Now & Next E03 – Wattpad: Leveraging a billion data points a day

Full interview transcript

Leora Kornfeld (LK):	I'm going to start by throwing one of your lines back at you, and hopefully this is okay. You have said, "I have one of the easiest jobs in entertainment because I don't have to guess what's important. I only have to listen to millions of people telling me what's important." But I only kind of believe you because, in general, the entertainment industry is run on something like a 97% failure rate, whether it's music, or books, or TV, or film, because it's really, really hard to figure out what will resonate with an audience. Please, I want you to come clean. Your job is not easy.
Aron Levitz (AL):	I would say that when you're working in a moment of disruption across an entertainment industry, starting a new methodology of finding new IP and developing IP in new ways, that's . . . A harder part is to have an industry who's been based for hundreds of years, as you talk of all those creative industries, off their gut, and helping them learn that audience and data can really make a difference. I'd say that's the hardest part of what we have to do. And the nice thing is, is while it makes my job easier to not necessarily have to understand how good something is if a million people tell me that it's important, my job is to understand why they think it's important. And that can be part of the trick in working with great creatives like directors, producers, etc., to really understand why the audience likes something.
	But, absolutely, the fact that we can use data to better predict success, to your point whether it's a 90% overall fail rate or 80% of shows that can't make year two, that audience is going to make the difference, and it is going to make it easier for us to better see our chance of success. Not guaranteed success, but what an audience can do. There's a lot of things that can go wrong between the written word and the screen, but we really believe that following the audience using data-driven stories, we can have a better chance of success.
LK:	We have a sense of what you're talking about, data-driven insights and built-in audiences, but what does that look like in practice?

AL: Really, if you look at Wattpad as really the world's largest social network for readers and writers, what it means is 65 million people coming every month to re-write and tell their stories. Four million people contributing stories to the platform, 565 million stories across genres. What it really means is finding a way to listen to what people like, what resonates with them, whether on a micro-level—and by that I mean what story may be shot to the top of a list—or macro-levels, being able to look at what subgenres of horror are trending this month.

In practice, it really is looking at the billion points of data that we collect every day in trying to understand and analyze that so we know what audiences are resonating towards. It's not always the biggest, it's great to look at some of our biggest stories, be it a *Chasing Red* or an *After*, that have both seen off-platform success already in publishing, and the cast of *After* has already been announced for the feature. But it's often being able to say, "Okay, this story got to a million reads before any other story," or "This story had six-and-a-half more reading time than any other story in that genre."

Really, the practice looks like a lot of really understanding what audiences want and what audiences are attracted to, not what my gut feel is or my development team's gut feel is, but if *Chasing Red* is one of the most-read stories over the last two years, why did audiences love it? We're able to bring that kind of insight to our producing partners, not just to find *Chasing Red*, but to actually be put in to develop it where we can say, "Look, this paragraph," —because on Wattpad you can comment directly down to the paragraph level— "This paragraph has 5,000 comments, so something is obviously important here, but let's see what the subtext of it is." Or, "This chapter has no comments whatsoever, if we have to cut something to write the script, that seems like a prime candidate for it."

LK: And how do you do that? Without giving away any secret sauce, but what kind of proprietary systems do you use to determine these things?

AL: We definitely have a unique set of data around genre, around storytelling, around audience, that nobody else has. It is something that only we get to look at along with our partners that we're working with. It's not something that's publicly available that you can just go look at Wattpad. Now in Wattpad, if you went on you would see reads, the amount of reads the story has, and that's a great metric and it can show that a lot of

people like something. But we're often looking at metrics that aren't publicly available, things like the amount of reading time spent. You can imagine . . . It's not important that someone opened the book, it's important how long they spent reading it or how many people finished it, how many people re-read it, how many times it was shared, what sections are most shared, what quotes were most shared.

We're really diving into data that nobody else has access to, which does give us that unique view of how stories are growing, how genres are changing, how writers are writing. And that's a very important way we look at data. We can look at what readers are reading and that might be a little more akin to how industries have looked at content to date, which is, "Here's a story, did people like it?" We can flip that whole piece on its head and look at what writers are writing. We have four million people that contribute to the platform every month. We can look at what they're writing.

This doesn't have an audience yet, this is just what creators are interested in creating, and that gives you a totally different view of the data. With the billion points of data every day, we really can cut that in almost an infinite number of ways to learn new insights about the audience, genres, writing and writers themselves.

LK: And you've mentioned a billion points of data every day twice. I've heard you use that number twice, so I'm guessing that's not just a generalization, that's actually what you're working with.

AL: No, that is actually what we're working with. It is a phenomenal amount of new pieces of information we gather every day.

LK: What you're describing actually sounds not unlike Netflix. Every start, every stop, every rewind. How simpler would your system be, would you say?

AL: It's actually quite a different use of data. They're really focused on taking the insights that they could draw from their community, creating something, but they can't figure out if audiences like it until they post it, until they put it up on their platform, until an audience starts streaming it. And then it goes back to a really traditional way to use data and your Nielsen rating: "I made this. Did people like it?"

What we're doing is actually flipping that axiom on its head and saying, "People like it, let's make it." We can see that something

like *White Stag*, which is a phenomenal fantasy story on Wattpad, written by an author who goes by Pandean, that it was getting, I think, seven times more reading time than any other fantasy story on Wattpad. And we read it and it is a beautiful, beautiful universe. And we say, "Okay, there's an audience here who really enjoys this." Well, that's going to take some of the risk out of developing it further.

If we publish it in a book and it got a three-book deal with Macmillan, for example, they know that people are going to come from Wattpad, show up at bookstores and pick that off the shelf. That's a real difference than saying, "Okay, well, people like political drama. And people like this actor, so let's put them together and hope the show turns out well." And then it goes into a purely traditional development process, versus saying, "Look, we're trying to be predictive of what people like, but we have the underlying IP and the storylines, and the characters, and the universe lying right in front of us."

LK: Something I read about the creative industry, I found it a really interesting way to have it expressed, is that the big challenge is actually maximizing your creative risk because nobody knows what's going to hit. If anybody knew what was going to hit there wouldn't be that 90-something percent failure rate we already talked about. So you have to maximize your creative risk while actually managing, if not minimizing, your financial risk.

AL: I think that's what happens when you shoot from the hip, when you have traditional development, which is four people in a room, the same four people with the same backgrounds, and lack of diversity they've had for the past 70 years, 80 years, making these decisions. And saying, "Look, we know what's best." At that point, do you have to maximize creative risk? Maybe because you want to try to stand out in a crowded room, but if you're starting with the audience, with the data that people already like story X or series Y, here's why they like it. Here are the parts they like, here are the parts they don't like. Here's the subtext on that character that while you think this is where everyone falls in love with the male lead, this is actually the point where the audience splits and half of them love him, and half of them hate him.

With that kind of context, it's not about maximizing creative risk, it's making really smart creative decisions. And it's also not about just letting the audience say, "Okay, we all said go left, so you have to go left." It's still dependent on taking something from the written word to screen to have great development execs who have a vision on how something goes from written

word to a visual format. It needs great show-runners, it needs great actors, it needs great directors.

Those creative layers still have to be there, and there's different risks you can take in any of those creative layers. But what the audience is going to let you do is to just remove a creative risk where the development exec's job does not have to be to choose the next best story, it should be to have a view on how to make a story someone's favourite. And I think that's a real difference that we get to bring to the whole development process, not just finding the story.

LK: Right, a lot of people think that there are these things called answers that exist in the data. And I can hear you chuckling. I'm dying to hear what you have to say about this because it's not, you don't get answers. How would you describe what you get from the data?

AL: I think what you get is . . . Here's how I put it, if you've ever sat in a writers' room for long enough, there comes a point where everyone gets stuck. It doesn't matter how many people are in there, how many awards they've won, and how long they've been doing this. You have a point where everyone starts to stare at each other, refilling their coffee, and really trying to break this scene. We can't figure out if the premise is right.

What if there was a way to take a dark room and brighten it up a bit? What if there was a way to say, "Look, we don't know whether to cast person A or B in the show. What if we could ask people who already love this IP, who love this universe, and they could help us?" Data is not the end to be all from the standpoint of the amount of creative layers that have to go in to get something from written work to screen. What the data is going to do is brighten that dark room, it's going to give you a better starting point. It may even give you a better ending point, because you can go back and market to the same people who helped you find the story in the first place to show up and tune in night one, or show up at the box office, or go to a book store.

But at the end of the day, the data is a tool. It becomes part of a development executive's, director's, show-runner's and editor's repertoire that they don't have today. By no means does this negate the necessity for a great screenwriter. We need a script to be generated, we need someone with a creative vision on how to take 300 pages and turn it into 90 for a feature, for example. But the data is there to help you understand what people have loved already and what people will love in the future.

LK: You've given some examples of some really great success stories. I'm curious to hear more about times when data actually didn't steer you in the right direction.

AL: Very good question. I think we've been pretty lucky in how our projects have proceeded early in our studio's lifetime. We really started the studio about two years ago. We launched it at the Banff Media Festival and we've just seen that, by using audience, we have had great successes. We look at something like *Cupid's Match*, which was a story we found that, at the time, probably had about 20 million reads on Wattpad, maybe a little less. It's now close to 40, 45 million reads on the platform. And we created a totally new development methodology for it, we took it actually to a community of 120,000 filmmakers and said, "Pitch us how you'd do a trailer for this."

So now we've taken our data that said this is important and taken it to a massive group of creators, who all pitch differently. We get three that we create trailers from, the audience chooses the production type they love. We take it to distribute on CW Seed, which is the CW channel in the U.S. digital platform. And within a matter of weeks, it becomes the second-most-watched pilot on that platform of all time. Why? Because the audience continues for us to be predictive. It continues to show that if you adapt, the audience is going to go check it out.

Now, it has to be adapted well for them to stay there. But we lucked out because the audience helped us then choose the producer and director who had a vision that they thought resonated with them. And by the way, it resonated with people who had never heard of *Cupid's Match* before, and those people went back and read the story on Wattpad. We continue to show with our successes that the data is predictive. It is about changing an 80% year-one failure rate to 70 or 60%. We've been quite successful to date, I'd say, in really showing that the data does have answers for the creative industry to help them make better decisions.

LK: That's interesting. So far you're seeing that incremental bump of about 10% and, because you're a data guy, you wouldn't just throw a number out if you didn't have science behind it.

AL: I would say that's the goal. The goal is to change, not a 100% success rate. And I think that's the biggest thing to take away there is. I'm never going to tell you that if you start with a story that 10 million people liked, you're going to get a 20-season TV show out of it, because there's a whole lot that goes wrong during the adaptation process that has nothing to do with the

underlying IP. We all know that one big snowstorm can wipe out an audience for a first weekend of a movie and that takes it off half the screens in the country the next day.

But what I can say is that the aim of the studio is just to improve that success rate. It's to improve it 10 or 20%, maybe 30%, maybe one day 40%, but it's not talking about necessarily guaranteeing a 100% success rate if you start with a data-driven story.

LK: Right, and that's one of the misconceptions about data in general. I'm sure you find that when people who aren't necessarily in the industry talk to you, it's what we refer to as, "Oh, you get these things called answers from the data and you have this thing called a guarantee of a success."

AL: Yeah.

LK: And how do you address that? How do you start talking to those people to guide them down the right path?

AL: We've taken a really creative approach to our own process. We have people with great TV, film, publishing backgrounds on the Wattpad Studios team, who help us look at the data and literally translate it. Translate it so that traditional players in entertainment, this group that's running through this moment of disruption, can understand and better imbibe what we have to show them. This is never just about putting ones and zeros up on a screen, this isn't that scene from The Matrix where all the numbers are just flowing down. That's not what it means to use the data. It doesn't mean we're just going to send a producer who says, "I'm really interested in spaceships and aliens sci-fi," a link with a bunch of numbers after it.

We do take a creative approach to our own data. We will look at a story and say, "Look, everyone loves this story, but we know when you turn this from a novel to a TV series, that this character loves inner monologue, you're going to want to add a best friend." We can take a view that, "Look, the data is showing us, and the ups and downs are showing great three acts, we want this to be a film," or "The ups and downs in this one," because all of our stories are serialized, they're written chapter by chapter, "This one really lends itself to a long-running series and here's the arc for season one, here's the arc for season two. And by the way, if you want to do season three, focus on these four characters because the data really shows they're the four characters that should be in your A, B, and C plots."

It isn't just about sending over the ones and zeros, spreadsheets, and just the raw data, it really is helping our partners work through that data, understand how it can be a useful tool, and helping them make use for it. For that matter, often our partners have totally different questions they want to ask the machine, as I often call it, that we haven't thought of, which is why our great partnerships, like with eOne or with NBC Universal, or Bavaria Films in Germany. They're just as important in this process that helps us understand our data even better.

They may ask, "Well, look, we know the networks are saying that, right now, nobody is interested in this type of main character. What does your data say?" And we can go and check it and confirm or deny it. It really is a two-way street.

LK: I'm curious to know how threatened Hollywood is by this scrappy Canadian company? Are you still getting invited to the good parties?

AL: We only go to the best parties, in fact, I'm pretty sure. I think because of the way we work, because we're here to partner with the industry, and because we're open with our data with our partners, it really sets us apart. And I'm sure you've heard this before, the streaming services, if you sell your show to the streaming services, they're not going to tell you how many viewers you have. They're not going to tell you if 100 people viewed it or if 10,000 people viewed it. The best you can hope for is someone goes in an article and mentions your story, or that you happen to turn on your streaming service every night and you still see it in the top left corner. But you get no data from it.

I think why Hollywood isn't scared of us, as you said, is that we're helping them. We're the ones that are going to help them move through this disruption and then really be able to harness data that they have no access to otherwise. Otherwise, they're reduced to using the data they've always used, which is we made a show, we invested 10 million, 50 million, 100 million into it. Was it good or not? And we're going to help them use that data in an earlier stage, in a more predictive way, whether it is to create TV, film, digital series, AR, VR, it doesn't matter. There's a place for a Wattpad story literally everywhere in every language around the world. And the data is going to help those players that don't have access to data make better decisions as they continue to grow their businesses.