to make a trade-off between what might be perceived as "surveillance" (though it should be short-term and time-limited) and health benefits. Bubble-Bursting, in particular, requires continuous monitoring of

---

[54] BBC, "Coronavirus: Test and Trace Puts People in 'Difficult Position'" (BBC, 7 August 2020) available at https://www.bbc.co.uk/news/uk-england-manchester-53693290 (accessed 17 August 2020).

[55] European Data Protection Board, "Guidelines 04/2020 on the use of location data and contact tracing tools" (21 April 2020) available at https://edpb.europa.eu/sites/edpb/files/files/file1/edpb_guidelines_20200420_contact_tracing_covid_with_annex_en.pdf pp11-12 (accessed 14 August 2020).

[56] Sydnie Cobb, "Harvard to Track Affiliates' Wi-Fi Signals as Part of Contact Tracing Pilot" (The Harvard Crimson, 2 August 2020) available at https://www.thecrimson.com/article/2020/8/2/tracefi-wifi-contract-tracing-coronavirus/ (accessed 17 August 2020).

both location and inter-personal contacts. It may be helpful to point out the direct personal health benefits, as well as the societal benefits that have been the focus of national publicity. Both these concept applications are fragile to obstructive behaviour by those who object to them, whether by changing device configuration, swapping devices, making false reports, or simply ceasing to follow guidance. Some of these behaviours could even increase the risk of virus transmission, if individuals deliberately move to locations where their contacts will not be visible in our data, or its severity, if testing and treatment systems become overloaded by false alarms. Clearly this is an outcome we should take great care to avoid.

Would it Fly?

The final stage is to take a broader view, considering effects on, and from, the whole campus and its surroundings. Effects on individuals will be identified and assessed as part of the Data Protection investigation that should follow the selection of proposals. But given Selwyn's "concerns … over the unequal agency that individuals and social groups have when engaging with digital data",[57] we should also review what inequalities may be inherent to our proposal. How might it affect students living at home, in private accommodation, in residences, or from overseas? Are there different effects – for either staff or students – as a result of different study patterns? What about those at increased or decreased risk from infection? Differences may be inevitable, even desirable if they accurately reflect different needs or risks, but we should be conscious of what they are. More problematic may be any differences arising out of different access to, or use of, technology, digital/data literacy, or language. These we should do our best to address.

An inevitable consequence of using WiFi transmission data is the risk of discriminating against those who do not carry transmitting devices, or who have turned them off, for example because they work in medical or other settings where mobile phone signals are unsafe. Jisc's 2019/20 Digital Experience Insights Survey found that 17% of students did not have a smartphone.[58] If these people are invisible to a Capacity Warning or Bottleneck Detection application it may not matter, provided they are evenly distributed and can be allowed for in the device/human conversion ratio. But if they, and those they encounter, miss out on the protection available from Location-Specific Information, Contact Recording or Bubble-Bursting, this is much more of a problem.

Next Steps

Once a small number of viable ideas – likely to help, work, comfort and fly – have been identified, these should be investigated in more detail, to ensure that they can be implemented and operated in trustworthy, and trusted, ways. Even if the ideas do not appear to involve processing personal data, or to have effects on individuals, Data Protection law provides a useful framework, as well as ways to demonstrate that

---

[57] Selwyn, "What's the Problem with Learning Analytics?" *supra* n. 6, p.13.

[58] Jisc, "Student Digital Experience Survey" *supra* n 16.

effects and side-effects have been properly considered. Information gathered, and discussions held, while considering the “four questions” will provide much of what is needed for this investigation. For example, applying the ICO’s “12 steps” model:[59]

- Awareness: identifying which of the organisation’s processes might be enhanced or supported by the use of data (“Will it Help?”) should have raised awareness among relevant stakeholders of what is being proposed;
- Data Protection Impact Assessment/Data Protection by Design: investigating “Will it Work?” in the technical and organisational context should have resulted in at least a high-level understanding of the technical, process and procedural requirements; investigating “Will it Comfort?” should have identified risks and, at least, begun a consultation process. These can be further refined using Data Protection by Design and Data Protection Impact Assessments (DPIAs): the latter likely to be a legal requirement for many COVID-related applications, and a reassuring part of communications even when they are not. Jisc’s Intelligent Campus Toolkit contains tools both for assessing the level of assessment required and for conducting DPIAs.[60] Since the pandemic is unlikely to have been foreseen in any statement of purpose or privacy notice, Purpose Compatibility is likely to be a relevant tool in determining what new information needs to be provided;
- Information Lifecycle: the understanding of process gained during “Will it Help?” should naturally provide most of the information needed for the information lifecycle. In particular it should allow realistic time limits to be set, both for the retention of data and for when temporary measures should be reviewed and, if no longer needed, withdrawn;
- Legal Basis: as discussed above, the likely legal basis for COVID-19 measures is likely to be public interest or legal duty. Considering “Will it Work?” may, however, have identified solutions or data collection that can be offered on a voluntary, opt-in, basis. Once the appropriate bases are identified, the organisation should confirm that its processing will still meet the relevant legal requirements;
- Privacy Notices: “Will it Comfort?” and “Will it Fly?” should have identified the kinds of information that campus occupants need to have confidence in the proposal, and some methods by which these could be presented;
- Individual Rights (including Subject Access): As for Privacy Notices, “Will it Comfort?” and “Will it Fly?” should identify the kinds of remedies that occupants need. However the aim should be to achieve sufficient consensus and trust that remedies are only used in unforeseen circumstances: they should not be relied on as the main way of discovering – too late – problems with the choice, design, or implementation of solutions;

[59] Valerie Surgenor and David Gourlay “12 Steps to GDPR Compliance” (Macroberts, 25 May 2018) available at https://www.macroberts.com/insights/gdpr/12-steps-to-gdpr-compliance/ (accessed 21 September 2020).

[60] Jisc, “Intelligent Campus: Data Protection Impact Assessment” (August 2020) available at https://repository.jisc.ac.uk/7972/1/data-protection-impact-assessment-toolkit.pdf accessed 21 September 2020.

- Consent Processes (including for Children): if consent is used as a lawful basis then “Will it Comfort?” and “Will it Fly?” should have identified the kinds of mechanism by which this can be sought, granted and withdrawn.

Finally, the kinds of discussion prompted by the four questions should make a significant contribution to meeting the GDPR’s Accountability principle. Clear accountability is essential if activities are to be trustworthy, and trusted.

Conclusion: Benefitting from Data Reuse

Reopening campuses safely during a pandemic will require new systems, processes and behaviours. Data and technology may well have an increased role in supporting and informing these. However this use falls at the nexus of several existing fears: about dataveillance, about career opportunities and about health. Simply asserting “public interest” might even make things worse. Addressing individual concerns – building trust that solutions are effective and not themselves harmful – will need discussion and agreement between those who manage campuses and those who live and work on them. Do we, as Ferguson and Selwyn warned, understand the context sufficiently well that data’s inevitable simplification will not matter? Does our understanding of the incentives and pressures for students and staff give us confidence that technology will be seen as a reassurance, rather than a threat? Or will students or staff, as Edwards warns, actively resist it? This paper has suggested a framework to try out ideas: to increase the chance of our uses of data and technology calming, rather than exacerbating, those fears.