Within educational technology, tech companies can acquire data via multiple routes. The most direct way is via a direct signup: a teacher creates an account to use a service, and the teacher is the only person using the service. BetterLesson is an example of a site like this - a teacher creates an account, and only teacher data gets collected. The site is primarily teacher focused.

A second model includes sites that get their initial data from teacher or school signup, but then as part of the service offered in the site, they acquire student information. Basic gradebook applications and some simple student information systems work like this - the teacher or school signs up for the application, and in the process of using it they enter student names, grades, notes, parent information, and other details - the actual data added will vary based on the needs of the application, but information is shared about people without their direct involvement or consent. Online IEP programs also fit this description. While the teacher or school provisions the account, the vendor ends up getting additional data through the teacher's use of the account. In this model, students and parents do not use the service directly, but data about them is collected and stored in the service as teachers use the service.

A third model involves a teacher, school, or district creating an account on the service, and either creating accounts for their students/parents or using an invitation process. In this version, teachers, students, and potentially parents sign up for and interact with the service. The data trail here involves information about all participants that includes personally identifiable information, location data (via IP addresses and/or phone GPS location), and behavioral and interaction data pulled from time spent using the service. Examples of services like this include many EdTech products out there, including Edmodo, Remind, ClassDojo, Schoology, most digital text offerings from traditional publishers like Pearson and McGraw Hill, learning programs like Agilix Buzz, and app ecosystems like Amplify tablets, iPads, and Chromebooks.

A fourth model that collects both learner and parent data includes apps for kids, marketed either to parents or to children. Examples of applications like this include most educational apps sold in the Apple and Google app stores, and some online learning sites. Because these apps are built to be used outside of schools, the data collected by them is not considered an educational record under FERPA, and is therefore covered by the privacy policies and terms of service put in place by the app vendor. If the app is primarily intended for children under 13, parental signoff (and therefore parental data) is often required for use.

The fifth model includes highly structured datastores to collect longitudinal data, and glue services that integrate multiple external services. These applications can support both storage of and analysis on data collected in a variety of applications. A very incomplete list of examples here include Knewton, Infinite Campus, eScholar, Schoolnet, Learnsprout, or Clever.
Context of EdTech Data Collection

The context around educational data is arguably different than data collected in consumer technology. In both K12 and Higher Ed, schools can sign up for services that students use directly, and in many of these cases student data is uploaded before students or parents are consulted. For example, if a teacher signs up for Remind, parents aren't asked if their contact information gets shared as part of an "invite" feature. While many consumer tech apps include invite features, EdTech apps are used within a different context. When a student or a parent sees an app or an invite coming from a school or a teacher, there is a level of implied trust. Increasingly, the implicit trust that students and parents give schools and districts appears to be unearned.

Data Privacy Plays Out Over Time

Unless data collected by an app is deleted or destroyed - and this includes data in backups and in systems that provide redundancy - we need to start thinking of data trails as timeless. This means that a data trail can be transferred from one entity to another if the conditions allow these transfers to occur. In technology, the Terms and Conditions and Privacy Policies are where we can see the conditions in which our data trails are preserved.

While Privacy Policies and Terms of Service should be read in full (free login required for access), for the purposes of this post we are going to focus on two specific sections that can be used to gut the terms in any policy: how policies can be changed, and how data is treated in case of sale, merger, or bankruptcy.

Changes to Terms

Over time, a site's policies should change. Unfortunately, many sites specify that terms can be changed at any point, with no notice to end users, and with no explicit signoff from end users. Many sites state that visiting a site or logging in to a site means that a user accepts the updated terms - so, the simple act of reading updated terms gets interpreted as "acceptance" of the terms. Even on sites that have better notification policies (and at this point, the "best" policies generally include an email and a banner on the top of the site), users often have no recourse (aside from stopping their use of the site) if they don't like the updated policies. Additionally, many sites do not allow users to delete their data from a site, so even if they stop using a site, their data is still stuck in the site.

The ShareMyLesson Privacy Policy uses pretty typical language:
To summarize: most sites reserve the right to change their terms whenever they want, with minimal notification, and no option to remove data. In the case of a site where a learner has been added to a site by a school or district, learners have even less recourse.

**Data Transfer During Sale, Merger, or Bankruptcy**

While a sale, merger, or a bankruptcy are three very different events, most privacy policies treat them identically: if it happens, user information is an asset that gets sold.

**Edmodo’s privacy policy**, shown in the screenshot below, is pretty standard for EdTech terms:

> 5. Business Transfers
> If Edmodo, or some all of its assets were acquired or otherwise transferred, or in the unlikely event that Edmodo goes out of business or enters bankruptcy, user information may be transferred to or acquired by a third party.

For a start, the weak terms used throughout EdTech could be improved by the following four changes:

- Users opt in to changed terms, and/or export data as part of an account cancellation process;
- User opt in when data is transferred to a new owner;
- User account cancellation/data deletion as a regular feature available to users of the app;
- In case of a bankruptcy or going out of business, user data gets destroyed and is not treated as an asset.

These four changes would ensure that user awareness and buy in is included as a factor when terms are changed. There are additional ways that vendor practices could be improved, but starting with these four would be a solid beginning.

**Who Cares Where Data Ends Up? It’s Only School!**

**Data brokers care.**

They have created lists of victims of sexual assault, and lists of people with sexually transmitted diseases. Lists of people who have Alzheimer’s, dementia and AIDS. Lists of the impotent and the depressed. There are lists of “impulse buyers.” Lists of suckers: gullible consumers who have shown that they are susceptible to “vulnerability-based marketing.” And lists of those deemed commercially undesirable because they live in or near trailer parks or nursing homes. Not to mention lists of people who have been accused of wrongdoing, even if they were not charged or convicted.

In the housing market, we have examples where information from data brokers was used to discriminate based on race. If you’re looking for a relatively benign example (if you can get through the marketing speak) of how data from brokers can be mashed up to create profiles, spend some time on this [zip code profiler put out by ESRI](#). It’s worth noting that the profiles were created based on aggregated data of individuals, so that the summaries here are the result of millions of data profiles on
individuals. The **description of how the site was put together** provides a superficial glimpse of how data on individuals from multiple sources can be combined to tell a story.

Now, imagine the increased accuracy that could be added to personal profiles if they are fleshed out with datasets that contain personal information, starting with habits formed in elementary school.

The life of a data trail matters.

**Conclusion**

When a student is signed up for a service by a school or teacher, data is collected from people who have no say in forming the relationship or shaping the terms of the deal. In some cases, **this involves student work being sold without student knowledge**. The fact that EdTech companies treat student data (which really is a track record of learning, personal interest, and growth) as an asset to be bought or sold is on very shaky ground, both pedagogically and ethically. Given that a learning record is also - to an extent - a snapshot of behavior, and that behavioral info is gold for marketers, it raises the real question: why should edu records ever have the possibility of ending up outside an educational context? Treating records as a financial asset that can be acquired in a merger/bankruptcy ensures that some records end up being used outside an edu context. The combination of the "fail faster" mantra of **VC funded tech**, combined with ongoing deals - **over 8 billion dollars worth in 2014 alone** - ensure that student data is getting sold and used outside an educational context.

When a technology company reserves the right to sell user data in case of bankruptcy, they are hedging their bets. By claiming a stake in user data - instead of getting firmly behind their product or service - they are telling us that they do not have complete confidence in their product. And when a company tells us that: we should listen, and return the favor. If a company doesn't have enough faith in their product to leave user data off the table as an asset, we should match their level of faith - and not use their product.