

My name is Adam Mosseri. I've been a product designer at Facebook for about two years now, and today I'd like to talk about how we at Facebook use data to inform certain types of decisions, but how we also are very skeptical of being overly data driven. This is all really about decision making and what informs our decisions, so I'd like to start quickly by talking about who makes the decisions.

It's important to understand that, at Facebook, we believe in particularly small teams. Most projects are about six or seven total. We believe in small teams because we believe they are more efficient, and speed is something that's incredibly important to us. It's also important to note that decisions are made by those teams.

I'm also a manager. I manage about nine product designers at Facebook. I don't approve any of their work. I give them feedback, and participate in feedback systems, so they get feedback from other designers as well. But teams, like a Photos team, for instance, will make a decision about the Photos product, pending only Mark Zuckerberg our CEO's approval, so it's a pretty flat decision-making structure. So I'm gonna walk you quickly through our team structure.

There's always a designer. This is Francis. How many of you guys are designers here? I'm sort of curious. Okay, so a bunch, nice. Product designers at Facebook are responsible for visual design, for interaction design, also for what we call product design, which is essentially product strategy, and we even do some front-end implementation as well.

There's almost always also a researcher. This is Gabe. Gabe loves Post-Its, so I wanted to show this. Do we have a lot of researchers? Not a lot of researchers. Interesting. A couple over there.

This wasn't the case two years ago when I started at Facebook. Researchers were involved in some projects, only the biggest, but not a lot of them. But over the past two years, I've seen us sort of grow to accept the importance of both qualitative and quantitative research, so this is now sort of an integral part of our team.

We also have an engineer, usually between one and four. This is Ola. He's actually our most prolific engineer in the entire company, despite the blanket. But you don't have a lot of engineers, I don't think.

And then product managers. This is Blake. Blake is actually a director of product at Facebook. Product managers at Facebook are responsible not just for project managing, not just for making sure things ship on time, that everybody has what they need, but also for the quality of the product. They're sort of like mini-CEOs within their projects usually.

And, again, like I said, just so you know where I'm coming from, I'm a designer first and foremost. My main interest is in ensuring a certain quality of experience. But today

we're gonna talk about how these teams use data, and I'll stop at talking about the ways we do.

And we do use data. We value it. We store an incredible amount of data. We have about 20 people on the data team: 10 engineers, 10 data scientists. We record about four terabytes of data a day. We invested a lot in the technology to store and query all this data. We have, I believe, about ten petabytes' worth of storage, which is an incredible amount. And we believe it's important, but we use it in sort of certain particular ways. And the first way I'd like to talk about today is how we use it to sort of optimize usually workflows or interactions.

Data helps us understand how users use the product, which then in turn helps us understand how to optimize the product. And the most tangible and recent example I could think of for today was photo uploading. I've been working on photos for the past few months, and we recently, about two months ago, replaced our photo uploader. To give you a sense of scale, about – I believe it's – over 200 million photos are uploaded a day, and a few weeks ago we hit 50 billion photos in the system. That's a ton of photos.

But we thought we could do better; we thought there were problems. Actually, this is pretty interesting. How many of you guys use Facebook? Nice. How many of you guys have struggled uploading photos to Facebook? Yeah, that's about a third of you. So it's pretty bad considering we're, I believe, the largest photo site on the Web.

So we started with a hypothesis, as we usually do. The way we use data can generally be divided into two areas. There's hypothesis generation. And that usually includes sort of exploratory data analysis: we believe it's difficult to upload photos. And then there's hypothesis evaluation: iteration, testing, and that sort of thing.

So the hypothesis was quite simple in this case. It was that users were having trouble uploading photos. We knew this anecdotally, from our own experiences, but also because, as you can imagine, any time a friend, a relative, a loved one has trouble uploading photos, they call us personally.

So I'll walk you quickly through the current upload flow on the site. You start on the Photos dashboard. This is what we call the dashboard. It's on the homepage. You can also get here through the composer or the profile, and you click on Upload Photos on the top right of the page. You then get a form that asks you for some information about the album. You fill out some information; you say where it is or what it was about. You describe it.

And then you get to this page, which is sort of the start of the problem. This page has too many actions. First you select photos, and then you upload 'em. So you click Select Photos, then you get an OS dialogue, or an operating system dialogue, that allows you to select files. You select them. You hit okay. And then it tells you you've selected six. You can change your selection. Then you click Upload, and hopefully you wait patiently as it compresses and uploads your photos to Facebook. So this is a lot of steps.

Putting this presentation together was sort of – felt sort of like airing our dirty laundry. So a few months ago we decided to do a waterfall analysis of the photo-uploading experience. A waterfall analysis is simply taking a look at each step in a flow and seeing what happens.

So of users per session who try to upload photos now, only 87 percent reach what we call the ready state – that means that page you saw where it says Select Photos, Upload Photos – and everything's working. We lose some people to – some people decide not to upload photos 'cause the page takes too long to load. Some people don't have the most recent version of Flash, and we're currently using Flash for photo uploading. So we lose a bunch of people right off the bat. This is actually pretty bad.

Only 57 percent of users actually select photos. In this case, selecting photos means not only clicking Select Photos but also finding some files and successfully selecting them. Fifty-two percent actually upload photos, so that's click the Upload button, 'cause you can change your selection. And then 48 percent are actually successful. We lose 4 percent to poor load times, bugs, etc. It's pretty bad, but it's actually significantly better than where we were.

If you look over the past two months, you can see the photo success rate has increased from 34 percent to the mid-40s. Now, this was a new Flash uploader. I'll talk about the old Flash – I mean a new photo uploader, and I'll talk about the old photo uploader in a little bit. But we're continuously iterating on it, removing bugs, removing pain points, removing steps, etc. And we watch this – this is sort of data driven. This is one of the types of products that are data driven.

But I wanted to dive into one specific change we made. We found that 85 percent of users, when we first launched this, were selecting only one photo for an album, which is clearly not ideal for us or for them. And we wanted to figure out why, so we took a look at the UI that users used to select photos, and they use this.

This is called an operating system file selector. We don't actually have control over this interface. This is the Mac OS version; there's a Windows version as well. But it's very difficult here to select multiple files. You have to click on one and hold Option or Shift and then click on another, and this proved to be very difficult for the vast majority of our users.

So we did what we don't like to do: we added another step. After you click Select Photos, we gave a little – we showed you a little tip that said "You can select multiple photos; this is how you do it." There was a little bit of friction, but we believed it was important, 'cause clearly a lot of people were struggling.

This resulted in a drop on the number of people who were uploading only one photo, from 85 percent to 40 percent, which was huge. We also only show you this dialogue until you successfully select two, and then we never show it to you again. This meant

that photos – per attempt. This 90 percent success, this is per attempt, increased from 3 to 11, which is a big win for us.

This graph is actually interesting. This is photos per upload attempt, and you can see that – this is over two months. You can see there's about eight spikes. Does anybody have any idea why there's spikes? What was that, weekends? Sundays. People upload photos of their interesting weekends on Sundays. We see really interesting patterns in the data all the time. It's about 150 percent of the average, every Sunday. So this was an example of how we use data to sort of optimize a workflow. We're very comfortable doing this.

Another type of way we use data which is significantly different is to sanity check decisions we make for non-data reasons. So it's a little bit complicated, but we do things for all sorts of reasons, and we have key metrics that are very important to us. And so we generally sanity check our changes with it, by running A/B tests. At our size, we can launch a product to a small percentage of users, like half a percent, and get statistically relevant data very quickly, which is really just sort of convenient.

So I wanna talk specifically about what we call the composer. The composer is what we call internally the "What's on your mind?" input field that you see at the top of your homepage, right above the News Feed. I'll try to show it to you here. This is the way that most – not most users. This is the way that a lot of users update their status. But once you click on it, you get a few other options: add a photo, post a link – what was the other one? Oh, add a video, etc.

But recently, we've started to roll out our Questions product. I think we've finished rolling out in the U.S. Do you guys all have the Questions product? Yes, no, maybe? Well, anyway, for Questions, it was important for us to surface a really easy way for you to just ask a question on the homepage, and this "What's on your mind?" input field wasn't gonna cut it. It didn't really afford us that sort of flexibility.

So we wanted to test moving the composer to more of a selection-based model, so we tested a couple options. We tested a version that was just four links across the top, so we could add an "ask a question" link. We were worried, though, that this would decrease the number of status updates, because the relative prominence of statuses is less here. And it did; it decreased status updates by about 1 percent.

We also tested a version here where we tried to incentivize users to update their status by showing them their most recent status. The idea would be that stale content would motivate you to upload – to update your status. And it worked marginally. It was about a .5 percent increase in status updates. That wasn't actually statistically significant, but we saw it in the data.

And then we tested a big option. We always test the big option. This was links and an input field. This resulted in a 2 percent increase in status updates and a 2 percent increase in photo uploads. But the real – the truth at the end of the day and what we were actually hoping for – and we actually ran around eight different versions, not just three that I

showed you here – was that none of these really significantly moved our key metrics, which was what we really wanted, actually. We wanted to make sure that we could go with the UI that we thought was the best, and so we went with the simple one, 'cause the homepage is very complicated; I'd argue it's significantly too complicated. It's very important that News Feed is high up on the page, so we went with the simplest, lightest version. This is actually a pretty recent example. This is – I'm not even sure if it's fully rolled out.

And the third way I wanna talk about how we do use data, and comfortably, is to evaluate retroactively projects. This usually is for small projects. I'm gonna talk a bit about the deactivation page. The deactivation page is the page you get when you decide to leave Facebook, which we find sad. And Lee Byron, a designer at Facebook, actually spearheaded this project. It was his idea. He designed it; he built it; he ran the tests; and he shipped it.

And the idea was that the current version of the deactivation page – this was in mid-2008 – was just a form. We wanted to know why you were leaving, but we didn't ask you to stay. We didn't give you a reason to stay. So he thought about being somewhat emotionally manipulative, and he did this.

*[Audience laughter]*

As you can see, it says – there's a picture of my friend Aaron, and it says, "Aaron will miss you," and then Kevin, and "Kevin will also miss you"; "Send Wayne a message." To just hit that emotional chord, to give you a reason to stay, to make you feel guilty about leaving. And it was wildly successful. It reduced deactivations by 7 percent. Seven percent at this point is millions and millions of users still on Facebook, 'cause this was a year and a half ago and when we were about 70 million users.

And so this project was entirely data driven, but it was in sort of a mid-size project. This wasn't a homepage redesign or a new version of Photos or a new version of Groups, and so we're comfortable with it in these areas.

But I think it's fair to say that, at Facebook, in product – we call product "product management and product design" – that there's a healthy skepticism of being overly data driven, maybe even too much so. But I thought a lot about this for this presentation, and I tried to articulate why: why we're really so skeptical of overusing data.

And the most straightforward reason I could think of was that it's very difficult for a set of metrics to fully represent what you value. There are a lot of factors that go into making any sort of product decision, as I'm sure you guys all know.

Quantitative data is one. We use it, as I've showed you over the past three examples. Qualitative data is another. Our researchers run qualitative tests all the time. We have a usability lab; an eye tracker, which is pretty amazing, actually, if you ever get the chance to use one; etc.

Strategic interests are another factor we use in making decisions, as I talked about with the Questions product. User interests are another: what people complain about, what can people ask for.

Network interests, which are actually significantly different. Competition clearly factors into our decision making. Regulatory bodies at this point. At our scale, we have to deal with privacy advocacy groups. The European Union had a lot to say about Questions – oh, no, sorry, about Places. So we deal with them, and we looped 'em in on decisions.

And business interests. This is actually, on purpose, small because explicitly we value revenue generation right now less than growth and engagement, growth being defined as how many users come onto the site, engagement being defined as how often users use the site. So these are all important factors that we use in making our decisions.

And so this is sort of implicitly understood at Facebook, and every once in a while we experiment with something that's a little bit maybe too data driven, I'd say. So I'd like to talk about a pretty recent example, which we called internally the engagement team.

We've gotten away with a lot of designing for ourselves over the past six years. And recently we've decided to really invest in trying to understand why users use the product as much as they do and how to sort of motivate them or to persuade them to use it more. So we created a team we called the engagement team, which was tasked with understanding engagement and increasing it significantly, but also with quantifying it, which was sort of the dangerous piece.

And our first attempt at quantifying engagement was RAW, reads and writes. So what do I mean by that? We talk a lot about the social graph internally and externally. The social graph is the digital representation on Facebook of real-world entities. Your relationship with a friend, we call a friendship. You going to a party, we call a – you are RSVPing to an event. Your football team is a group. And we believe that – or we talk about that the social graph is just objects and connections between objects within the system that represent real-life entities. And we talked about what reads and writes are. Reads are creations of either objects or connections between objects – sorry, writes. And what reads are, are what they sound like: reads of that information.

And so we just decided to treat all writes equal and all reads equal, and start to try to optimize for that. We did this over the past few months, and we ended up with products like comment liking. Comment liking is what it sounds like. We produced – we put the product that allowed you to quickly and easily like a comment. Here, Saleo said, "Fine," and for some reason I liked it. This, actually, biometric was wildly successful. It produced an 11 percent, I believe, increase in likes throughout the entire system. This is really good for our metric goals.

But there was sort of a feeling within the team or within the company that this really might not be the best thing to optimize for. We sort of got what we asked for. This type

of write, the fact that you like the comment, is explicitly or obviously less valuable than you telling us that you had a baby or that you switched jobs or that you moved companies. So clearly, all writes weren't created equal, and we started to struggle with this.

We also started to struggle with what we found ourselves optimizing for. We found that 85 percent of the content in the system was generated by users who log in more than 25 days a month. That's a lot of users – or, sorry, a lot of days. That's actually around 20 percent of users in the system.

But we realized that if we start optimizing for this percentage, for these heavy users, for these power users, that'll be invariably at the expense of our more casual users. Casual users are important to us too; some of 'em become heavy users. There's no reason why you can't log into Facebook once a week instead of 25 days a month. And we realized that we were over-optimizing for a small user segment at the expense of the rest.

So what we're doing – and I actually have a brainstorm today back at the office – is trying to reevaluate this metric. We created this team; we're committed to understanding engagement, but clearly the metric-driven approach isn't working for us, and specifically that metric is poor.

So another reason why – to move on – why we are skeptical of data-driven design is that we find that overreacting to data often leads to what we call micro-optimizations. A micro-optimization is when one interest over-optimizes for itself at the expense of another, and this is a very difficult thing for us as we scale.

As we scale, a division of labor becomes invariably sort of more intense, and you have different people representing different interests. We have a Photos team; we have a growth team; we have an engagement team; we have a News Feed team, etc. And all of these teams optimize in good faith for their own interests. But sometimes these interests can be sort of opposing or distracting from each other, and sometimes you can get lost in the specifics of a decision and sort of miss what we think of as the big picture.

So I'm gonna give you an example of something that launched that was core to our product for a long time, that I think was a poor decision. And it's what I call the application menu, or we called internally the application menu.

This was the site in early 2008, and the navigation was on the left. That's how you navigated to what we call applications. Applications are photos, groups, events, notes, but also platform applications – PackRat, Mafia Wars, FrontierVille, etc.

And we redesigned the whole site – right when I started at Facebook, actually – with the idea that we wanted to move the navigation to the top, explode the frame, and allow content on the site to sort of thrive. The application menu moved from being a list on the left to a dropdown at the top of the page. And this resulted in a significant decrease in traffic to applications, and this was a big problem for developers.

So we were committed to this though, so we started to explore how can we increase the prominence. So we moved it to the bottom of the page, as you can see it here. Not particularly prominent, but it increased traffic significantly. We even tried the "big blue button" approach. This resulted in a 5x increase in traffic, but we all hated it, so we actually didn't launch this, so maybe we didn't do as poorly as we could've done.

But what we were doing here is we were optimizing for a local maximum. Within this framework, there was only so much traffic we could funnel to applications. And what we needed was a structural change. Our premise was sort of off. Our interests were basically leading us down the wrong path, and we didn't realize it, and we launched this.

This existed on the site for a year. But it did spawn a few of the conversations about navigation and how navigation should operate and persistence and about platform navigation versus internal navigation. And it ended up resulting in a team that designed this, which is the current crumb – what we call the crumb – the current navigation of the site, which was sort of a half-step backwards, to a left nav with a wider frame and a bit of a more flexible system.

So we were optimizing for something locally, and we needed to be somewhat disruptive to sort of get out of it. And this resulted in an increase in application traffic, but this was about a year later.

Another example of a local optimization – or a local maximum where we got lost chasing a local maximum is the old photo uploader, the one that existed temporarily, briefly before the one I showed you, and it looked like this. And this photo uploader was awesome.

Basically, within the context of Facebook, we allowed you to browse your file system, see thumbnails of photos, and select what you like. You could select photos from different folders. You could click Upload, and you can continue to navigate the site while it was uploading. It was a really great experience.

But the problem was, to enable this, to give us access to the file system, we had to build a browser plug-in, a downloadable – something you had to download and execute. In Safari it looked like this. You got a very scary warning that said, "An applet from Facebook is requesting access to your computer." It was actually much worse in IE. In Internet Explorer, you got an ActiveX control. If any of you have seen that, it's a 11-pixel, tiny yellow thing across the top of the page you're supposed to find. In certain browsers, you had to download and install something. A lot of users actually don't understand the difference.

And we did a waterfall analysis, and we found that, out of the roughly 1.2 million people a day that we asked to install the uploader, only 37 percent even tried to. That means that 63 percent said, "Piss off," like "This sounds like – this is either spam," or "I don't trust you," or "I don't know what this is," or "I don't understand," or "I got lost somewhere in

the process," and they didn't even try. And only 23 percent were actually successful. This is abysmal.

Twenty-three percent is a little bit misleading. Out of the 1.2 million who tried to install, there was another 600,000 that already had it installed. So the overall success rate was around 40, 45 percent, which is about where we've gotten recently, if you remember this slide from earlier.

But what we needed to do was start over, really. We had reached our local maximum, which was around 40 percent, 45 percent. We had been optimizing for months. We had made substantial gains, but we had plateaued, and what we needed is to move to a completely new uploader. And we've reached our previous performance, and we're still on an upward trajectory 'cause it's still a new project.

But this is somewhat disruptive, which leads nicely I think into my last point, which is why – or my last reason why we're pretty wary of being data driven, which is that we really believe – and this is a little bit controversial – that real innovation invariably involves disruption. And disruption is usually – involves a dip in metrics.

And this is core to our sort of culture. It's core to our product beliefs. It's one of the main reasons why I joined Facebook. I joined in 2008, but I started to try to join in 2007 because I saw News Feed, and News Feed was an example of a project that was executed in lieu of, in spite of, or just oblivious to data.

Does anybody remember this version of the site? This is pretty awesome.

*[Audience laughter]*

This is Ezra. Ezra was actually employee No. 6 at Facebook. This is what it looked like. And the way people used this site back then is they navigated from profile to profile, essentially trolling for interesting information. And we had – we knew that this wasn't ideal. It was actually good for the standard metrics at the time for engagement, i.e., page views. People loaded a lot of pages in search of something interesting.

But we thought we could do better. We thought we could surface what was interesting to you right there on the homepage, create a custom social newspaper for you, and we called it News Feed internally but – yeah, both internally and externally. And it looked like this when we first launched it.

And this, if you remember, had a massive backlash. Users hated us. We got a ton of bad press. This is one of my favorite quotes: "Generation Facebook is taking action – against Facebook" – *Time* magazine. TechCrunch hit us. There has been an overwhelmingly negative public response to Facebook's launch of two new products. The two new products were News Feed and Mini-Feed, which is your feed on your profile.

But we stuck to it. We believed in it. We added some privacy settings. Mark wrote a

---

letter to the entire user base about – explaining what we were doing and why. And eventually it ended up becoming the primary driver of traffic and engagement on the site. It is probably our greatest success story.

But I think it's only fair, if I'm talking about bold moves in spite of data that were successful, is to acknowledge some of our failures.

Beacon was a project where the basic idea was that Facebook shouldn't just be about what happens on Facebook. When you go to Facebook, you should see what your friends are doing all over the world. And the way that manifested was that we would allow third-party sites – that is, sites that are not Facebook – to funnel your activity back to Facebook. So if you created a review on Yelp, it would come back to Facebook. If you wrote a review on Rotten Tomatoes, it would come back to Facebook. And it looked somewhat like this. Basically, in your News Feed at the top of your homepage, we threw those stories. We created stories about things you were doing off of Facebook.

But we did this – it was opt-out. We did this implicitly. And the classic terrible story was when Christmas was coming up, and you bought your girlfriend a nice bag on Amazon. And then she logged into Facebook, and she saw that you bought a nice bag on Amazon, and either you had spoiled Christmas by letting her know what she got beforehand or she found out that you were buying a bag for somebody else, and this was just terrible. It blew up, and we tried to stick to our guns, and eventually we had to sort of roll back and make it opt-in – yeah, make it opt-in.

And it's a pain point. It's actually difficult for us to talk about. But I wanna acknowledge it. It's real. Along with trying to innovate and trying to make bold moves comes – you run the risk of failure, and you have to just understand failure, acknowledge it, and move on.

A couple other projects that were made sort of in spite of data or in lieu of data were homepage redesigns. We've done this a number of times. I've actually worked on the last two, so if you hate them, you can e-mail me later.

This was the homepage in 2009. This was March of 2009 we launched this. This had nothing to do with data. The idea here was that we wanted to make the News Feed entirely about what your friends were saying. So instead of algorithmically deciding what we thought was interesting, we showed you everything your friends were saying: the photos they were posting, the status updates they were writing, etc.

It being all-inclusive also meant that we sort of were focusing on recency and voice. We didn't say what your friends were doing, so if your friend RSVP'd to an event, we didn't tell you; if your friend joined a group, we didn't tell you. But if your friend posted a status, we told you. So the idea was the focus on voice, the focus on recency – 'cause it updated in real time 'cause it was all-inclusive – and to focus on simplicity, determinism. You knew how News Feed worked.

---

It did increase comments significantly, and it did tank wall posts significantly, but we were gonna go forward with this no matter what, 'cause we believed in it for strategic reasons.

Another one – homepage redesign which we talked about briefly was this one, which was about improving and simplifying the navigation on the site. Instead of the way – the way before this to get to photos was you had to find a 16-by-16-pixel icon in the bottom left of the screen in a gray bar, which was very difficult for a lot of users, and we knew this was wrong. This resulted in an increase in traffic to applications, but we were gonna move forward with this.

But I do wanna take a moment to acknowledge that when you make these big changes, when you don't take baby steps – and I actually do believe in baby steps – there's a cost. There's a very real cost.

This is a real group on Facebook called "I automatically hate the new Facebook homepage."

*[Audience laughter]*

This is awesome. This is, like, my favorite group. There's over 23,000 members, and some of the hate groups had millions of members. Sometimes it was – you would have a hate group about one homepage redesign that would then exist until another homepage redesign, and people would join that, at which point you had people who were hating on both, and so you wanted to go to one or the other. It gets really complicated. But it's real, and we need to understand it, and we need to be somewhat sympathetic.

The way I think about it is that the average user who logs into Facebook today will spend about 46 minutes on Facebook. That's crazy. That's a lot of time. Now, if you spent 45 minutes every night at your desk organizing your photos, writing letters to your friends, doing your thing, your social sort of activity, and then I came by with no provocation, with no heads-up, and I just rearranged your desk for you, you're gonna be pissed. It's gonna happen. I'm messing with your desk. That's real. That's a real sense of entitlement, and it's – you can argue with it all you like, but the truth is we need to understand that.

So moving forward, we need to understand how to message our motivations behind our major decisions better by explaining value add to users better. There's clearly a lot of room for improvement. But I do believe that you'll see us continue to make big changes that you'll be like, "How is this good for data?" or "How is this good for anybody?" but there's a reason behind it. Usually it's either because we believe that this is where the market is going, or it enables a product that's gonna come later, or we're worried about being stagnant and we wanna continually innovate.

For us, the greatest risk is really taking no risk at all. That's why I'm at the company. That is, like, pervasive from Mark, all the way down through to all of engineering, all of product, etc. And I believe in this. I really, really do.

That doesn't mean we can't do better. That doesn't mean I'm not pushing for us to do better. But it means that at the end of the day we make decisions based on common sense, on interests, on strategic interests. And we use data. We acknowledge it's important, but it's really just a small piece of the pie.

So that's actually all I have today. Again, my name is Adam Mosseri. My e-mail is Mosseri at Facebook. If you liked this or you thought I'm totally off base, I'm really actually open to feedback. But thanks for your time.

*[End of Audio]*