

What Do We Owe Students When We Collect Their Data - A Response

Autumm Caines

Editor's Note

This was originally posted to [Autumm Caines's blog](#) on October 26, 2018.

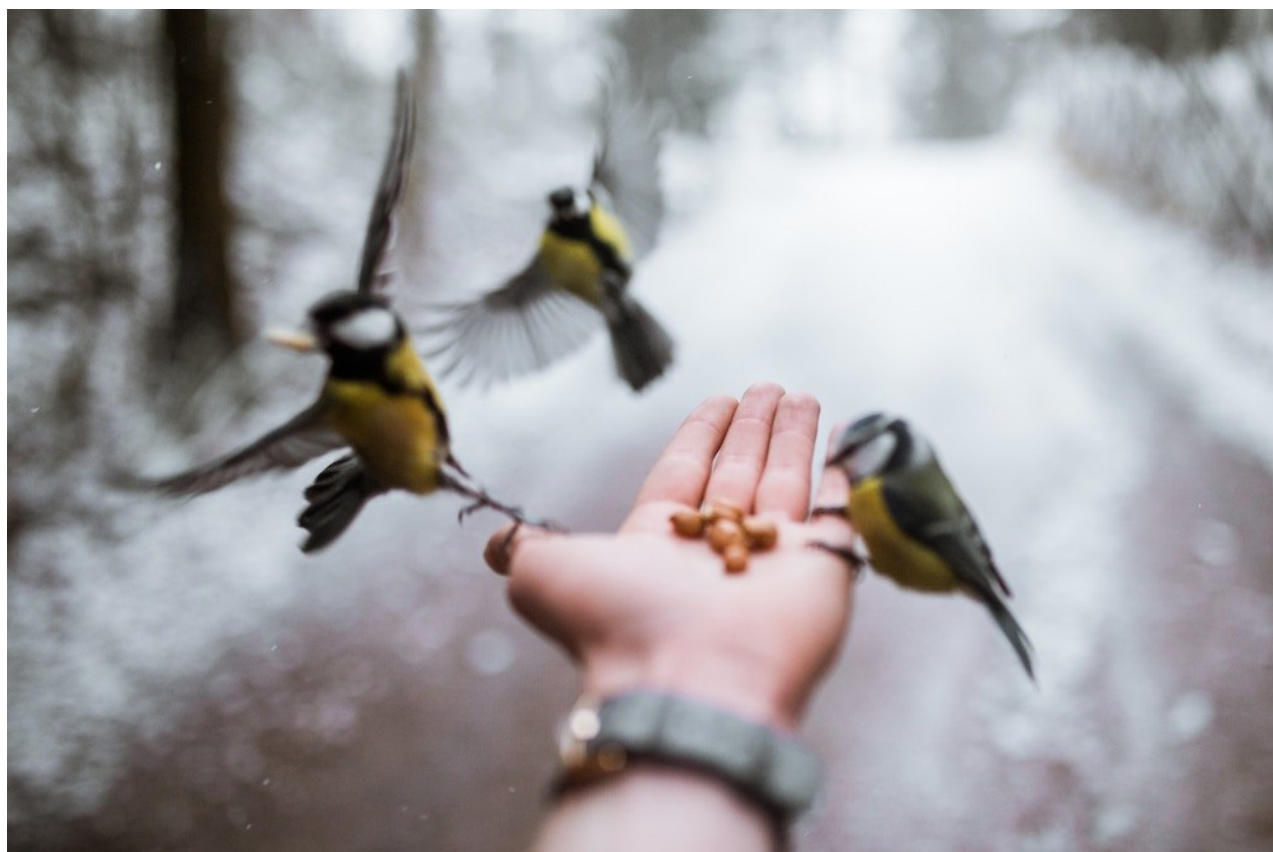


Photo by [Taneli Lahtinen](#) on [Unsplash](#)

It has been a few weeks since we issued our [#DigCiz call](#) for thoughts on the question “What do we owe students when we collect their data?” and there have been a few responses. The call is in conjunction with the [interactive presentation at the EDUCAUSE Annual Conference](#) that I’ll be helping to facilitate with Michael Berman, Sundi Richard, and George

Station. The session will be focused around breakout discussions both onground and online during the session. We don't necessarily have "answers" here – the session (and the call) are more about asking the questions and having discussion. The questions are too big for one session and often there are not easy answers; so we released the call early hoping that people would respond before (or after) the session. I've yet to respond to it myself so I'm going to attempt to do that in this post.

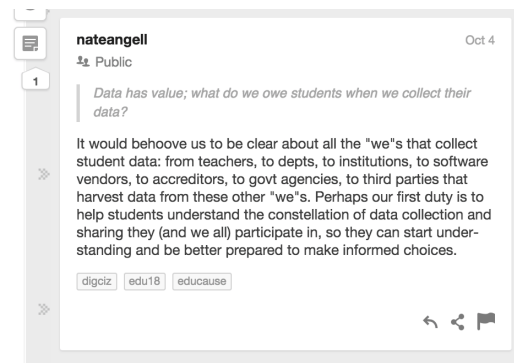
The #DigCiz Call

We want the call to be open to everyone – even those who don't know a ton about student data collection and we want people to respond using the tools and mediums that they like. We have had some great examples already and I wanted to thank those who have responded so far. I threw the call out to some of our students at SNC and I was super honored that [Erica Kalberer responded](#) with an opinion piece. Erica does not study analytics, she is not a data scientist or even a computer science major. She didn't do any research for her post and it is an off the cuff, direct, and raw response from a student perspective – which I love.

Additionally, Nate Angell chose to leave a [hypothesis annotation on the call itself over at digciz.org](#).

- **Data has value; what do we owe students when we collect their data?**
- Does the widespread and routine collection of student data in ever new and potentially more-invasive forms risk normalizing and numbing students to the potential privacy and security risks?
- What is our responsibility as educators to increase students' agency around their own data, so that they can better understand what these data could tell someone about them, ways these data could be used to benefit them or harm them, and what they are giving away when they consent to data collection.
- How can our policy and decision-making frameworks keep up with the ever-increasing opportunity to collect and analyze student data?

Then, on November 2nd at 8am MDT (the time at the conference – which is: 7am PDT / 9am CDT / 10am EDT) we will deliver our session at the EDUCAUSE Annual Conference in Denver, Colorado.



Nate points out that there are many "we"s who are collecting student data and that students often have no idea who the players are that would want to collect their data let alone what data is being collected and what could be done with it. What do we mean when we ask "What do WE owe students...." Who is this we? Instructional designers may answer these questions very differently than accreditors would, or as librarians would, or even as students themselves would. I hope that by hearing from different constituencies that we can bring together some common elements of concern.

Framing Things Up

I am really intrigued by our question but I also have some issues with it.

The question is meant to provoke conversation and so in many ways it is purposefully vague and broad. It is not just "we" that could be picked out for further nuance. So many simple definitions could be picked out of this question. What is meant by "data" and more specifically "student data".

What are we talking about here? Is this survey data? Click data from the LMS or other educational platforms. What about passive and pervasive collection that is more akin to what we are seeing from the advertising industry? The kind of stuff that does not just track clicks but tracks my where the cursor moves, the speed of how my cursor moves, where eyes are on a screen, text that has been typed into a form but has not been submitted. What about if we are using wearables or virtual reality? Does the data include biometric information like heart rate, perspiration, etc. Is this personally identifiable information or aggregate data? Some of these examples seem particularly sensitive to me and it seems like they should all be treated differently depending on context.

We could keep going on.... What is meant by “collect”, “students”, “owe” ... a whole blog post could be written just about any one of these things.

Another of my issues is that the question assumes that student data will be collected in the first place. I’m setting that issue aside for this call and presentation because if I like it or not I am part of field that is collecting student data all of the time. As an instructional designer I make decisions to use technologies that often track data and to be honest if I wanted to avoid those technologies completely I’m not sure that I could. Over the course of my career faculty and administrators have often come to me asking to use technologies that collect data in ways that I consider predatory. How do I respond? How do I continue to work in this field without asking this question?

People who know me or follow my work know that over the last few years that I have often struggled with considering our responsibilities around student data. Even though I have been thinking about these kind of questions for a few years now I don’t think that I will be able to dive into all of the nuance that any of these could bring. (I want to write all the blogs – but time). So, I just have to resolve that – that is why this is a broader call for reflection and conversation and invite others to respond to the call around things that I may have overlooked.

Though I am still new to this conversation, I’m not so new or naive to think that there are not already established frameworks and policies for thinking about the ethical implications of student data collection. I’ve been aware of the work that JISC has been doing in this area for some time and had just started a deeper dive on some research when I attended the Open Education Conference in Niagara Falls a few weeks ago.

Somehow I missed that there were two important data presentations back to back and though I only caught about ¾’s of the [*Dangerous Data: The Ethics of Learning Analytics in the Age of Big Data*](#) presentation from Christina Colquhoun and Kathy Esmiller from Oklahoma State University, I got the slides for [Billy Minke and Steel Wagstaff’s “Open” Education and Student Learning Data: Reflections on Big Data, Privacy, and Learning Platforms](#) which I missed completely.

Both of these presentations looked at different policies and ethical frameworks around using student data which was a goldmine for me. Dangerous Data’s list did not make any claim about quality of the framework’s while the Open Education and Student Learning Data presentation did specifically state that their list was curated for policies that they were impressed by.

Open Education and Student Learning Data listed:

- [The Policy on Ethical use of Student Data for Learning Analytics](#) – Open University
- [CSU Learning Analytics Code of Practice](#) – Charles Stuart University
- [Privacy and Online Monitoring Policy](#) – University of California, Berkeley
- [Data Privacy Practices \[Draft\]](#) – University of California

Dangerous Data listed:

- [The Policy on Ethical use of Student Data for Learning Analytics](#) – Open University (same as above but links to broader set of resources)
- [Code of Practice for Learning Analytics](#) – JISC
- [Responsible Use of Student Data in Higher Education](#) – Stanford CAROL and Ithaka S+R
- [Consensus Principles on User’s Digital Privacy in Library, Publisher, and Software-Provider Systems](#) – National Information Standards Organization (NISO)

My Response

I’ve started reading through the policies and frameworks listed above and while I have not had a chance to dive deep with each one of them, I’ve found a lot of overlap with what I have identified as four core tenets that I believe start answer the question “What do we owe students when we collect their data?” at least for me – for now. I’m personally

identifying with “we”s as in instructional designers, college teachers, IT professionals, librarians (as an official wannabe librarian) and institutions – at least on some level.

I’m still learning myself and I could change my mind but for the purposes of this post I’m leaning on these four tenets. I feel like before we even start I need to say that there are times when considering these tenets, in practice, that the answers to the problems that inevitably arise come back as “well, that is not really practical” or “the people collecting the data themselves often don’t know that”. In these cases I suggest that we come back to the question “what do we owe students when we collect their data?” and propose that if we can’t give students what they are owed in collection that we think twice before collecting it in the first place.

I will list these tenets and then describe them a bit.

- Consent
- Transparency
- Learning
- Value

Consent

This one seems of the most importance to me and I was shocked to see that not all of the policies/frameworks listed above talk about it. I understand that consent is troubled, often because of transparency – more on that in a bit – but it still strikes me that it needs to be part of the answer.

There is a tight relationship between ownership and consent; there is a need for consent because of ownership. If I own something then I need to give consent for someone else to handle it. But not all of these frameworks recognize that. The Ithaka S+R/Stanford CAROL project, listed above, talks about something called “shared understanding” where they basically envision that student data is not owned solely by the student but is a shared ownership between the school, the vendors, and third parties. In a [recent EDUCAUSE Review article](#) some of the framers of the project actually said “the presumption of individual data propriety is wishful thinking”. This, after they put the word “their” in [scare quotes](#) (“their” data) when referring to people being in a place of authority around the data about them. Ouch!

I mean I get what they are doing here. One looks at the Cambridge Analytica/Facebook scandal and says “oh how horrible” but their response is: you are a fool not to realize that it is happening all of the time. And maybe I am a fool but I still think it is horrible. The article points to big tech firms, how much data they already have about us, and how much money they have made with those data and uses it as a justification. But here is the thing, we are talking about students not everyday users. I think that makes a difference.

In another [EDUCAUSE Review article](#) [Chris Gilliard](#) points out the extractive nature of web platforms and the problems of using them with students. What of educational platforms? Is it really okay to import the same unethical issues that we have with public web platforms into our learning systems and environments? I’m comforted that most, if not all, of the other frameworks listed above and those that I’ve come across over the years do understand the importance of consent and ownership.

I’ve read broader criticisms of the notion of consent that I found quite persuasive by [Helen Nissenbaum](#) (Paywalled – sorry) but even she does not abandon consent completely. Rather she points out that consent alone, in and of itself, is not the answer. We need more than just consent – especially now when our culture grants consent so easily and thoughtlessly. Nissenbaum’s criticisms of consent are in thinking of it as a free pass into respectful data privacy. But here I’m thinking of consent in terms of what we owe students – I see it as a starting place and the least of what we owe them.

What do we owe students when we collect their data? We owe them the decency of asking for it and listening if they change their mind.

How we ask for data collection and how we continue to inform students about how it is changing is not easy to answer and I want to be very careful of oversimplifying this complex issue. I think that, at least in part, it also an issue of my next tenet – transparency.

Transparency

Asking for consent is no good if you are not clear about what you are asking for consent to do and if you are not in communication about how your practices are changing and shifting over time. In the policies and frameworks it seems like transparency is sort of a given – even the guys over at Ithaka S+R/CAROL have this one. We need transparency in asking for consent around data collection as consent sort of implies “informed consent” and we can’t be informed without transparency. But we also need ongoing transparency of the actual data and of how it is being used.

I found a [blog post from Clint Lalonde published after the 2016 EDUCAUSE Annual](#) that pretty much aligns with how I feel about it:

“Students should have exactly the same view of their data within our systems that their faculty and institution has. Students have the right to know what data is being collected about them, why it is being collected about them, how that data will be used, what decisions are being made using that data, and how that black box that is analyzing them works. The algorithms need to be transparent to them as well. In short, we need to be developing ways to empower and educate our students into taking control of their own data and understanding how their data is being used for (and against) them. And if you can’t articulate the “for” part, then perhaps you shouldn’t be collecting the data.”

What do we owe students when we collect their data? We owe them a clear explanation of what we are doing with it.

But I actually think that Clint takes things a bit further than transparency at the end of that quote and it is there that I would like to break off a bit of nuance between transparency and learning for my third tenet.

Learning

Providing information is not providing understanding and while I can concede that in consumer technologies providing information for informed consent is enough, I think that we have an obligation to go further in education and especially in higher education. We have an obligation because these are students and they have come to us to learn. While they will learn from “content” they will learn a lot more from the experience of the life that they lead while they are with us. If that life is spent conforming and complying to data collection practices that they don’t understand and never comprehend the benefit of then, at best, they will graduate thinking all data collection is normal and they will be vulnerable to data collection practices from bad actors.

Of course this means that we ourselves need to better understand the data that we are collecting. It means that we need to know what is being collected and how it can be used ourselves before we start putting students through experiences where this is happening inside of a black box.

Inside of institutions we need to know what our vendors are doing. We need to create and articulate clear expectations about how we view the responsibilities of vendors around privacy and security. We need to vet their privacy and security policies and continue to check on them over time to see if any of those policies have changed. We need to build a culture of working with reputable companies. Then, we need to build that into the curriculum through increased digital, data, and web literacy expectations.

What do we owe students when we collect their data? We owe them an understanding, an education, about what their data are; what they mean; and what can be done with them.

Collectively, as teachers, librarians, instructional designers, administrators, product developers, institutions, etc. it seems that we will always have a leg up on this though – we will always be in a position of power over students. And so my final tenet has to do with the value of the outcome of data collection.

Value

Finally, if we are collecting student data I think that we should be doing it for reasons where we believe that the benefits to the student outweigh the potential costs to the student. This means putting the student first in the equation of what, when, why and how of student data collection.

I also need to be clear that I'm not talking about a license to forgo consent, transparency, and learning because it is believed that the best interest of the student are intended. This is not an invitation to become paternalistic or to do whatever we want in the name of value.

My point being that the stakes are too high to be collecting student data for the heck of it, or because the system just does that and we are too busy to read the terms of service, or because someone is just wondering what we could do with it. If we have data we should be using the data to benefit students. If we are not using it we should have parameters around storage and yes even eventual deletion.

Collecting student data makes it possible to steal or exploit those data; while we can take precautions and implement security measures no data are as secure as data that were never collected in the first place and, to a lesser extent, data that were deleted. If we are going to collect student data then we have to do something of value with it. Having piles of data stored on systems that no one is doing anything with is wasteful and dangerous. If there is not a clear value in collecting data from students then it should not be collected. If student data has been collected and is not serving any purpose that is valuable to students and no one can envision a clear reason why it will hold value in the future then maybe we should discuss deleting it.

Amy Collier speaks to how data collection can particularly impact vulnerable students in [Digital Sanctuary: Protection and Refuge on the Web?](#) (at the end of which she presents seven strategies that you should also read – no really, go read them right now – I'll wait). Collier starts with a quote from Mike Caulfield's [Can Higher Education Save the Web?](#)

"Caulfield noted: 'As the financial model of the web formed around the twin pillars of advertising and monetization of personal data, things went awry.' This has created an environment that puts students at risk with every click, every login. It disproportionately affects the most vulnerable students: undocumented students, students of color, LGBTQ+ students, and students who live in or on the edges of poverty. These students are prime targets for digital redlining: the misuse of data to exclude or exploit groups of people based on specific characteristics in their data."

What do we owe students when we collect their data? We owe them an acknowledgement and explanation that we are doing something that will bring value to them with those data.

Summation – Trust

Policy is great but I think taboo is stronger.

I can't get that power difference out of my head. I mean it is like the whole business model of education – knowledge is power and we have more knowledge than you but if you come to us we can teach you. There is this trust to it; this assumption of care. We will teach you – not, we will take advantage of you. And to offer that with one hand and exploit or make vulnerable with the other – yeah...

I've been working in educational technology for fifteen years and when I first started there was very little that I heard about ethics. Security, sure – privacy... that was a thing of the past, right? It seems that we are starting to see some repercussions now that are making us pause and I'm hearing more and more about these things.

Still, I see these conversations happening in pockets and while I'm seeing lots of new faces there are ones that are consistently absent. I wonder about new hires just entering the field, especially those in schools with little funding, and

what kind of exposure they are given to thinking about these implications. I wonder if a question like “what do we owe students when we collect their data?” ever even comes up for some of them.

There is a whole myriad of issues that are now coming to light around surveillance and data extraction. What is happening to trust in our communities and institutions as we try to figure all of this out?

Perhaps more than anything, what we owe students when we collect their data is a relationship deserving of trust.

Don't forget

So, don't forget, [the #DigCiz call is open](#) for you to respond how you see fit. Share your creation/contribution on the #DigCiz tag on twitter or in the comments on the #DigCiz post.

We go live Friday, November 2nd at 10 AM Eastern Time with a twitter chat and a video call into the session. Please join us!

~~~

*Thanks go out to Chris Gilliard, Doug Levin, Michael Berman, and George Station, all of whom offered feedback on various drafts of this post.*



This content is provided to you freely by EdTech Books.

Access it online or download it at [https://edtechbooks.org/wild/owing\\_students](https://edtechbooks.org/wild/owing_students).

