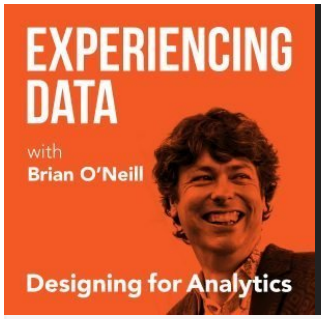
The Experiencing Data Podcast Presents:

036 – How Higher-Ed Institutions are Using AI and Analytics to Better Serve Students with Professor of Learning Informatics and Edtech Expert Simon Buckingham Shum



036 – How Higher-Ed Institutions are Using AI and Analytics to Bet...

00:00 / 00:45:34

1X

Download file | Play in new window | Duration: 00:45:34

Subscribe: iTunes | Stitcher | Google Play | Spotify

Simon Buckingham Shum is Professor of Learning Informatics at Australia's University of Technology Sydney (UTS) and Director of the Connected Intelligence Centre (CIC)—an innovation center where students and staff can explore education data science applications. Simon holds a Ph.D from the University of York, and is known for bringing a human-centered approach to analytics and development. He also co-founded the Society for Learning Analytics Research (SoLAR), which is committed to advancing learning through ethical, educationally sound data science.

In this episode, Simon and I discuss the state of education technology (edtech), privacy, human-centered design in the context of using AI in higher ed, and the numerous technological advancements that are re-shaping the higher level education landscape.

Our conversation covered:

- How the hype cycle around big data and analytics is starting to pervade education
- The differences between using BI and analytics to streamline operations, improve retention rates, vs. the ways AI and data are used to increase learning and engagement

- Creating systems that teachers see as interesting and valuable, in order to drive user adoption and avoid friction.
- The more difficult-to-design-for, but more important skills and competencies researchers are working on to prepare students for a highly complex future workplace
- The data and privacy issues that must be factored into ethical solution designs
- Why "learning is not shopping," meaning we the creators of the tech have to infer what goes on in the mind when studying humans, mostly by studying behavior.
- Why learning scientists and educational professionals play an important role in the edtech design process, in addition to technical workers
- How predictive modeling can be used to identify students who are struggling—and the ethical questions that such solutions raise.

Resources and Links

Designing for Analytics

simon.buckinghamshum.net

Simon on LinkedIn

#experiencingdata

Designing for Analytics Podcast

Quotes from Today's Episode

"We are seeing AI products coming out. Some of them are great, and are making a huge difference for learning STEM type subjects— science, tech, engineering, and medicine. But some of them are not getting the balance right." — Simon

"The trust break-down will come, and has already come in certain situations, when students feel they're being tracked…" — Simon, on students perceiving BI solutions as surveillance tools instead of beneficial

"Increasingly, it's great to see so many people asking critical questions about the biases that you can get in training data, and in algorithms as well. We want to ask questions about whether people are trusting

this technology. It's all very well to talk about big data and AI, etc., but ultimately, no one's going to use this stuff if they don't trust it." — Simon

"I'm always asking what's the user experience going to be? How are we actually going to put something in front of people that they're going to understand…" — Simon

"There are lots of success stories, and there are lots of failure stories. And that's just what you expect when you've got edtech companies moving at high speed." — Simon

"We're dealing, on the one hand, with poor products that give the whole field a bad name, but on the other hand, there are some really great products out there that are making a tangible difference, and teachers are extremely enthusiastic about." — Simon

"There's good evidence now, about the impact that some of these tools can have on learning. Teachers can give some homework out, and the next morning, they can see on their dashboard which questions were the students really struggling with." — Simon

"The area that we're getting more and more interested in, and which educators are getting more and more interested in, are the kinds of skills and competencies you need for a very complex future workplace." — Simon

"We obviously want the students' voice in the design process. But that has to be balanced with all the other voices are there as well, like the educators' voice, as well as the technologists, and the interaction designers and so forth." — Simon on the nuance of UX considerations for students

"...you have to balance satisfying the stakeholder with actually what is needed." — Brian

"...we're really at the mercy of behavior. We have to try and infer, from behavior or traces, what's going on in the mind, of the humans we are studying." — Simon

"We might say, "Well, if we see a student writing like this, using these kinds of textual features that we can pick up using natural language processing, and they revise their draft writing in response to feedback that we've provided automatically, well, that looks like progress. It looks like they're thinking more critically, or it looks like they're reflecting more deeply on an experience they've had, for example, like a work placement." — Simon

"They're in products already, and when they're used well, they can be effective. But they can also be sort of weapon of mass destruction if you use them badly." — Simon, on predictive models

Transcript

Brian: Welcome back to Experiencing Data. This is Brian O'Neill, and today I have Simon Buckingham Shum on the phone from Australia. How's it going?

Simon: Good morning, Brian, or good evening to you.

Brian: Yeah, exactly. We were joking around, you're already in the future. And for those listening, he can report that tomorrow is doing just fine, despite everything that's going on in the world, so good news to report from the future. Thank you so much.

Simon: We're still here.

Brian: You have an interesting title here. You're a professor of learning informatics and director of the Connected Intelligence Center at the University of Technology in Sydney. So, to a layperson or a lay analytics person who may not know much about learning analytics and data in the field of education, which is our topic today, tell us what that means. Like, what is the day in the life of Simon look like?

Simon: [laughing]. Right, well, so yes, I'm an academic. Most of the people you talk to are more in industry in business and government. But, we are very applied here at UTS. We're interested in how the Big Data revolution and analytics, and blah, blah, blah, what does that actually mean for education or learning. And of course, when we say learning, we're not just talking about what goes on in school or in university, it's lifelong learning these days, everybody is continually upskilling, so, I hope that the things that we're talking about today will be relevant for governments and businesses as well. A day my life is a whole mix of working with my team, who are Ph.D. students and researchers who have got their PhDs. We have data scientists on our team. And our whole job here is as an innovation center for the university, for sharing what we do widely and open source as well. Basically asking, "Well, what is analytics?" and increasingly moving into some AI. But what does that mean for the operation of university? How does this make a difference to business processes, to teaching and learning, to researchers?

Brian: Got it. Part of the reason I wanted to chat with you, and I'd love to hear your perspective on this is that, so you have a Ph.D. from University of York in Human-Computer Interaction, or HCI as we call it, which is in the design field, basically. And now you're working with data scientists and in the analytics field, so tell us about the collision of these two worlds. Like, what perspective do you think that that brings, and do you sometimes feel like an outsider a little bit with your lens being maybe different than some of the people that you bring in? And I imagine it's complimentary, but what's that unique angle that you think your background in HCI brings to the table in your work?

Simon: Right, yeah. So, I mean, I come from the human sciences, I came through psychology, ergonomics, and then into human-computer interaction and was just completely fascinated by what happens when people engage with technology and the different ways that they see the world. And so, that's been my whole worldview since the late 80s. And I've gradually moved into learning and now educational data science. And so, that's the perspective I bring. So, whilst we have the technical people on the team, and in the university who are going to talk machine learning, and statistics, and so forth. I'm always asking what's the user experience going to be? How are we actually going to put something in front of people that they're going to understand, like, putting up a dashboard is not the same thing as providing insight, or providing good feedback? And we can talk a bit more about how we're thinking about that later, perhaps. But, yeah, so human-centered data science for learning. That's our strapline. And that means all the usual things you'd expect from user-centered design, about the way you design, who gets to have a voice in the design, who gets to have a voice in the requirements analysis, and the user experience, all that good stuff, but as academics, we're also bringing quite a critical eye to the whole infrastructure that's emerging. So, we want to ask hard questions about, well, where's our training data coming from? And increasingly, it's great to see so many people asking critical questions about the biases that you can get in training data, and in algorithms as well. And we want to ask questions about whether people are trusting this technology. It's all very well to talk about big data and AI, etc., but ultimately, no one's going to use this stuff if they don't trust it. So, what does it mean to trust an analytics or AI-powered learning infrastructure?

Brian: And we'll probably go into this a little bit later, but is there a particular recent example you can give me of where that trust factor came into play, and your design eye on this—the human factors element—either got you in front of it and prevented that, or use it was a "told you so," situation where it's like, well, this is the risk when we don't think about designing a solution with people involved? This is the risk that we run into. Is anything recent come to mind in your space?

Simon: Yeah, there's lots of examples now emerging and, for example, there are lots of edtech companies now. You can't buy an edtech product without a dashboard of some sort. And so, finally, you will know what your students are doing. But often the system is logging relatively simple things; clicks and page views, and so forth. And many educators are saying, "Well, that's all very well, it doesn't really tell me anything about learning." If the technology is not rolled out in an appropriate way within a school or university, then, again, hard-pressed teachers and educators feel just dumped on. Oh, here's another tool we're supposed to be using and nobody's trained us adequately, and what's this got to do with the kinds of teaching and learning and assessment that I'm interested in? We are seeing AI products coming out. Some of them are great, and are making a huge difference for learning STEM type subjects, you know, science, tech, engineering, and medicine. But some of them are not getting the balance right. And we've got parents and teachers and students literally on the streets protesting that they're spending

way too long in front of screens, and they're not engaging with each other and with their teachers enough, and they're not picking up the skills they need for future jobs. So, people are getting it wrong. And that's all part of the hype cycle around big data and analytics, which is now starting to pervade education. The way we're trying to get it right, the way that people are getting it right, is really engaging the computer science and the data science with the learning science. That means really understanding how are people learning? What are they needing? What do teachers really need to know about students' progress? And here at the university, because we have the capacity within my own center, we are developing and adapting software, we have control over the code. So, we can actually bring educators in and go through co-design sessions with them. We can bring students in and use participatory design techniques which don't require high technical knowledge, which gives them a voice in envisioning what the software should be doing.

Brian: Got it. When we met, I was really fascinated by this area, and so I actually went out and called up the head of data governance and data at my university and just to do a little research and learn about what they were—I figured, well, I might as well start where I went to school—and see what they were doing there. And it was pretty interesting. She reported to me that the President, I think, had an accounting background, so was very interested in using data to make decisions across the university and so they've designed that dashboards and apparently all the deans are required to use these dashboards, and so I started asking her how did they go about figuring out what stuff to measure and what the right dashboards were, and she said, "Well, we do training and the President has decreed, these are the metrics and KPIs that we're going to measure, and everybody will be using these." And apparently, she said, there's a lot of willingness from the deans to attend the training, and a couple things concern me there. A) I always get worried when there's a lot of training required for technology like this, and B) I couldn't help but wonder if the geology department and the performing arts department really need to be tracking exactly the same metrics. And it's like, I just wonder if we're tracking stuff that's trackable. I mean, I'm sure there's some practical use with enrollment numbers and some of the financial aspects of running a university, etc. But, I don't know, I don't know, because I haven't talked to deans from multiple departments, but I would just think that even with those two examples, there must be different needs, right? And is this still at a very basic level where we're tracking low-level data and trying to inform intention from that, or—any thoughts on that?

Simon: Oh, yeah. So, we spent a lot of time explaining to people that there are some quite different kinds of analytics in the world of education. So, the kinds of things you two talked about, where a university is essentially using business intelligence, the way that any organization uses BI, right, to try and track some high-level indicators about how things are going. So, the extent that, a big university or school is like any other organization that has business processes and KPIs that give you a sort of bird's eye view of how things are going, well, universities can make use of that and, the fact that you're example involve

deans, well, deans operate at a very senior level, and they're just looking for some key outcomes like student numbers in, [crosstalk 00:11:09] the end. How are finances going? How many of them are getting jobs at the end? What are the demographic breakdowns? So, these are the kinds of things that you'll find in any educational organization, the standard kind of metadata you'd have about any student. Okay. And so, that's fine. That's BI for organizations and for managers. Now down where the sort of exciting stuff is, really, is about tracking students, day by day, week by week, in order to try and pick up quickly enough how things are going to make a difference. Right? So, it's all very well for the dean to be seeing that we have this dropout rate by the end of the semester. But how about data and analytics to actually improve retention, to improve engagement, to improve the depth of learning? And this is where a lot of excitement is, about closing the feedback loop to teachers and students about how things are going. And of course, in big data in any sector, all the excitement about that is about studying a complex system, trying to see what it's doing in time to make a difference, not just conducting a post mortem afterwards. So, when we translate that into teaching and learning, we are trying to track how the system is going, and whether we can close the feedback loop with students to empower them, as well as the teachers. So, ideally, we don't want these tools to just be perceived by students as surveillance tools that are tracking them at high fidelity, and we could be talking about anything from every click they make to their movement around campus because we can geolocate them. The trust break-down will come, and has already come in certain situations, when students feel they're being tracked and this seems to be no payback for this intelligence that the university has gathered.

Brian: Yeah, I would also think that there's a fair amount of friction to overcome here, with this technology being adopted by the teachers as well. I mean, you've got—there's so many stakeholders here, right? I mean, you've got students, you've got faculty, you have the administration of the school, who owns the data? The vendors, I mean, I guess a lot of this maybe exists in business capacity, too, but I wonder if universities aren't supposed to be primarily profit-driven, how do you get teachers—I don't get too far off on that, but—

Simon: Wash your mouth out with soapy water.

Brian: Exactly. Purest me of thinking that way. But tell me about the teacher friction here, in terms of getting them to actually, A) care about this, believe about it, participate in the design of a system. I could see this being difficult to get adoption when maybe they don't—Are we at such an elementary level with this technology right now, where the first technology output that we get that's supposed to be assisting with the learning process is so low-level still, that it may be perceived as just a tax on the teachers time? And—

Simon: No, I mean, there are lots of success stories, and there are lots of failure stories. And that's just what you expect when you've got edtech companies moving at high speed. And to get product out there

quickly, analytics is the current buzzword, so everyone's got to have analytics, and less attention is going into actually, is this useful, [inaudible 00:14:49] products out there, and people pick up that it's all being hyped, then you'll get natural skepticism. And educators are obviously smart people and they don't like being dumped on with bad technology any more than the next person. So, we're dealing, on the one hand, with poor products that give the whole field bad name, but on the other hand, there are some really great products out there that are making a tangible difference, and teachers are extremely enthusiastic about, and our job here at the university and also, I work with the Society for Learning Analytics Research, SoLAR, which is a global organization to try and educate and raise the quality of dialogue and the conversation amongst educators and researchers and the vendors, so there's incredibly useful stuff here that teachers will get very enthusiastic about when it's quite clear that it's giving them some value. So, there's good evidence now, about the impact that some of these tools can have on learning. Teachers can give some homework out, and the next morning, they can see on their dashboard which questions were the students really struggling with? Well, let's just prioritize our time this morning on reviewing those questions, and not the ones that everyone's already know. Students can progress at their own pace with certain kinds of adaptive platforms. So, that if they're struggling with something, it's going to rewind, it's going to diagnose that they're really struggling with a threshold concept of some sort, it's going to walk them through more slowly. Or if, at the other end of the spectrum, we're going to stretch and challenge you. It's really useful to know. I've set this pre-work before we have our exciting collaborative session, but it's not going to work if the students haven't watched the video, or read the PDF, you know? And we can ask all sorts of interesting questions just about, did they do the pre-work? Did the students who did the pre-work a good week before the session do much better than the ones who are cramming the night before? Has the student mastered these core concepts? As someone with this track record in her subjects likely to struggle with this new subject that they're contemplating? So, these are questions which are extremely important and relevant for students and teachers. The area that we're getting more and more interested in, and which educators are getting more and more interested in, are the kinds of skills and competencies you need for a very complex future workplace. Your collaboration ability, your ability to think critically, your ability to reflect on how you're developing as a professional in training at university. So, these kinds of competency, your appetite for risk and uncertainty—because that's what we can guarantee the future holds, and pretty much we don't know what else is coming—so we're very interested in diagnostics, which will give feedback to teachers and students about these kinds of competencies and that's sort of the leading edge of the field at the moment.

Brian: Got it. Got it. You said in one of our email exchanges, in the kind of edtech world, you said there's a challenge of the, quote, "ignorant users." Can you unpack that for us a little bit here?

Simon: Well, yeah, this is one of the interesting things about working in learning compared to most other contexts. So, if you're doing user-centered design or participatory design, or whatever you want to call it, how do you find out how people work? How do you find out what the activities and the workflows and the work practices are, that your system is going to have to support and hopefully not completely disrupt? Well, you talk to the users. But in learning, we've got this rather interesting situation, because obviously, you've got students, they can tell you what their experience of studying of course is, they're the world expert on their experience studying, but they're not expert learners. Students are not necessarily very good at studying. They don't really often have very good revision practices. They're obviously not experts in the subject matter, because that's why they're learning. And nor are they very good at teaching. I mean, they're not expert educators, they don't know what's good for them. They know what they like, they don't necessarily know what's good for them. So, we have this interesting situation where the quotes, you've got inexpert users compared to most situations where the user is considered to be quite authoritative about what their work entails, and who they need to talk to, and what's important. So, we obviously want the students' voice in the design process. But that has to be balanced with all the other voices are there as well, like the educators' voice, as well as the technologists, and the interaction designers and so forth.

Brian: Mm-hm. Yeah, I feel like that's maybe not just localized to the learning space, I feel like one thing that I find myself talking to clients about or even in training context is unpacking or discovering the latent problems. So, because so much of the time, you're not handed this requirements document that perfectly outlines, this is precisely the problem, and the outcome that I want, you never get that. You usually get something on the surface, and you have to balance like satisfying the stakeholder with actually what is needed, and to me, that process is usually trying to get an agreement on this hidden thing that you've discovered and getting the team to agree that, actually, this is the problem space that we're operating in. Do we agree this is really what we want to solve for? And it takes some soft skills sometimes to—

Simon: Absolutely, yeah.

Brian: —negotiate that. Is that how you [crosstalk 00:20:44].

Simon: Yeah, so we have to deal with all of that, you know that users often don't know what they want, or they don't know what the technology is capable of. And so, that's a conversation that we had and when they realize, "Oh, you can do that now, can you? Oh, well, in that case, that opens up all sorts of interesting possibilities." So, we have to deal with all of that, plus the fact that students are not experts in the domain, and aren't necessarily experts in what their work practices should be.

Brian: There is another interesting quote in your email that our exchange, and you had said, “Learning is not shopping.”

Simon: Well, right.

Brian: I mean, you can shop for schools, we all know this, Newsweek or whatever. [laughing].

Simon: Yep. Well, again, so learning is not shopping. So, this is shorthand for—we’re talking about something which is going on inside the mind, okay? And although people are starting to slap skull caps on students in experimental labs to try and understand what their brainwaves are doing, and well, we even see in China, they’re rolling out some of this stuff at scale in really worrying ways. Basically, like any data analytics situation, we’re really at the mercy of behavior. We have to try and infer, from behavior or traces, what’s going on in the mind, of the humans we are studying. If it’s shoppers, well, we know they bought something, and we’re going to take a guess they might buy something else. Those are fairly discrete transactions, and we know in the end whether we are correct or not. Did they buy this thing or did they not? So, when it comes to learning, well, this is rather a more complicated thing. It’s in the mind. It’s social as well. It’s emotional, as well as cognitive. Yet, we’re working off activity traces that are logged by online platforms, or other kinds of sensor technology, which can actually track physical embodied activity, and we have to figure out, are they learning? So, that just makes things rather more interesting as well as complicated. And this is why we also need to have learning theories as well, the death of the theory really does not apply in the world of learning. We need to have theories which explain how to tie together the kinds of activity traces we’re seeing, and how they might point to whether someone is actually learning in the ways that we want them to. And that’s a big part of the whole R&D effort field as well. We need learning scientists and educational searchers as well as the technical people.

Brian: And is that like—can you kind of, on our theme of the future, since you’re from the future, what does that look like, the destination, and how far we are from getting to the point—I assume the leaps look like this—there’s, there’s some learning model that starts to get a little bit more concrete, and then at some point, we try to encode that in our technology solution so that we can actually start to then go and measure it. Like, if those are—I don’t know if those are the correct three bullet points of what that journey looks like, but where are we on a timeline with those?

Simon: Yeah, well, there aren’t many products that do this yet. But where we are now is that they are methodologies being developed so that we can say, “Okay, if we see this, this, this, and this happening in the data traces, we’re going to hypothesize that that’s a property for some higher-order construct.” And this happens, of course, in any area of data science where you’re trying to build a bridge between low-level activity traces, and higher-order constructs. A higher-order construct might be the likelihood of switching mobile phone provider. Or it might be the likelihood of wanting to buy a house in the next

three months. And people will be building models that map from low-level data traces to higher-order constructs. So, this is what we're doing in the world of learning now. So, we might say, "Well, if we see a student writing like this, using these kinds of textual features that we can pick up using natural language processing, and they revise their draft writing in response to feedback that we've provided automatically, well, that looks like progress. It looks like they're thinking more critically, or it looks like they're reflecting more deeply on an experience they've had, for example, like a work placement. So, we have developed software tools here at UTS, which give instant feedback to students on their writing. We're not talking about spelling and grammar and plagiarism here, we're talking about, are you demonstrating the ability to work with ideas and to reflect critically on yourself, and how you're changing as a person? And we have developed mappings from low-level textual features to these higher-order capabilities, which we're really interested in. Or, we are tracking how a team is working, a team of nurses working together face to face around a patient. That patient is a mannequin streaming data, but we're also tracking the students in that position, and who's talking, and who's using different kinds of equipment ,and administering different kinds of treatment. And again, all that data can be aggregated, and we will then say, "Well, if we see this happening in this kind of way, at this moment in time, in relation to your peers, that looks like you've understood how to do X, Y, or Z." And it's that critical mapping from low-level traces, which is the world of machine sensors, up to higher-order capabilities, which is the world of learning and skillful performance.

Brian: Got it. One thing I was thinking about as you were talking about that is that I would hope, and maybe this is where your design background comes in, and, hopefully, that that skill set as present in other people that are working in this space, but there's some funny opposites here where, in a business context, writing skills, like, short, brief, clear, use clear language, etc., and then you look in the university space, where there's an academic writing style which could be completely different. The vocabulary could be very domain-specific, could be very technical. It could be very, very detail-oriented, and so I start wondering, when we talk about these models for something as basic as writing, how do you—I guess there's different scores, right? Are we talking about how, well, is this student preparing to enter a Ph.D. program is different than how is their written communication skills for the business world?

Simon: That's right, so there are different kinds of writing in the world of, of academics, and so the way that we would use the tool is working really closely with the educators. The educators are trying to get the students writing a decent business analysis, or a decent reflection on their workplace, in a pharmacy or in an engineering company, or a law student putting together a really good essay. So, these are different kinds of writing, but actually underlying it, there is some sort of DNA that we have identified, that shows evidence of critical thinking, bringing ideas together, contrasting tensions or disagreements, identifying missing knowledge. And actually, these kinds of forms of critical thinking are things that are just as relevant in the business world, and the world of government and policy analysis, or market

analysis, as it is in academia. So, this isn't just about, how do you write to get published in a journal? This is actually about critical thinking and making that thinking visible clearly to the reader. And you're as interested in missing knowledge or expressing surprise about something that you weren't expecting in the world of business and policy analysis as you are in the world of research or studying subjects.

Brian: Yeah, that's—you could really see here just from the picture you just painted how important, if you're using predictive technologies, the training data is to be able to say this is a this is an excellent legal brief, or this is an excellent business case study, or this is an excellent piece of poetry. How do you—modeling that stuff such that if you're going to score on it later or predict a student's aptitude, that seems really, really important. But it's interesting that you're finding, if I understood it, you're finding some commonalities across all the different domains from the work, is that correct?

Simon: There are certainly some commonalities across the domains. I should emphasize that we're not going anywhere near poetry or creative writing. [laughing]. Because, you know, machines really don't do that very well.

Brian: Yeah.

Simon: We are talking about more analytical forms of writing, where you're making an argument of some sort. Or reflective writing where, increasingly, we get students—and this is also done in senior leadership development as well, coaching, to write reflectively about how they're making sense of a challenging experience. What they're learning, how they might be changing, [inaudible 00:31:01] and so forth. And those are kinds of writing that span many [inaudible 00:31:07]. Some of the work we do is—in fact, this text analytics work is actually not machine learning-driven, it's actually a rule-based grammar. I won't get into the details of that here, but certainly within the field of education, there is interest in predictive models. For example, we're interested in predicting, well, which students might be most in need of some extra support, because it looks like they're struggling. So, we're not involved in that work ourselves, right now, but there are quite a few universities around the world who are very interested in predictive models for sort of spotting at-risk students as they're called. There are some interesting examples where that's really paying off. There are also some examples where it's being used really appallingly, to try and weed out poor students and almost stop them enrolling, just by predicting their statistical likelihood of failure based on demographics. And this is not what education is about, of course. We don't want to be seeing these kinds of technologies closing down opportunities. So, we can use predictive modeling to identify which students might be most in need of support? Or—but then there's some really interesting ethical questions around that, as well. So, well, what do you do if the model is flagging a student red, you can't pick up the phone to them and say, our algorithm thinks you're at risk, okay? That's going to freak the student out. You have to deal very sensitively with that. Do you only call the students that the algorithm is flagging, or do you check in with all the students? Exactly how

is that algorithm working? So, we know, for example, most universities know, that students from a particular area, postcode, zip code, or students with particular demographic profiles, are going to be higher risk because they've got more challenges. They may be the first in their family at university, they're coming from less opportunity, etc. Statistically then there's a shadow over them. But we certainly don't want to classify, pigeonhole, or flag somebody at risk simply because of historical injustices, or inequalities of opportunity. So, there's some really interesting ethical issues here. Once they're on campus, once they are signed up, supposing our algorithms says, "Hm, well the students who got high grades, those are the ones who go to the library a lot. Those are the ones who also seem to write much longer assignments. They also seem to be much more enthusiastic buyers of university merchandise like the sweatshirts from the shop. So, we've got predictive variables that correlate with high outcomes. But of course, those are not necessarily causal. Buying the university sweatshirt is not causing you to perform better. Going to the library, well, okay, there's some causal relationship there, one might assume, but you don't simply tell students to go to the library more, because that's what the high performing students do. Learning is a complex social and emotional and cognitive business. And so, simply because you found correlations does not mean you've got causation. And so, these models, if they're black boxes then we have potentially a problem here because there's a question around accountability. And if we try and explain how they work, that can be also pretty tricky as well. So, these predictive models are a double-edged sword for sure.

Brian: Mm-hm. Do you foresee those as being a ways off before they're going to be properly constructed, adopted safely, understood, interpreted, trusted?

Simon: Yeah.

Brian: Is that a ways off?

Simon: Yeah. I mean they're in products already, and when they're used well, then they can be effective. But they can also be sort of weapon of mass destruction—

Brian: Sure.

Simon: —if you use them badly. It's a culture shift, as in many sectors, moving to making decisions where you've got input from a model is not something many people use today. And there's going to be questions around trust, and accountability, and transparency. And education is no more exempt from those concerns as any of the other sectors which are now coming under scrutiny from the AI ethics and data science ethics type critiques, that's entirely appropriate.

Brian: Do you think design matters in this case. Is it going to make a difference in how this—

Simon: Yeah, yeah, yeah. So, we've talked about—I feel like what we're doing is we're talking about the whole infrastructure, which includes the human infrastructure as well as the technical infrastructure, and at every level of this infrastructure, we have to think critically about that. I'm still really interested in user experience, and how we design these things. And if you have a rubbish user experience with something, then, these days, people have quite high expectations. And so, we absolutely need the best design thinking going into the design of those user interfaces, as well as into the design of the processes that lead to those systems in the first place. Something we're really interested in is data storytelling, and I know that you've spoken with people like Cole Knaflic and others before. So, we take great inspiration from some of that data storytelling work, and we think that it's absolutely critical. You've got a huge amount of data you could present to somebody. Well, how are you going to do that in a manageable, intelligible way? How are you going to foreground the important stuff? How are you going to make sure that a student, if you're going to put up a dashboard or some sort, or a visualization of some sort, or for a teacher, how are they going to quickly see the signal amidst the noise? Their time is precious. They do not have hours to go doing exploratory data analysis, even if they had the skills. A dashboard is meant to be something that you can glance at quite quickly and see the key messages. And so, we've been taking on board some of the great work on information design and data storytelling, and exploring whether we can render and foreground the important messages, so that that really drives the visual attention around the user interface.

Brian: Sure. Yeah, I mean, it seems so multi-factored. There's—leadership has its lens on this whole system. The teachers are obviously integral to this, and of course, the students themselves, and how do they all interact, and who's inputting data [laughing] versus reading out what's happening? How much is the student in control of any of these things? I mean, I could see it being really frustrating if you feel one way, and you see—you find out that you're being flagged by this surveillance system that thinks you're at risk or thinks that you're in a cohort that you don't identify with and not feeling like you can do something about that could be really frustrating. A lot of friction there on multiple levels.

Simon: Yeah. And in one sense, this is no different to any other complex human organization where people are doing interesting stuff, and we've got machines trying to make sense of it. This just has [inaudible 00:38:39] particular forms in education, and as a flag, learning is a particularly complex form of human behavior, which needs handling with care. Of course, when this works well, then the data is a provocation for reflection. And so, what you might think you're doing right, actually, you're not. And, it's our job here at university or in a school to actually confront you with that. That's what you pay for. But let's try and do that in a constructive way, and let's make sure that we understand how this technology is working and what its limitations are. Because there is this aura of omniscience around big data and AI in some quarters. And people can be tempted to just go with the machine's recommendation sometimes because they're busy. And we have to make sure that the humans are in the loop in an appropriate way.

Brian: Sure, sure. Just wrapping up here is there anything particular you would like to share either, from your background in human-computer interaction or from an educational standpoint, that you think could transfer over to the business world, like, any methods of working or, things that you feel are working well in your space that you don't see happening so much in the business world?

Simon: Well, to the extent that businesses and governments and other sectors are interested in their own staff's learning and development, what we're doing in the world of what's called learning analytics or educational data science, that's relevant for you. So, there's a conversation to have there about workplace learning, and how that's different from formal education. But there's definitely relevance of what we're doing to those sectors. In terms of design processes, I don't think we've invented any radical new user-centered design, co-design, or participatory design techniques. What we have been doing is adapting well-known techniques from co-design, for example, using playing cards to structure conversations. We've developed a deck of cards that's particularly tuned to thinking and talking about data, and analytics, and feedback. We've used user journeys, which are a well-known technique mapping the user journey as they enter our organization's digital space, etc. Well, we've done a variation of that for eliciting from learners, what their experience is when they engage in learning where data and analytics might help them. So, we've been doing a bit of adaptation of well known low-tech, high-touch kinds of design techniques. I think perhaps the most interesting work is the stuff that I mentioned earlier, which is, how do you map from low-level machine logs to higher-order human capabilities and skills. That's something that I think organizations will be increasingly interested in doing. Has to be handled with care. There is an emerging field of human resource analytics, which is trying to get at their staff capabilities by using data traces. That's a conversation that I think could be really worth having between the HR world in the world of learning and education, as well, because in that sense, we're actually trying to do something similar, but it's also fraught with difficulty as well. If you make the wrong inference there can be quite high stakes attached to concluding that an employee is really not doing their job based on an analytics dashboard. So, let's walk in with our eyes wide open, but it's a very exciting field. And when it's done well, it has huge benefits. When it's done badly, then it's pretty destructive.

Brian: Simon, this has been really great. Just again, for our listeners, I've been talking to Simon Buckingham Shum, who's a professor of learning informatics and director of the Connected Intelligence Center at the University of Technology in Sydney, Australia. And on that note, Simon, where can people follow your work and what you're doing? Social media, LinkedIn, what's your—

Simon: Yeah, I'm pretty active. Simon.BuckinghamShum.net and go to LinkedIn.com/in/Simon. You can find me.

Brian: Awesome, great. [I will definitely link those up], and thanks for coming on Experiencing Data and sharing all this information about the learning analytics field with us.

Simon: Thanks for having me. It's been really interesting talking to you, and I love the whole podcast series. You have great people and such interesting backgrounds, so very happy to be part of it.

Brian: Well, thank you so much. Yeah. Glad to have you as well.

Simon: All right, thanks very much man.

Brian: Cheers.

More Episodes of Experiencing Data:

- Browse past episodes - each episode has a full text transcript and audio
- Join the mailing list below to be informed when new episodes are released

Subscribe Now

I help data product leaders turn machine learning and analytics into simple, innovative decision support software using human-centered design.

Menu

Privacy & Legal