Magisto's A.I. Helps Anyone Produce Polished Video (Full Interview)

Magisto wants to do for video what Instagram did for photos—edit and improve them so they can be easily shared. The company was founded in Israel in 2009 by two experts in computer vision who applied artificial intelligence tools to video editing. It's product enables a user to select photos and videos on their smartphone and then have its software automatically turn them into a slick edited product in minutes. While its basic product is free, if a user wants to make more than a few short videos they have to pay for the privilege. Magisto launched in January 2012 at the Consumer Electronics show, won an app competition there, and now has 20 million registered users around the world, up from 3 million a year ago. With 30 employees, the company has offices in Tel Aviv, New York, and San Francisco. Techonomy sat down with Magisto CEO Oren Boiman for a wide-ranging talk about the future of video, the state of social media, and how people want and need to express themselves. Boiman, like so many Israeli entrepreneurs, began in high-tech while in the Israeli military, where he worked in intelligence.

Techonomy: How did Magisto get started?

Boiman: When my daughter was born, it was the first time I actually did real video, not just in a lab. You have a newborn and you want to capture their life experience. I bought a Sony Handycam in 2005 and I was shooting for two months. I thought, "I'm curating the family experience, it's going to look great." Then we wanted to have a party and share our young family's experience. I looked at the videos and they were just horrible. We had to do something. Me and my wife—both engineers—said, "Let's do video editing. It will probably take us a couple an hour or so." We spent two weeks slaving to get to five minutes we were proud enough to show. On the one hand I'm doing my Ph.D. in video analysis. On the other hand I cannot do anything to help with our home videos. But then again, when my mother saw the final video, she cried.

Techonomy: So after all, it worked.

Boiman: But because of that experience I understood that the solution needs to be as simple as pressing a button.

Techonomy: So you took the expertise you developed doing your PhD and began to apply it in this context.

TECHONOMY

Boiman: My PhD is about one mechanism which can address many, many problems in computer vision. But when it comes to video editing, you have to deal with freeform consumer video and photos. You cannot do just face detection or face recognition, because you have a ski mask, or there's a dog here, so it doesn't work. You have to have a general understanding of what is in a video, who are the important characters, what are the important scenes, when do they interact with each other, the speech parts, the audio. You need to create something that moves people. That's what video editing is about. But even one video has so many pixels you need to analyze, so people were doubtful whether we can do it. There was nothing that analyzes videos and photos so deeply at that scale when we started.

Techonomy: So you do really do deep analysis of the imagery?

Boiman: You have to. Without getting to the level of understanding characters, scenes, interactions, actions, speech, you cannot get to storytelling, because this is what stories are built from.

Techonomy: Does that mean for instance you can match faces so you can put them in different scenes or whatever?

Boiman: That's a starting point. That's the lowest level of understanding. Beyond that, you need to understand, "here is the main character, because I see him again and again in all the footage". Without this level of understanding, you're not going to do any good video editing. I'm pretty sure we have one of the most sophisticated platforms for understanding photos and videos.

Techonomy: What's the long term vision for Magisto?

Boiman: There are endless ways to improve our platform. For instance, we introduced different editing styles. So when the user chooses "travel," or "cute," we don't change just the looks. We actually use different algorithms and choose different parts of the video. So if I want it to feel sentimental, we might take dialog between the main characters. But if I'm going to do fast paced music video editing with the same footage, I'll choose completely different parts. It's not about what's important, but what's important to tell the story in a certain way. What we're doing with Magisto is injecting emotion.

Techonomy: Did the incredible success of the Facebook "look back" movie surprise you?

Boiman: Look back was a pretty straightforward slideshow that used the metadata you have on Facebook posts, like likes and comments, to choose the right parts. For most of the footage that we process, there's no metadata. You have to go to the pixels. That's what we do. We're doing look back for every life experience you have.

Techonomy: Have you ever talked to Facebook?

Boiman: We talk with pretty much everyone.

Techonomy: So everything you do is about video.

Boiman: Yes. When it comes to mobile, video is the best form of consumption. Video as a medium is like a thousand times more powerful than photos and text. Video has another layer of emotion, another layer of story. Today's social media is very shallow. I mean, we use smileys to try to put some emotion there, but it's not there. But when people see things in video, they really laugh. With Magisto if you want to make it cute, we'll make it cute. You want to make it sentimental, we'll make it sentimental. You want to make it funny, let's make it funny.

Techonomy: You could do all three with the same set of video?

Boiman: Yes. That's why people are paying for this. People are paying quite a lot to create Magisto movies. It's not like they cannot do it themselves, but we save them a lot of time. But it's mostly about them being moved. People say they cried from Magisto videos.

Techonomy: So talk about the market reception Magisto has had. How are you doing with your "freemium" business model?

Boiman: The premium offering right now is still kind of limited. We help you make longer videos, if you want to upload more photos and more videos on the web. But already we see amazing conversion rates among people who have families, way beyond industry norms. We will offer more and more as part of the premium offering. Now people are paying the $5 bucks or the $18 per year using an emotional justification. It's an impulse buy. Around 50% of our users are teenagers, but this model is not right for them. They will not pay $5 bucks. We will have offerings that enable them to buy, for instance, premium themes, instead of a subscription.

Techonomy: What's happening with the company. You're growing.

Boiman: We recently announced a deal with Deutsche Telekom for launching LTE in Europe. Today, when operators try to market LTE, it's just technology and a bunch of numbers. Why should a user pay more? Consumer-facing video is an enabler for many, many things--operators, LTE networks, storage, new cameras, better devices. Why do you need 4K video if you don't do anything with your video? But if you do something with it, if it suddenly becomes emotional and an important part of your life, and suddenly many things in the ecosystem can flourish. So Magisto is in many cases becoming an enabler to the next steps. We think now we are ready for the next level of

social media, which is video. Entertaining, emotional video can turn social media into something more entertaining.

Techonomy: And you're starting to work more closely with marketers and brands?

Boiman: We are producing people's life experiences. That's what we do. We're turning something boring into something moving and entertaining. And sometimes those life experiences are branded. So when you go to a Dodgers game and put your shirt on as a fan, you're branding yourself. What we can do is help the brand overcome their biggest problem right now with social media, which is that they're trying to broadcast instead of using the audience themselves. Brands on social media right now have bidirectional conversations with fans using text or photos. But with video it's a one-sided conversation. They have a channel on YouTube and essentially broadcast commercials. But with Magisto, by clicking a button, suddenly, brands can call their fans to action to share their life experience, the brand essence with video. So we can work with the team or the sport channel, for instance.

We just announced a relationship with NBC Sports. They have a show called *I Was There*. It's about fans, prosumers, shooting video of a game, of them watching the game. It's kind of social video done on TV. But they get to like five or ten movies, and they're not that good. Think about calling all the fans to do that, everyone at the field, everyone watching at home. Tell them, "Show us your life experience of watching the game," and we can also take highlights of the game, and put it in your video for you. So you get personal video, but it's professional. Hopefully this will turn out to be viral. You go to NBC Sports and claim a new theme, because you're a fan. And on Magisto you see an *I Was There* NBC Sports theme. It's life experience produced and shared between people with their friends. It goes from the fans to the friends of fans, which is exactly what a brand wants with viral video. But this happens in large scale and it's not an ad. It's essentially a one-minute commercial of the brand essence, from people who are like you. It lets your fans become your evangelists.

Techonomy: Where is the Magisto technology headed?

Boiman: We're trying to reimagine the next stage of photography. We want Magisto to be your 24/7 video editor, to put Magisto in the background to always be there for you. So if you capture things with photos and videos on your phone's camera, Magisto will surprise you with a produced movie. You don't even have to open the app if you don't want to. We announced a technology with Qualcomm at CES, called CamCrew. It's a camera crew inside your phone. We take some of the storytelling brains of Magisto and run it in real time, so now you have not just an editor, but you suddenly get a 24/7 photographer and director to help you get better photos. When you're shooting video or photos, you get notified by Magisto, "Here's some people you might want to highlight."

[He picks up his phone.] So let's highlight Josh [who is sitting at the table]. When I move the camera, Magisto knows that and will start to track him, resetting the focus and exposure, instead of me obsessing on the camera controls. Okay, smile for a second.

It reframed the picture, saw you smiling, and it's taken it for you. This is starting to do the thinking for you. If I'm starting to cut off your head or something, Magisto will tell you to put it in the middle. It helps the average photographer start to capture like a pro. When you press "done," Magisto produces the movie in minutes.

Techonomy: Does that happen on the camera or in the cloud?

Boiman: It's kind of hybrid. The heavy lifting of processing and editing is in the cloud. The directing happens in real time on the camera, using the [Qualcomm] Snapdragon processor. We started with the most powerful kind of application processor, which is the Snapdragon 800. But we'll add more and more platforms for lower and lower-end phones. But the idea is that your smartphone is already a monster in terms of computational power that's generally not used when you're shooting video. We have a smartphone but the camera is dumb. If you have the capabilities to change the photo to video, to go to slow-motion mode, or whatever, to do the right thing at the right time, at the end you're also going to get a much better edited movie. We're trying to change the paradigm. Today you assume that most of the things you capture are going to be archive. We want most of the things you capture to be shared immediately. That's what Instagram did for still photos.

Techonomy: It's really deploying AI in more and more ways as you go forward.

Boiman: Yes.

Techonomy: Who are some of your investors?

Boiman: We have Magma Venture Partners, Horizons Ventures, Li Ka-shing's private investment arm. We have Qualcomm and SanDisk. When you have CamCrew starting to work for you and starting to capture, you're going to need a lot of storage. And we recently announced that Mail.ru is an investor.

Techonomy: They want to help you in Russia?

Boiman: We already closed a deal with Odnoklassniki. We're integrated with them already.

Techonomy: That's the second largest social network in Russia.

Boiman: Yes. So they have like 33 million active users. We integrated Magisto with them. It's about producing everything that flows on that network. We think this is already

the most sophisticated integration of video storytelling into a social network anywhere. Everything that runs in to Odnoklassniki is turned into something that is shared. It's a look back for everything.

We have a feature coming soon called tweaking. You get a very, very simple UI that helps you correct mistakes or just state your preferences. You can highlight or de-highlight people and say, "That's the one who's important for me. I want to eliminate them from the movie." We're also going to show you a list of selected scenes, so you can just wipe them out and say, 'I don't want it, I don't want it, I don't want it." We're also going to show you already a selection of scenes we deleted from the final cut and you can just drag them back in if you want. It's not full editing, but in most cases, it's going to be good enough.

Techonomy: To have the ability to have a little decision-making power in the editing could be useful.

Boiman:Yes. It's a balance we're trying to keep. The user experience of full-fledged video editing is very bad on mobile, so that would really turn off most of our users. Most of them don't even need to tweak. We know that because we ask them all the time.

Techonomy: You have an ambitious agenda.

Boiman: Disruptive technologies often start from the root, not with high-end users. The first lousy digital cameras were used by consumers. They were not used by prosumers or by professionals. So that's exactly what's happening here. We cannot yet address the high end needs of professional video editors and brands, but this is exactly what consumers want. But we will just build in more and more sophistication, and add more and more tools. It has to happen this way.