Tom Ohanian
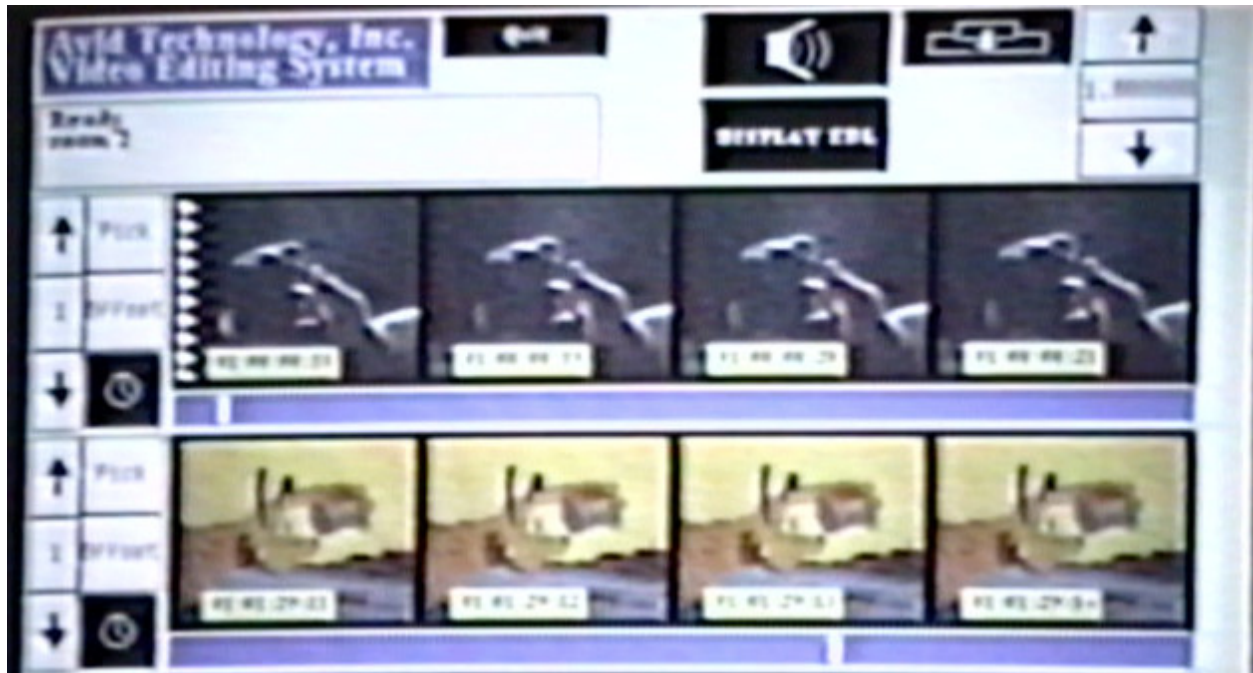**On The State of Digital Nonlinear Editing Series**
Copyright 2019 Tom Ohanian

Avid 1988 on the Apollo Workstation

On the State of Digital Nonlinear Editing
Number 1

The years 1988-1991 brought with them significant activity in the development of Digital Nonlinear Editing Systems (DNLEs). Systems coming from Avid Technology and from EMC—no not *that* EMC, but from Editing Machines Corporation were aspiring entrants onto the NAB show floor, then dominated by CMX, Grass Valley, Paltex, et al. Clearly, these weren't the first DNLE entrants. Back in 1971, the CMX 600 was created and used a combination of monitors, signal processing (video and audio), a PDP-11 computer and computer disk drives, each of which could record just over five minutes of NTSC video. And…it used a light pen for system control and as part of the GUI. Now that was cool…

Well, the past, indeed, was prologue and by 1988, various companies were in the early years of demonstrating DNLE possibilities. At NAB 1988, Avid did this off the show floor, demonstrating a rudimentary GUI, running on an Apollo workstation (see pix):

By 1989, Avid had moved onto the Macintosh platform while EMC was operating on the PC. These were software-based picture-decimating methodologies to reduce the overall image file size. Avid's software compression was at a staggering 250:1 ratio and those 24 bits per pixel were reduced to 4-8 bits. No matter. The fact is that DNLEs from both companies (and,

thereafter, systems from many other companies) took off in popularity and the end was inevitable for the linear videotape edit controllers and their companies. Unfortunately, those companies just simply couldn't get out of their own way, face creatively destroying themselves (they waited for the DNLE newcomers to do it to them), or figure out how to transform their business models.

Think about it—a nasty (yes, I said *nasty;* or I suppose I could have said, *pixel-challenged*) 250:1 compressed picture. And, in the case of EMC every other frame at 15 fps, and yet these systems took off. I know, because I was there designing those products and, at least at Avid, $253 million of combined recognized revenue and pipeline were on the books. And, remember, hardware-based JPEG compression hadn't yet been ubiquitously implemented.

Obviously, we cannot deny the dominance of DNLEs. They have forever changed the content creation, post-production, and distribution aspects of the digital media supply chain. We went from dedicated workstations, to workgroups, to distributed systems, and, soon enough, to virtualized systems operating over wide area networks.

So, now we've had DNLEs for 28 years or so. But, have DNLE constructs really—essentially—changed? Obviously, if you've read this far into the post, you probably can figure out where this is going. Because my fundamental posit here is, well, no, they haven't.

There are 7.5 billion people in the world. What has to change for digital nonlinear editing to be as ubiquitous as mobile phone usage? Based on all the data we've seen reported, 48% of the population will have access to the Internet this year. If you tallied up all the devices with an IP connection, there would be eight devices per person in North America. And yet, what must change for DNLEs to be used—daily—by not just the professional or prosumer user base but by, say, 50, 150, 500 million people? We'll talk more about that next time…
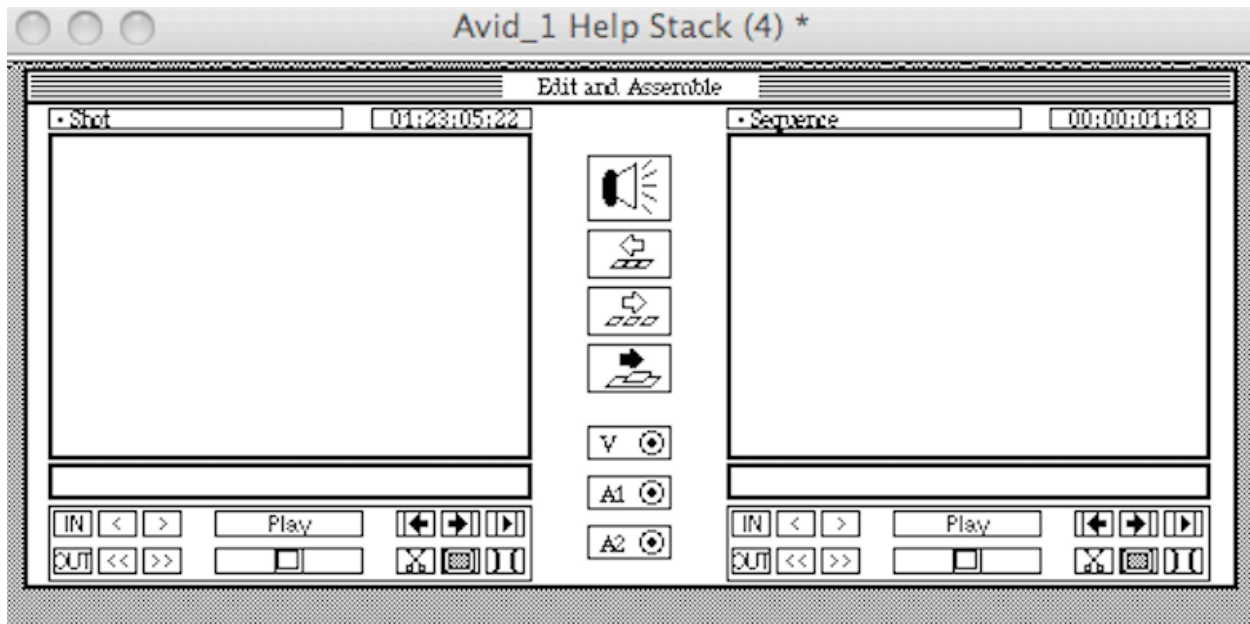
The 1989 pre-release version of the Avid/1

On the State of Digital Nonlinear Editing
Number 2

Well, first, let me thank those who have been kind enough to read my last post and who have left comments. It's great to see the passion around this subject from both a business and technical perspective from the many accomplished people who I have had the pleasure of meeting or working with over the years. Alas, dear readers and colleagues, the interest has spurred me on to continue…

The DNLE revolution occurred due to a confluence of technologies that continue to affect, improve, and, in some eyes, dilute our everyday lives. From a technology perspective, compute and I/0 capabilities were the main factors for the DNLE growth curve. Let's face it—relatively low cost personal desktop computers whether Mac or PC and lower cost video I/O processing were the two main enabling factors that got this started. We can't include storage—yet—because in 1987 it had a price tag of $15 per megabyte. Two years later, the price had dropped to $7.48 per MB and ten years later the cost was—wait for it—$0.02 per MB. Okay, so, great, ride *that* cost curve any day of the week. Similarly, video I/O (bless the Truevision VIDI/O box…) units for desktop video capture were affordable. Just think what would have happened to DNLE adoption if every unit required an SGI Sirius Video board which ran you a cool $24,900 in 1993 dollars... Network speeds weren't a factor and they wouldn't come into the picture until we needed those individual workstations to start sharing content in workgroups.

But, 25 years later, that triad, compute, storage, and network—inevitably invited another to the party—virtualization. Ooo—and that's where things are going to get *very* interesting. We'll get to that in a bit…

To understand where we are and where we are going, take a look at the graphic at the top of this post. That's early 1989, ladies and gentlemen. Thoughts that come to mind? Nice—not too many buttons. Too many buttons? Intuitive? Confusing? Uh…that's one big, big audio speaker (LOL)! Could you figure out how to use this product without a manual? And a thousand other questions.

If you'll indulge me, let me, for a moment, take you behind the firewall a bit because I can certainly say that there were days and weeks of trying to figure out and agree upon basic issues. Hours and hours of what, on the surface, seem to be the most mundane, boring issues. When you Mark In and Mark Out, should the duration be inclusive of the points or exclusive of the points? That took a long time to finalize. Because on that one, film editors expected one duration and videotape editors expected a different duration. Ever wonder why those first hard-coded keyboard maps had the In and Out marks under the, respectively, middle finger and index finger? It certainly wasn't a random decision. Lot of time spent on that one. Simple reason. You know why. They're the two strongest fingers, generally speaking, in humans and can take repeated abuse.

Along with the improvements in technology and ultimately the decrease in prices of those enabling technologies, DNLEs took off because they were similar enough to the operational status quo. Whether the status quo was the film editing convention of *feed and take-up reel* or that of video editing with *source reel and record reel,* or even, if we harken back to the CMX 600, and live broadcasts which use the *preview and air model*, they conformed to the thought process and thus the buying proclivities of these target markets.

The point is that these early systems had to be accessible and they had to easily adapt to the mindset and skill set of the user. If you were a film or videotape editor, it wasn't a stretch to learn the left to right model, the source/record modality, and—best yet—that ultra-cool thing called a timeline. That was a first for the video editing community who obviously never had a visualization structure like that. For film editors, that visual representation of their synchronizer was an easy visual adjustment. Similarly, the concept of *the bin* lent itself to the sensibilities of the film editing community while finally giving the video editing community a common view into all those source tapes on the shelf.

So, naturally, what could you disagree with if you were editing film or editing video? Okay, the two obvious answers were the *nasty* (well, *there*, I've used *that* word again) picture quality and the cost of storage. But the immediacy, the ease of use, and the ability to see your decisions later realized with the full resolution images was a simple straight line to draw.

As we leave January 2017, there's a story to relate. Wind back (or better still, rewind) the clock to January 1990. At Straight Cut Editorial, Tom Schacte, who was one of the original five Avid owners, had been a commercial editor working on film. For a two-week period, I sat with Tom in his Hollywood-based editing facility while he attempted to cut a series of national television commercials on recently released version 1.0 code (The Avid/1 shipped in December of 1989). I'll spare you the gory details but with a lot of work on both sides—manufacturer and

owner/editor—Tom cut the spot for Chiat/Day for Budweiser. And then a very interesting thing happened. He was contracted to cut one spot and ended up cutting 12 of them. Now, that was something he had never been able to do before without all that is entailed in getting black and white film dupes and beating the tar out of them.

And I think, at that moment, we both saw that beyond being a replacement for editing on film, there was something different at work here. I certainly never forgot the sheer delight he took in being able to give his customer something he could never do before—that is not without blowing through a budget. And something that—artistically—gave him joy.

But, there was something else going on in that room and soon the notion was pretty clear: choice—and not having to commit to a specific outcome translated to great value. Value to the owner, value to the editor, value to the client, and value to the manufacturer. Now, this was something different—ease of use (not having to learn something foreign) and choice were going to translate into increased earning potential for everyone involved. And that soon came to bear fruit.

When we next return (and I hope you will do so), we're going to talk about those editing constructs, revenue growth, and why those who are interested in reaching 500 million plus people are going to start chipping away at those conventions. They already have…

NOTES FROM TOM OTTAVIAN — LA + NYC

Mark O.

CONTINUING ISSUES FOR ENGINEERING
----

1. We should be able to click on an image (frame) in the frame mode of the Timeline and the Record master should travel to that location.

2. We should provide the means to generate a video only or an audio only EDL from the Master EDL.

3. The ability to resize the record monitor or the source monitor to approximately 80% of full raster and have this feature easily accessed by a toggle command.

First Feedback from Avid Beta Sites. Love that dot matrix printer...

On the State of Digital Nonlinear Editing
Number 3

Today's graphic is brought to you by the very first visit to get feedback from early DNLE customers. That's June, 1989, prior to any v 1.0 shipment.

And how far we've all come. But, to some, today's basic editing constructs of source/record, feed/take-up, preview/air seem quaint. And, to some, they may be very outdated. Well, that's to be expected. Remember—the very thing that made these systems take off is also the same thing that precludes adoption of them by really big numbers. Again, why aren't we seeing 500 million people "editing"? Well, it has a lot to do with what's natural and what isn't natural in terms of *communicating and connecting.* We're going to spend some time on those two things a bit later…

But for now, let's just think practically about DNLE adoption and growth. Top line revenue: growing, declining, or flat? It's certainly not straightforward to determine that answer. Because when we look at the 10K reports for various manufacturers, the DNLE component is never specifically called out. Sure, we can speculate, but it's really not a fair comparison without having the real numbers. But, segments do show us what relative value they have for the manufacturer. A case in point is to examine the quarterly numbers of Adobe's Creative Revenue (within the Digital Media Segment) which was $885.6 million in Q416. Those are quarterly numbers. Quarterly.

The takeaway, clearly, is that the segment—the digital media applications segment—is lucrative and from YoY examinations, growing. The key going forward, of course, is what magic combination of functionality, accessibility, and pricing is going to be necessary to expand the user base. That, in turn, will inform us as to the relative strengths and weaknesses of companies within this sector.

We know, from the great work that Joe and Josh do over at Devoncroft that the M&E industry is populated by some 3,000 vendors, 85% of which have less than three products. That type of fragmentation causes a lot of concern for customers regarding the business continuity contingencies that they may have to adopt if one or more key vendors either goes out of business, doesn't have enough capital for R&D efforts, or is acquired.

A fine point here—put yourself in this management seat—you're in charge of the operational aspects that drive $3B of revenue for television syndication and a key application that you depend upon comes from a company that does less than $15M in annual revenue. This is not fantasy. It's an actual situation. And your problem is that there just isn't another vendor that provides that functionality and it's intertwined with a lot of your systems. What do you do? For you, it's critical that this vendor stays in business and, hopefully, has enough capital to continue improving the product. Or do you roll your own because you can't put your company in that position? So, are you now a software developer? That's a problem if you don't really want to be in that business.

Alas, it's a stretch for many of the vendors within this sector because of the highly fragmented nature—too many companies...not enough revenue for a majority of the product categories. And that's why it's so important to grow the base within your area. Not the top end customers. There are just finite numbers of them. You have to grow the base. Well, technically, you can keep your base the same and do everything you can do become as profitable as possible, but you'd better hope that you keep those top end customers and that the average revenue per customer increases. That becomes harder and harder as technology improves and changes (remember *compute, network, storage, and virtualization* for now) because technology will always, inevitably, trump (no, not *that* trump) customer loyalty.

And, the magnification on technology and R&D is profound now and going forward due to the transition from SDI to IP-based technologies. Remember—when all is said and done—a single spine and four leaf IP topology provides the same configurative functionality as an 876 x 876 SDI-based matrix router. And as 10 GbE goes to 25, then 40, then 100, then 200, well…

So being able to make it through this next transition where applications ride atop generic COTS offerings is crucial. I am not downplaying the performance-related benefits of purpose-built appliances. It's true that—today—in some cases, virtualized instances of applications cannot touch purpose built appliances for some processing routines. However, that will all change very rapidly. So, if you're in that business, ride it as long as you can but prepare for the virtualized instances, transportable licenses, and flex licensing offerings that will be required. It's going to be a tumultuous time in the vendor community because the changes that are required at both technical and offering levels are coming very quickly.

So, when you starting putting the pieces together, we'll see changes throughout the M&E industry. Next time, we'll be spending some time examining the delicate relationships between content providers and service providers, the hypergiants who have higher capital valuations than those providers and the challenges and opportunities that are abundant in the M&E industry. We're going to look at valuations, growth, advertising spend, and moving at web scale and at web speed. We'll be doing that because the relevance that vendors have to their customers will be influenced by how content is made and how it is consumed.

Last, a question from a previous post was why ubiquitous editing is important and why everyone with a mobile phone should be doing it. Now, I don't think everyone with a phone should be editing, but I do think that the very fundamental way that pictures and sounds are put together will radically change. We must look beyond editing and come back to those two words—*communicating and connecting*—that's what people do. And how they communicate and connect using moving pictures and sounds is not going to conform to the current modes of operation of DNLEs. It will be something different. And that's what we're going to be interested in exploring.

But, let's follow the money first…
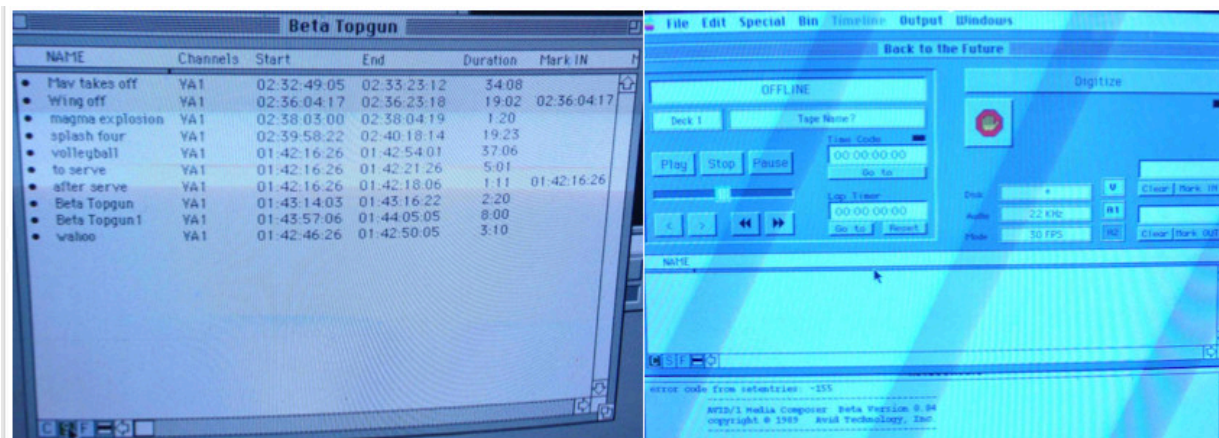
Beta Software v. 084 running today in 2017!

On the State of Digital Nonlinear Editing
Number 4

What is driving the incessant move to IP and to virtualized infrastructures? How will those transformations affect the landscape of M&E manufacturers? What will it mean for the companies who have provided products and solutions for the broadcast and post industries? And, you may well ask, 'Why do we care?' and 'How is this pertinent to the DNLE discussion?' Okay, fair enough. But, if you'll bear with me—the driver of technology is not because of what a manufacturer desires to make available to the customer. That power—the 'we decide what product(s) you get and when you get them' has moved from the manufacturer to the industry drivers. If that sounds funny or suspect, allow me this example:

Television programming for international distribution, in the old days, used to be based on a three-week schedule. Once that program was done, it would be available in Italy—or for that matter, Australia—three weeks later. Satellite transmission of the content or shipping videotapes were the methods for both aggregating as well as distributing content. Today, worldwide availability happens within 48 hours—out of an L.A. post facility on Friday and in Australia and almost everywhere else by Sunday. And, with 105 different format permutations (and sometimes many more), moving that much content in so little time and doing so in a cost-effective manner is what led to the rise of terrestrial IP aggregation and distribution solutions.

Now, it's not as if the vendors of satellite or leased lines for video transmission were lining up in the streets to creatively destroy themselves. But if you were a major content owner and you were paying that monthly nut for OC-3 and OC-48 lines which were unbelievably overprovisioned and underutilized, you were looking for solutions. The dirty little secret is that a lot of content and service providers—when they weren't using those lines to deliver real-time video—were using *deadly* (yes, I said it, *deadly*) FTP to send content via those expensive lines.

But that's all over now—the bulk of the highest value-per-second primetime content is traversing inexpensive links via secure, encrypted, UDP-based aggregation and distribution

solutions. A bunch of companies—you know them—Aspera, File Catalyst, Signiant to name a few— (and also some open source UDP code and talented engineers) have killed it and a lot of content providers and service providers have benefited from decreased CapEx and OpEx and have been able to fulfill the pressing industry demand—content delivered where it needs to be in the format it needs to be in because you—the consumer—are out there watching in continuing evolving ways (Yes, I know, unbelievably long run-on sentence; don't chastise me). And the incumbents—the leased video line providers and the satellite services providers were late (if at all) to the changeover.

Creative destruction is all around us—technology, business models, and for content providers and service providers—there are amazing opportunities and challenges to go along with this terrific journey we are experiencing. And whether you creatively destroy yourself (well, not *you*, unless you're a *complete* train wreck)—something will come along to make you either do so or you'll be left behind.

Okay, enough bluster—time to put some meat on this one. Follow me for a second here.

Let's look at some numbers* from 2016:

116.4M, 96M, 17M, $106B, $13.41B, $34B and, finally, -2.8%, -13.9%, -16.8%, 110.2%.

Hmm… now let's take a closer look:

116.4M Television Households
96M Pay Television Households
17M Households who get their television over-the-air (*free!*)
$106B in Cable TV Revenue
$13.41B in OTT Revenue
$34B in Network Programming Expense

And, more importantly, the trend:

The Pay TV Households: 100.9M (2011), 100.8M (2012), 99.3M (2013), 98M (2014), 97.2 (2015), 96M (2016), and 94.6M (projected 2017).

The Households who get their TV Over-the-air: 11.3M (Q412), 11.3 (Q413), 12M (Q214) and 17M (Q316).

Cable TV Revenue: $85B (2012), $91B (2013), $96B (2014), $100B (2015), $106B (2016).

OTT Revenue: $2.48B (2008), $2.77B (2009), $3.18B (2010), $4.18B (2011), $5.24B (2012), $6.57B (2013), $8.27B (2014), $10.48B (2015), $13.41B (2016).

Network Programming Expense: $22B (2011), $24B (2012), $27B (2013), $29B (2014), $31B (2015), $34B (2016).

Annual Growth of Advertising Revenue:

Television: -2.8%
Newspaper: -13.9%
Radio: -16.8%
Mobile: +110.2%

While there may be no huge surprises here, it's important to be mindful of these trend lines. Pay TV Households are down. Put aside the issues of whether the pundits are right or wrong regarding cord-cutting, cord-nevers, skinny bundles, switching, etc. *The numbers are down.*

Over-the-air households—growing and not an insignificant total number—that's almost 19.8 million households who do not pay for broadcast television. And if you could get them to pay the average pay television bill of $65/month? Well, that's $780 buckaroos a year multiplied by 19.8 million households. And that is *very* significant…So where are they going for their non-broadcast programming?

Cable TV Revenue—Growing. But the question here is deceivingly straightforward—can it be sustained if subs are lost?

OTT Revenue—Well…the numbers tell the story…but behind those numbers, comes a rolling ball like that big one in Raiders of the Lost Ark—what's the associated content cost burden?

Networking Programming Expense—an echo of the flipside of the OTT revenue issue—content costs are growing faster than other operational expenses.

And, finally, look at the percentage growth of mobile advertising revenue. Sure, we all well understand that mobile ad rates and television ad rates and the CPM behind those rates are quite different. However, we're looking at annual growth rates and nothing—*nothing*—is close to the growth of mobile. Better have that in your strategy whether content provider or service provider.

Next time, we're going to look at the disruptors—the Hypergiants—and we're going to see massive war chests of cashola, spectacular YoY growth, and we'll see where the real power lies—With the Content Providers? With the Service Providers? Or with the Disruptors? (Hmm… And what did Facebook do on the first of February and why did they do it?)
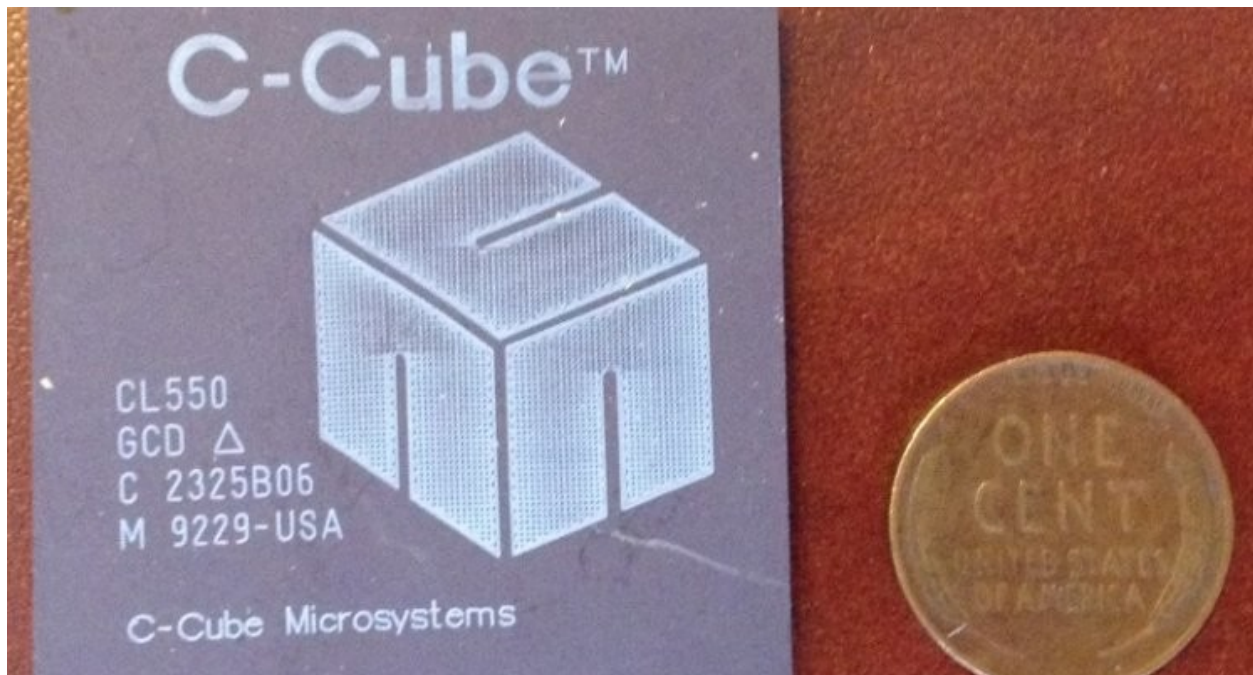
The Chip from 1991. I pulled this off a Janus Board...

On the State of Digital Nonlinear Editing
Number 5

With the last post, I highlighted some of the trends that are occurring in the content provider and service provider space. Revisiting them*:

Pay TV Households: Down 6.24% over six-year period (2011 to projected 2017).
Free Over The Air Households: Up 50.4% over a four-year period (2012-2016).
Cable TV Revenue: Up 24.7% over a four-year period (2012-2016).
OTT Revenue: Up 440.7% over an eight-year period (2008-2016).
Network Programming Expense: Up 54.5% over a five-year period (2011-2016).

And, finally, Annual Growth of Advertising Revenue:

Television: -2.8%
Newspaper: -13.9%
Radio: -16.8%
Mobile: +110.2%

What do the gross figures indicate? Pay TV subscribers going somewhere else. Increasing numbers of people who get their TV for free. Cable TV Revenue going up. OTT revenue soaring. It's costing more to create network programming.

Now, let's talk about some corporate valuations. As of December 23, 2016, who had the highest valuation, worldwide? If you said, "Apple" you'd be right—$621.3B. Okay, probably no surprise as to numbers 2 and 3—respectively, Alphabet and Microsoft. Berkshire Hathaway is number 4 and Amazon is number 6 followed by Facebook at number 7. For telcos, AT&T is number 12 and Verizon's number 16. And when do the media companies show up? Comcast is first at number 27 followed by Disney at number 28.

Okay, and now you'd be justified if you said, "Tom, what exactly are you getting at? Apple isn't Comcast and Disney isn't Amazon." True, true, true. But, bear with me, the journey will get us there. Now let's take a look at Facebook's revenue for Q416: $8.629B from Advertising and $180M from Payments and Other Fees for a total of $8.809B. And CBS? $1.796B from Advertising, $893M from Content Licensing and Distribution, $770M from Affiliate and Subscription Fees, and $59M associated with other income for a total of $3.518B.

Facebook and CBS are both terrific companies. They are certainly world class in the creation of value for their shareholders and both had stellar 2016 performance. But, again, we're looking at trends here and let's look at YoY growth percentages. Let's consider three different segments in the U.S. market.

The Content Providers: Approximately a $238B market with 4% YoY growth.
The Service Providers: Approximately a $165B market with 1% YoY growth.
The Web/OTT Providers: Approximately a $13.4B market but with 54% YoY growth.

54% YoY growth? Well, again, you're right if you say, "Yeah, Tom, fine, but that growth comes with a cost." True enough, but if the growth can be sustained and the cost structure over time is less than the others... Or, if a newcomer has *a lot of cash*?

Let's consider the movement which is actually occurring as these three types of providers go about augmenting their core businesses. For example, Content Providers are moving towards a direct relationship and direct distribution to consumers. To that end, they will take on more of the appearance of that of a Service Provider. Consider Disney as it prepares for some type of OTT offering for ESPN. At the same time, Service Providers are moving towards the content creators—the most obvious past example is the Comcast acquisition of NBCU. And AT&T's interest in TimeWarner is yet another example. Stepping back, it's easy to see. The dance is being adequately diversified in order to capitalize (literally and figuratively) as one line of business wanes while another grows.

And, in the middle, are the Web/OTT providers. But these companies have a different history and lineage. They're not the traditional content providers. They're not the traditional service providers. We can clearly point to the transition that Netflix and Amazon and others are making wherein they exclusively licensed content and then started to create their own, following the HBO success model... In other words, they're not a broadcast network and they're not a large MSO.

And those I lovingly call The Hypergiants? These are the Amazons, Apples, Facebooks, Googles, and Microsofts. Plural because there are going to be more entrants. Those companies have a lot of users, they have a lot of information about us, they have a lot of credit cards on file, and they have a lot of subscribers to some form of service. Let's, again, look at sheer numbers: It is now speculated and reported that Amazon Prime has 70 million subscribers. Now, we don't know how many of those subs are also part of the SVOD business or accounting for rental or EST revenue, but that's what we know…so far. Netflix has 93 million subs. As of February 2016, Apple had 782 million iCloud accounts and something more than 885 million iTunes accounts and over 1B active iOS devices. And they're making $41B a year from those connected devices… Vertical integration anybody? All you need is the content.

Now, let's turn to Facebook. A treasure trove of activity in February alone… Launching an application to stream their News Feed videos to STBs, Apple TV, Amazon Fire, and Samsung Smart TVs. Talking with MLB to stream one major league baseball game per week… But, most importantly, Mr. Zuckerberg starting an earnings call with these words: "I want to start by talking about our work around putting video first across our apps… …video is only going to become more important. …that's why we're prioritizing putting video first… …and taking steps to make it even easier for people to express themselves in richer ways."

Whether or not you believe in the power of The Hypergiants, consider that 23% of corporate cash is held by Apple, Google, and Microsoft. And with the Hypergiants growing faster than both content providers and service providers, that cash horde is likely to increase. Further, let's not forget what happened on December 17, 1993. For, on that day, this network called Fox successfully bid for the television rights for NFC games. And that, as we all know, was the major turning point for that network.

Which must bring us back to cash. Cash to acquire, cash to invest, cash to use for R&D and cash to redefine how people communicate and create and consume content.

And yes, it's absolutely going to influence content creation applications. More on that next time…

From Wired via Riot. 32M viewers for an eSports event!

On the State of Digital Nonlinear Editing
Number 6

We're into the sixth post now, and some of the feedback I've been getting is interesting. Some readers are intrigued by where these posts are headed. Some are very interested in the data and the metrics. A few readers aren't terribly happy with me because they're wondering when I'm going to get around to discussing digital nonlinear editing. After all, that's the title of the postings, right? But the fact is that we can't have a meaningful discussion about the state of DNLEs (and frankly of the media industry's vendor community) without discussing the business states of content providers, service providers, and the Hypergiants.

There are more than enough forums to discuss DNLE features and functions and to gripe or to commend. But that's not going to move the needle in terms of your business, your career, or the collective knowledge of these industries. Again, the reason that I've spent so much time discussing the landscape is because the applications that create, distribute, and generate cash for content are going to be influenced by much larger companies and much larger concerns. And that's why we must spend time thinking about strong companies, weak companies, well positioned companies, and companies that have a massive influencing hand which will dictate how these applications evolve.

Over the last few posts, I've covered the growth rates, valuations, some cash balances, and the structural shifts that are occurring. Content Providers who are becoming Service Providers. Service Providers who are becoming content providers and are buying content companies. And the Hypergiants—the disruptors who are both. And these Hypergiants—make no mistake—are

disruptors. Let's not forget: Netflix is winning Creative Emmys and Amazon funded films just took away three Academy Awards.

Now, when I wrapped up the last post, I outlined that three Hypergiants—Apple, Google, and Microsoft—hold 23% of corporate cash. And Amazon and Facebook are continuing to grow and acquire companies. And some of those acquisitions sit squarely within the areas of media software and content creation and design. In previous posts, we also began tracking the trends in viewership—both traditional television viewership and the growth of OTT viewers. And we already know that—putting per viewer ad revenue aside for the time being as a given—annual mobile viewing grew +110.2%.

So, consider: One billion hours of YouTube videos watched per day. That almost matches the number of hours of traditional television watched per day. And online viewing of content is no longer limited to short form (less than five minutes) content. The range for online viewing appears to be anywhere from 3 to 30 minutes, constituting a single view. And I think we can all expect that the trend to providing more compelling content—at longer durations—makes complete sense for longer period audience retention (and thus more advertising possibilities mid-roll or via subscription).

Okay, so consider. More content moving online. Content Providers and Service Providers are seeing viewers accelerating in their embrace of time shifted and place shifted content consumption. Conventional thought is that live sports programming is immune to these content viewing shifts. Well, there's a lot of money at stake if that isn't the case. Through 2022, ESPN has almost $50B committed to long-term content rights across 12 major sports franchises. The billion-dollar (more?) question is whether consumers will do the opposite of what they're doing now—which is to sit and watch those live sports broadcasts.

And if they don't? What happens if viewing habits and generational notions result in the viewing of 15-20 minutes of curated "best of moments"? What happens to the ad-dollar value of live sports programming? These aren't idle musings—do your own non-scientific method polling—ask your friends and I think you'll be surprised at the percentage of people who are—increasingly—perfectly fine with seeing highlights. I'm not saying that this is going to happen for the Super Bowl—but there are a lot of events which may be susceptible to viewers foregoing the live experience for the highlights. Then it will just be a question if the rights holders make those highlights available in curated form. My posit is that this will be inevitable.

And, now, let's really think about how these habits, the rise of the Hypergiants, and how traditional television viewing meeting social will fundamentally change the requirements of content creation applications. Sure, there will continue to be a need for craft editing—that's not going to go away. Neither is long-form editorial. But, if the nature of programming consumption is changing—and we know that is happening—then the creation of that programming is also going to change.

Take another look at the graphic at the beginning of this post. This is from Wired via Riot (say that fast three times). That's 32 million viewers who tuned in to watch The League of Legends World Championship, making it the most watched eSports event. And the others? We'll the numbers give you a very interesting sense of really large audiences…

Think about this for a moment: the fact is that creating programming for television has not had to change for a very long time. Yes, there is more visual sophistication (visual effects, etc.). Okay, file-based workflows, etc. But, fundamentally, creating commercials, episodic short form, long form television, theatrical—the workflows have been static for a long time. But, as viewers change how they view content, the methods by which that programming is put together is also going to change.

And that is going to have a direct and profound effect on the applications that are used to make content. DNLEs haven't had to radically—and I mean radically—change because the nature of content viewing hasn't changed. Until now… And that's where we're going next…

Facebook Live Streams at 11:03 AM ET in the U.S. Only the beginning!

On the State of Digital Nonlinear Editing
Number 7

At the end of Post #6, we landed on the fact that the content creation process has not changed all that much in a very long time. Again, I'm not downplaying—at all—the importance of file-based workflows or the current transition of SDI to IP infrastructures. I've covered that in previous posts. But the fact of the matter is that we have to look at where the eyeballs really

are going to be. And where they're going to be is absolutely going to have a major influence on the content creation to consumption digital media supply chain.

Disruption, Disruption, Disruption. Disruption in the media industry is happening so fast we may lose track of it. And the conventions by which we measure disruption must be examined. Now, it wasn't too long ago that the only comment associated with YouTube was about the inanity of cat videos. Today, YouTube is generating $4B in revenue and is launching its live streaming TV service. Twitter moving ahead with Periscope Producer… Let's turn to Facebook. Does Facebook Live peter out? Is it an experiment? Or, maybe trying to figure out if YouTube or Facebook or Apple can compete in Live TV streaming with affiliations to live local and network content may not be the pertinent issue at all. Maybe, in the long run, that's just meaningless. Does Facebook fail with video if it doesn't land premium content deals? Maybe, instead, the pertinent issue is the sheer amount of people spending time on those sites and that global audience. Who—by the way—are providing loads of information about themselves…and for whom video—both short form and long form are getting very sticky. And, oh, by the way, there'll be a whole lot of companies present within the Facebook Live area at NAB…

A pointed question. Do we really think it's inconceivable that the sheer number of content creation devices—the masses refer to these as mobile devices—won't lend themselves to a complete disruption of how content is created, contributed, and distributed? Let's take "news". First the dominant domain of the major broadcasters. You know—Murrow, Cronkite, Huntley, Brinkley… And then cable broadcasters, obviously CNN and for governmental affairs, CSPAN. But, today, those billions of devices, create the modern-day version of the stringer.

Ah, yes, the stringer. It's one thing to be a news organization with a staff and a network of professional stringers. It's an entirely different scale—and an entirely different proposition—if a platform enables millions (why stop there…billions) of stringers. Yes, sure, we toyed with the idea of citizen journalism. But that was before this massive device proliferation, the expansion of Wi-Fi, and most tellingly, the eventual effects of 5G. And, back then, there was no incentive that the platforms of today have the capability of providing. And, 1.3B people on a platform? A platform (whether from YouTube or Facebook or…) that can produce opportunities and create all kinds of audiences?

Take a look at the graphic at the top of this post. Those are the Facebook Live streams at 11:03 AM ET in the U.S. And that's only the beginning.

There are a number of different phrases that are being applied to these behaviors: micro-broadcasting, niche-broadcasting, micro-channels, pop-up channels. Whatever we call them, there is an inevitability that—when you have billions of people being able to self-contribute and self-distribute—disruption will ensue.

Okay, so even if some of that happens, and since we know that Facebook and YouTube et al are going to have all sorts of content offerings (because they're obviously already well down that

path), then I ask you—are the applications that we know today—are they ready for the new forms of content making that are around the corner?

Here's an everyday example. Broadcasters will eventually merge the broadcast and the social side of the enterprise. But, today, they are—largely—two separate parts of the entity. Separate just because of the rapidity with which social has had to be implemented at the broadcaster level. Acquiring or creating the content and then packaging it for broadcast is standard fare. But then that content has to be handed to the team responsible for the social side. Almost always, you will find that each group uses different tools. The metadata and essence of the content may, and often does, change. The two groups—the broadcast group and the social group—may not even be cross-trained on what each is doing, from common nomenclature to workflow. And, in some cases, the deliverables are completely reversed: social first, broadcast second.

Now, fast forward to those Hypergiants I've been talking about. When content starts to proliferate these social networks, there is a different set of editing, content manipulation and transformation tools that will be required. There will be less of a need for multi-function craft editing systems. (We could paraphrase Sir Sean Connery from The Untouchables: '…you don't bring a knife to a gun fight…'). In other words, fewer sophisticated editing functionalities. Instead, more templated approaches to being able to very quickly brand, post, analyze, update, and scale to an ever-increasing set of distribution outlets and networks.

I'm not talking about multiplatform distribution here—that's a volume and SKU exercise. I'm not downplaying its complexity—and we've talked about the logistical challenges of multiplatform distribution for television syndication. However, when content lands in the world of the Hypergiants and social platforms with a billion plus daily visitors, the significant net-new here are micro-channels, branding, and dynamic responsiveness to content consumption.

And that content is going to be tightly integrated with sophisticated AI-driven personalization models. After all, according to Facebook and Google, it has been the continued application of AI-based refinement that has been a major contributor in not only the increased numbers of viewers but also the increasing amounts of time that viewers are spending watching video (again, sorry for the brutally long run-on sentence; don't chastise me). And, we'll increasingly see companies like IBM and Sprinklr focusing on these massive amounts of video streams combined with real-time user interaction, all in service of brand management. Let me repeat. Brand management. Not video views, not trending videos. Managing the brand and all that entails as brands go online with dynamically changing media.

So, what applications, what technologies, and what companies—are going to be able to provide these types of functions? It's important to remember—it's not going to be about refining the cuts, or matching those cuts, or preserving the 180-degree rule. The growth—the massive growth—is going to be less on the bells and whistles and much more concentrated on the innate structure and personality of the content and the consumer. And, the content manipulation tools have to be simple and have a high degree of automation.

And that's where we're going next…

Automatic Editing of Footage from Multiple Social Cameras from Disney Research Hub

On the State of Digital Nonlinear Editing
Number 8

Over the last seven posts, I have been covering the shifts that are occurring in the media industry. Content providers, service providers, and the Hypergiants are no longer finding previous lines of division as distinct. And whether it is vertical integration (Comcast-NBCU, the possible AT&T-Time Warner merger, etc.), or the Hypergiants (Amazon, Apple, Facebook, et al) seeking to add live programming, we've talked about how quickly content distribution and consumption are changing.

I have been creating this overview of macro movements for us to be able to concentrate on the various changes that will also occur in the content creation, preparation, and delivery stages. I believe that there are six distinct areas that will need to be supported to address the different (and sometimes converging) needs of the current and evolving content consumption platforms. On the very first post, the questions I posed included: Why aren't a billion people editing content? What must change for that to happen? In the last two years, there have been such significant changes in the content creation to consumption experience that new approaches will be necessary to fulfill these evolving content exhibition models.

Okay, based on the feedback I've gotten, a lot of readers have really enjoyed the data regarding the growth and decline of both content providers and service providers. And, especially interesting to readers has been the YoY growth of the Hypergiants. How they are not only

dominating mobile content consumption and advertising, but are now negotiating for the rights to more expensive content.

On the other hand, some readers have asked me when I'm going to write about digital nonlinear editing. But there are Facebook groups and other forums that provide specifics of DNLE operations in support of the different DNLEs in the marketplace. So, rather than write about the feature sets of this product vs. that product or the strategy of this company vs. that company, I think it is, ultimately, more important to understand the industries and the economics that will require vendors to form and adopt relevant strategies in support of the digital media supply chain.

Given this incredible change of content distribution and consumption, what are the likely and required types of applications and functions? What segments are likely to rise, plateau, and decline? And how will those factors affect the types of required applications?

A long time ago, I wrote a couple of editions of my Digital Nonlinear Editing book (no shameless plug—it's out of print… though, ahem, the Digital Filmmaking book I co-wrote with Mr. Phillips is still available…grin). Now, in the DNLE book, I charted Six Waves of DNLEs:

The First and Second Waves of Nonlinear Editing: Tape and Laserdisc-based, The Third Wave: Digital Nonlinear Editing, The Fourth Wave: Digital Nonlinear Online and The Desktop Applications Explosion, The Fifth Wave: Intraframe Editing, The Sixth Wave: Digital Media Management: Open vs. Closed Systems, File Compatibility, and the World Wide Web.

Now, recall that these waves were outlined in 1994 and 1998, but, even so, they are all still applicable and in-play today. The WWW reference was originally intended for real-time WAN-based editorial though that hasn't yet become mainstream.

So, where are we now and how are content creation applications going to evolve? I believe that there are at least six distinct functional sets that will plot the transformation of content creation applications. Similar to what was outlined above with the different waves, let's consider the following functional sets and the different needs:

The six distinct operational characteristics that will impact the digital media supply chain and resulting content creation application functionality:

Craft Editorial.
Specialized Content Editorial and Packaging for Platforms.
Automated Content Packaging.
Artificial Intelligence and Machine Learning Applied to Automated Content Creation.
Interactive and Game Engine Applied to Content Creation.
A Completely Different Approach.

Craft Editorial. For this discussion, let's just agree that I'm not distinguishing being picture and sound. Let's just keep them both in the discussion as one. Craft editorial is, of course, the most familiar to people. This is the very well-known process of editing raw content to its finished form—whether commercials, short-form, long-form, television, theatrical, documentary, and so forth. Craft editorial on film and videotape have obviously been eclipsed by digital nonlinear manipulation. However, depending on the type of content, there is no longer an equal correlation of the amount of content required and the growth of craft editorial systems. This may seem contradictory—how can there be an increasing number of channels surrounding us and yet less demand for craft editorial? I agree, it seems very strange, but consider: on the theatrical side, studios are making fewer movies.

Six major studios and 10-12 mini-majors combine for between 120-150 movies a year in the U.S. But the number of films per studio has been consistently dropping over the years. In January 2017 Sony took a $962M write-down for its entertainment division. And even if you put that aside, look at the amount of films that Sony made through the last six years:

2011: 18
2012: 20
2013: 12
2014: 10
2015: 12
2016: 10

And those decreasing YoY numbers aren't exclusive to Sony. Fewer theatricals.

We must look at the raw economics of content that, in turn, fits into the craft editorial category. The average cost of prime time one hour slot U.S. television programming is approximately $3M. And for cable programming there's a huge range—$50K-$75K up to $5M. Five million? Really? True Blood came in at approximately $5M per episode versus $3.8M per episode for House of Cards. And those costs eventually did come down as multiple episodes were shot simultaneously. But those are the outliers, and, naturally documentary / reality come in at much less expensive per-minute costs.

If feature films are finite in their number and are plateauing, then the number of creative craft editorial seats for feature films cannot measurably grow. As channels expand and as OTT offerings become very prevalent (and ultimately dominant), the programming may not look different, but it will have to be created differently. The cost per minute of programming must come down as these new distribution / consumption channels grow and as they permeate different parts of the world.

Once we start to create programming for sets of channels where the viewership is no longer in the millions, not in the hundreds of thousands, but in the thousands, and eventually in the hundreds, the content will need to be prepared in different ways. In other words, as channels

expand, creative craft editorial will not correspondingly expand. The sheer amount of content required will have to be produced and post-produced at lower costs. That brings us to:

Specialized Content Editorial and Packaging for Platforms. In a previous post, I outlined the relationship between broadcast operations and social platform operations. Creating content for social platforms is different than that of linear broadcast and different applications are required. Here, the emphasis is on lightweight editorial functionality while tagging, branding, expanded metadata functionality, and reacting to social network feedback and iterating that content are emphasized.

Personnel will not be concerned with the nuances of craft editorial because they will need streamlined tools to make slight alterations, cutting down to length, pulling out snippets from longer clips and so forth. A few years ago, a well-known manufacturer of DNLEs was asked by a global financial services company to create such a streamlined application. This was to be used by employees to make shorter clips from archived recordings to circulate them to their internal and external network of analysts. And social media platforms which in aggregate generate billions of page views represent an ever-increasing real-time outlet for slightly modified and heavily branded content. Different sites + different platforms = slightly modified essence content with greatly modified branding and a great deal of volume and variations.

Automated Content Packaging. This is the classic ingredients and recipe analogy. However, imagine making 100 different kinds of cakes with 100 different kinds of fillings, in 5 different sizes, and with 200 different possible toppings. The problem becomes unwieldy without a supply chain mentality to solve it. Hence, you radically reduce the choices. Unfortunately, with content, possible combinations only expand. Languages, bit rates, aspect ratios, captioning and so forth only contribute to the permutations. In manufacturing, we'd equate this with a SKU-based system where multiple products (SKUs) are required and all the components to make each SKU are part of a bill of material. This is automated content packaging according to the various required deliverables. And every possible cost that can be reduced will be reduced—hence the outflow of many of these workflows to cloud providers.

Next time we're going to look at the additional areas:

Artificial Intelligence and Machine Learning Automated Content Creation.
Interactive and Game Engine Applied to Content Creation.
A Completely Different Approach.

A preview into one of those areas is referenced in the above graphic, based on the Disney Research Project regarding the "Automatic Editing of Footage from Multiple Social Cameras". We cannot afford to make light of how machine learning and AI are going to enable the automatic nature of content preparation and creation. This is about quick and logical assembly when we have massive amounts of content to contend with and which requires automation. And this is only a start.

Until then, another NAB is approaching. As you move through the show, think about the entire Digital Media Supply Chain. Keep in mind these next coming waves. Whether your perspective is that of a content provider, service provider, studio, broadcaster, corporate, state, local, federal, educational, individual—you play an important role in this complicated supply chain.
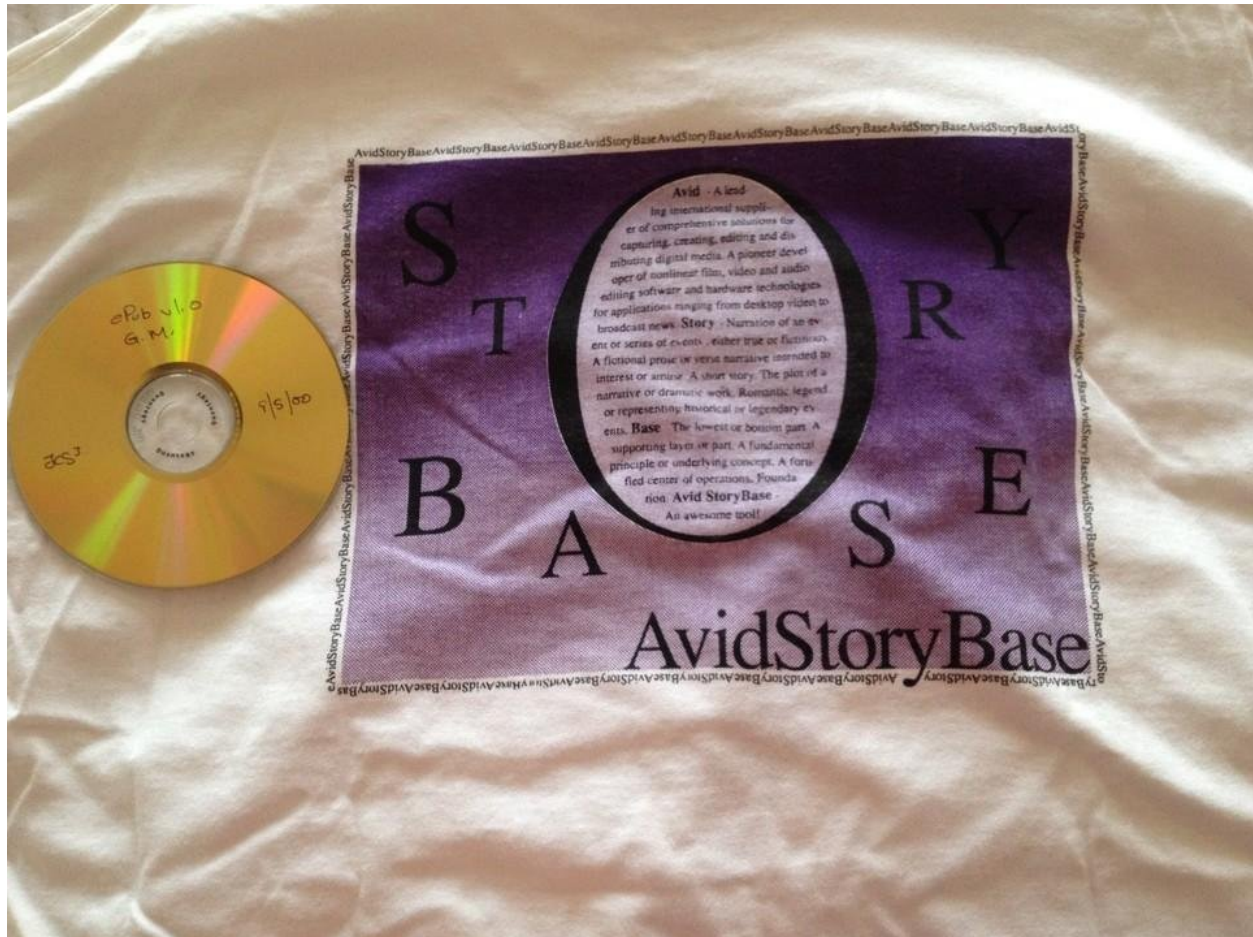
StoryBase: an early program from Avid that provided story-branching functionality.

On the State of Digital Nonlinear Editing
Number 9

It is now post-NAB and there are many sites and articles available to summarize the major introductions and themes that characterized NAB 2017. So, while I won't dissect every major piece, a few of them bear obvious mention.

As we expected: IP, IP, IP. The transition from SDI to IP continues and there were numerous venues showcasing these developments. And, as the industry requires, companies that

ordinarily compete at the broadcast controller level, switching level, or the destination device level showed successful interoperation. We saw real-world, demonstrable interoperating systems utilizing a combination of SMPTE 2022-6 and 2110; SMPTE 2059-PTP; and AES 67. What's particularly encouraging is that vendors at those levels (broadcast controller, network control / switching, legacy SDI, and end device) recognize that it's in their and in the best interest of their customer base to get on board. While not complete, the message is clear: interop or whither then die.

As we expected: OTT, OTT, OTT. OTT for content providers and service providers. Some solutions that focused only on multi-device encoding. Others focusing on mobile device-specific features. Some offering end-to-end solutions starting with virtual head-end implementations continuing through to set-top box and end device solutions. Your choices are distinct: all in with one solutions provider, or point products that are pieced together. But the looming questions in discussions is what does an outsourced managed service look like for the buyer and does that sell model provide for a pay as you grow of your subscriber base?

Some other notable developments. The Facebook Live Pavilion in South Hall. Yes, it was small and didn't have hundreds of vendors. But, it was a very well-traveled section. There were many innovative introductions. Lots of live streaming solutions for mobile, and software-only full-featured production studio offerings. That natural progression—from big to small, from hardware to software, from established companies to the next set of companies that are software-centric, built primarily for browser-based access, were in evidence.

And then there's Cloud and Virtualization. Here, it's still early days. Yes, some vendors touted virtualized instances of their applications which currently exist in a hardware appliance. Asset management, graphics, cloud-based video switching were among the categories. But there is still a lot of work to be done such that we can get to end-to-end solutions that will provide us with unlimited flexibility. And what is most interesting are customer requests for fine granularity regarding knowing exactly what is being used and only paying for what is being used. In other words, to harken back to one of my earlier posts, the notion of over-provisioned and underutilized equipment (hardware or software) is going to go by the wayside.

Most telling is that a clear shift has been taking place at NAB in two dimensions. First, the separation of entrepreneurial smaller companies at one end and, at the other end in terms of size, some of the largest companies in the world. Take the small start-ups in the Facebook Live Pavilion and contrast them with these companies that had significant booth presence: Amazon Web Services, Google, IBM, Microsoft, and many others. In the middle? Those companies that aren't small and they aren't big. These are the companies that have the bulk of the M&E customer base, but, as I have outlined in previous posts, there are just too many of them—1700 at last count and their revenue ranges from $12M to $500M!

There's nothing subtle about it—the Hypergiants (except for Apple) are all declaring that NAB was the place to be to showcase their relevance to the media industry. AWS, Google, IBM, and Microsoft were all showing and talking about automation, AI, Machine Learning, and Scalable

Services. And they were not only doing it in their booths but in very visible stage presentations. A case in point—this session from Google: "Media in the Age of Machine Learning in the Cloud".

And, then, during NAB, the announcement of a significant number of high profile jobs being cut at ESPN. In earlier posts, I outlined the significant capital expenditures to secure long-term licenses for sports franchise content. Those obligations and the loss of 12 million subscribers in six years continue to place a spotlight on the changing viewing habits of today's consumers. And that news was talked about on the show floor regardless of what hall you happened to be in. The net is that change is happening quickly in the content consumption arena. Previous posts have outlined the money flow and we can see—especially with the ESPN announcement and NAB developments—that rapid changes will now be required of all vendors.

Which brings us to the six distinct operational characteristics that will impact the digital media supply chain and resulting content creation application functionality. They are:

1. Craft Editorial.
2. Specialized Content Editorial and Packaging for Platforms.
3. Automated Content Packaging.
4. Artificial Intelligence and Machine Learning Applied to Automated Content Creation.
5. Interactive and Game Engine Applied to Content Creation.
6. A Completely Different Approach.

In the last post, I outlined the first three. And at NAB, we began to see the clear importance of 4-6. As we have discussed, consumers need to have access to the content they desire, where, when, and regardless of the device type. At the same time, more new content is being generated and content libraries are being mined. Here is but a sampling of AI and Machine Learning technologies that are being applied to automated content identification and creation:

Cognitive Metadata Extraction
Tonal Analysis
Visual Recognition
Speech to Text
Language Recognition
Near Human Voice Quality Dubbing

And this is just a start of how content will be automatically indexed, categorized, and, ultimately, combined into finished sets of products. A real-world example: You're a broadcaster. You have 300,000 hours of content in your library. Yes, that much. You've been in existence for many decades. You also have 80-plus individuals reviewing a portion of this content as files, trying to determine the who, when, what, rights holders, etc., of the file being reviewed. And the problem, of course, is that you simply can't get through it fast enough. AI and Machine Learning applied to this problem will result in more robust, accurate, and timely solutions. AI is continuing to advance in terms of the ability to automatically recognize

thousands of object types with accuracy and speed that often rivals that of a human. And as AI systems process more content, the ability to surpass human processing time will be significant.

Let's take another problem. This one has to do with content production and post-production. I'm talking about the classic green screen extraction issue. First it was a black background, then a blue background and now a green background. But, AI algorithms have the potential to change all that. Adobe and Beckman Institute are collaborating to create a new approach. Their "Deep Image Matting" paper outlines AI being applied using over 49,000 training images and 1,000 testing images, resulting in matte extraction without solid backgrounds being used and reporting success in daylight exterior conditions. And the reason this is even more important than ever has to do with the fact that not all content is of equal value. We know that a studio tent pole picture has a lot of value associated with it and that you will have a large staff to work on these images and composites. However, when you are creating content for an audience of smaller sizes, content must be composited in automated ways. Just think of the applications for call center CRM, training, and so forth. If it's easier to create content, will more be created? I think so.

Getting to content faster means getting to revenue faster. And AI and Machine Learning will be applied in ever increasing areas. That was clearly in evidence at NAB—sometimes in smaller companies and sometimes at much larger ones as indicated above.

Many of you, no doubt, have seen the use of IBM's Watson to create the trailer for 20th Century Fox's horror / thriller, Morgan. Now, a lot of people have downplayed this because they argue, "Well, a human editor had to do the assembly". True enough, but the process leading up to that is important for us to study. According to Wired, 100 horror film trailers were analyzed scene by scene for their visuals, sound, and composition. 90 minutes of Morgan was processed with ten scenes being isolated. From a time savings perspective, 10 days reduced to 24 hours.

Again, it's not a process without human intervention and review, but as I've stated in previous posts, AI, Machine Learning, and Automation are necessary going forward to create a responsive, efficient, digital media supply chain.

Interactive and Game Engine Applied to Content Creation.
Interactive programming in the 90's was primarily achieved using laserdisc technology. I worked on many these projects and, today, the process will read as very quaint, antiquated, and to some—medieval—in a Quentin Tarantino kind of way. Record all the narration on ¼" magnetic tape. Cut said tape with razor blade and splicing block. Splice good takes together according to script. Trim said tape to create proper cadence. Record video segments on analog videotape. Edit each individual videotape segment. Create graphics and menu pages. Create master videotape which is then transferred to analog laserdisc. Author program which branches to the appropriate video segment. These were the interactive laserdisc (or interactive videodisc) offerings that were widely used in the education, commercial, and retail industries.

Fast forward to today and the stated goals by Netflix to start offering interactive programming with various story branches which viewers can influence / choose. Game engines can be extraordinarily useful as the equivalent of an authoring tool and, certainly, as a preview engine. And, returning to the earlier example of the laborious process of creating content for branched interactive laserdiscs, it will not solely be the case of a human editor creating the various "scenes" that will be used. Increasingly, this content creation will be automated in a templated approach. And for today's graphic, I've reached deep into the Tom O. archives to find not only the t-shirt, but the CD for StoryBase (kudos to Joe DiMare), an early program from Avid that provided story-branching functionality. In other words, what's old is new again…

Want the same opening to a program 50 different ways where the deltas can be identified? Automation to the rescue. Once a template can be applied, AI, Machine Learning, and Automation gives us the scale that is required. Languages, subtitling, close-captioning, and future requirements cannot appreciably add to the OpEx of creating those deliverables. Faster, more precise and less expensive means of creating that content will be necessary. Now, there will be those who read this and think we're back to the future where humans will be left out of the content creation equation. No, that's not where I'm going. As I've written in previous posts, there are different tiers of content and some tiers will require the human element and others will require less of a human element (for example, review and approval of automated content), and still other types of content will be fully automated.

A very timely example? Those 3,000 people that Facebook is hiring to review live or posted videos in light of some of the recent and tragic live streams. Eventually, those hires will no longer be required. AI and Machine Learning will be the guardians… They will spot objects such as handguns, expressions, audio with terms that can cause concern, emotional distress, and so forth.

Those next waves are beginning to and will ultimately have an extremely significant impact on the Digital Media Supply Chain. In the next post, I'll describe #6: A Completely Different Approach.

The August 10, 1989 Beta Build

On the State of Digital Nonlinear Editing
Number 10

This is post number 10 on the State of Digital Nonlinear Editing and Digital Media. The previous nine posts have covered the landscape of content creation, content providers, and service providers. I've been covering the compound annual growth (or decline) rates of various

companies that are significant to the media industry. And, during the period since the inception of these posts, we've seen noteworthy developments that I've covered in the posts. Notably at NAB: Demonstrable interoperating systems utilizing a combination of SMPTE 2022-6 and 2110; SMPTE 2059-PTP; and AES 67.

OTT growth, subscribers cutting the cord, and the impressive NAB participation by most of the Hypergiants: AWS, Google, IBM, and Microsoft. And, most importantly, the types of technologies that are currently being utilized by these types of companies and which will completely redefine how content is made, distributed, and consumed. And that's a critical point, because the innovative technologies that will affect the media industry are being driven—more now than ever before—by the companies that fall into the Hypergiant category.

Let's take a brief stroll down memory lane. In 1989, one megabyte of hard disk based storage cost roughly $15. That certainly was an impediment to the adoption of DNLE systems. Or was it? Part of the pitch to a prospective customer was that technology improvements and ensuing cost reductions made by the computer storage industry would drive down storage costs. And, also, that density would increase while form factor size would decrease. Most importantly, that these improvements would come faster from the technology industry rather than from the manufacturers that were then catering to the broadcast or media industry. That certainly was part of the playbook in those early days at Avid. While picture quality of those early DNLEs was quite poor (250:1 compression) and inferior to the Betamax, VHS, laserdisc-based systems, the pitch was that storage costs would rapidly decrease, compression techniques would improve and the investment was a wise one to make. Of course, rapid market introductions were necessary and, at least in the case of Avid, within two years (1991), storage costs were decreasing and hardware-assisted JPEG doubled picture quality while reducing storage requirements by one-half. And the rest, as they say, is history.

Fast forward to today. Dramatic technological improvements are being driven by the Hypergiants. Five years ago, could we have expected Google to be at NAB with a session entitled: "Media in the Age of Machine Learning in the Cloud"?

In Post #9, I introduced the concept of the six distinct operational characteristics that will impact the digital media supply chain and resulting content creation application functionality. The six are:

1. Craft Editorial.
2. Specialized Content Editorial and Packaging for Platforms.
3. Automated Content Packaging.
4. Artificial Intelligence and Machine Learning Applied to Automated Content Creation.
5. Interactive and Game Engine Applied to Content Creation.
6. A Completely Different Approach.

And via #4, the following areas become applied to audio and video recordings:

Cognitive Metadata Extraction
Tonal Analysis
Visual Recognition
Speech to Text
Language Recognition
Near Human Voice Quality Dubbing

In Post #9, I covered the first five. So, what is a completely different approach? In Post #1, I asked a question—why aren't a billion people editing? And it comes down to the fact that editing is still too hard. I'm not saying that it is hard to figure out. There are many examples where gestural-based interfaces (e.g. pinch, swipe, etc.) have led to very fast learning, training, and adoption (the iPhone being the best example, naturally). But the actual mechanics of editing within the complete workflow are complicated.

Think about what goes into content creation—concept, writing, production execution, post-production across multiple disciplines (editorial, graphics, visual effects, multiple sound departments, etc.), and then there's the creation of the deliverables which can range from one format to 220. And if it's non-scripted? What, in the old days we'd call "reality"—where you could easily have 8 or more cameras shooting for a 24-hour period for multiple days, generating hundreds of hours of footage for a 44-minute show. Those are complicated endeavors—with multi-step workflows.

And a billion people undertaking those workflows? It just will not happen. Too complicated. And while progress is being made to make applications easier to access and easier to use, those applications are being relegated to mobile devices. An example—do a search for "video editing applications for smartphones"—and what you get are a variety of choices—all of which are fine—and there's nothing to criticize about them. And that's the point—they're all essentially the same. Note the way that the functionality is described and the terms that are used— "re-arrange clips, trim clips, add transitions, add titles…". The good news is that providing access to easy-to-use tools that enable people to capture, edit, package, and publish content is becoming ubiquitous—buy a mobile device, buy a very affordable app, capture your content, edit it, and publish / share it.

Simple right? I wouldn't say simple. I would say, simpler. But the issue to be examined is this— those products work in the same way as their predecessors. Clips, trims, layers, and so forth. Same model, different form factor, fewer choices, streamlined. Nothing wrong with that. But, not for a billion people. Not for someone who needs to determine the "IN" point and the "OUT" point. The main issue is that these products are trying to emulate constructs of other editing models. And that's what's actually holding back market growth.

In Post #9, I covered the improvements that are occurring in areas such as AI and Machine Learning and how they pertain to the content supply chain. Object, character, and speech recognition systems are becoming very sophisticated. Case in point: automatic, real-time, multi-language, 95-99% accurate subtitling systems. Only a few years ago, that level of accuracy was

not to be had; today, automatic speech recognition (ASR) has changed the subtitling product landscape (and, not to be dismissed, it has changed the labor requirement as well). And state of the art, today is real-time closed-caption creation in over 80 languages!

Now let's look at some markets that are much, much bigger than the M&E market, which tops out at approximately $49B. Homes, Cars, Consumers, and Augmented Reality (AR). R&D spending in the areas of home automation and AR is significant. Consider Amazon Echo and Alexa; Google Home and Google Assistant. Apple's HomePod and the purported Apple Neural Engine—a dedicated chip / chipset for AI. Autonomous vehicles, customer relations management (CRM). Intelligent, automated systems designed to realize as much automation as possible. And these core technology improvements—object, voice, character recognition—will change the content creation process. We will soon take for granted that the majority of subtitling requirements will be accomplished automatically—we won't even consider how it "used to be done". As AI and data driven experiences in the home permeate our everyday lives, more content will be automatically indexed for us. And image recognition (IR) systems aren't just for media and entertainment—though they certainly have their place in so far as a logo in a video stream could be used as a trigger for dynamic ad insertion. IR systems, as you would expect, have far ranging applicability—from security to retail.

In 2015, the year for which we have the most up-to-date numbers, there were approximately 7.3B people in the world. Of those, roughly 1.25B are illiterate—that is how they are classified versus the remainder who can read and write. *They can read.* And whether reading comes naturally or whether they struggle, more people in the world can read than are trained to edit video. Automatic content indexing assisted by object, facial, character, and speech recognition systems are profoundly going to combine to assist anyone who can read into being able to search for and rearrange dynamic video and audio content. When we think about hundreds of millions of people communicating using video and audio, Marking IN and Marking OUT to create edit points *will not be the model.*

Today, gestural interfaces are accepted and, in some demographics, already ingrained. Swiping up, down, left, right, pinch to resize, grab to reposition, and so forth are common place in mobile devices. Those constructs will be used to provide much more of a direct content manipulation model. But those are incremental changes in my view. Sure, it will be much better, faster and much more intuitive to use those methods rather than whack a slider left or right. So, while interface constructs will change and it will be easier to manipulate sound and images, that will not get us to a fundamental model change for how people will create content.

But people can read and write. And devices that record video and audio are now almost universally available (more mobile devices than toothbrushes on the planet as of 2016). And if you think about how those technological improvements above in terms of object recognition, automatic speech recognition and so forth will combine with gestural manipulation, then there is only one area left to be conquered.

The written word. Because a massive amount of people can read. And if they interact with content not first and foremost via video and audio, *but with words*, manipulation of content becomes really easy and very accessible. And it will / should work along these lines: Content that is recorded will then be processed by a variety of AI application suites. Each suite will provide different functionality (e.g. tonal analysis, speech-to-text, etc.) based on the characteristics of the content. When a live or recorded stream of content is digitized, it will be subjected to a variety of these suites. Very rich, detailed, and comprehensive metadata about that content will result without the large number of humans currently associated with these tasks.

At that point, the user will be presented with the text associated with the content. Each word, with exact reference to its precise positioning within the data stream, will be indexed. Manipulation of text (e.g. cut, copy, paste), will, in effect, be the method of editing that content. Picture and sound will follow along. Via automation, a high percentage of content will already be processed without human intervention (e.g. the earlier example of subtitling). When we have the ability to rearrange text as the primary means of interacting with recorded (read: digitized) content, that is the model by which picture and sound manipulation becomes easy, straightforward, and without any required training.

In the past, these constructs were attempted and some interesting applications were created. However, the difficulty with widespread implementation and adoption was because of the reliance on human transcriptionists.

But imagine the speed at which you can arrange and rearrange text, which would then arrange and rearrange picture and sound—and then you could go in for a trim, not by adding or deleting frames, but by removing words and portions of words. Remember, we are not talking about fine trimming. We are not talking about replacing the terrific craft editorial products that are in the market and that are currently being used for major feature films, television programming, and so forth. Those are products that are very feature rich—deservedly so—and that give you the precision that is needed to adjust frames, sub-frames, and provide intraframe pixel manipulation. You're editing a feature film comedy? As Chuck Jones once said, "The difference between a laugh and no laugh can be as little as one frame." Obviously, you need features that provide that type of manipulation.

However, to massively expand the marketplace, context-aware, textual-based, editing constructs combined with AI and Machine Learning technologies will make content creation and delivery extremely accessible. The creation of content for social network platforms alone requires that a fundamental change in how that content will be put together.

It is this confluence of technology and a recognition that the current methodologies for content manipulation, editorial, and delivery must change in order to be able to handle the enormous amount of content that is growing daily. Let's put it this way—Facebook Live is only one indicator of how popular live streaming and non-live VOD usage is growing. How is that content searched? How is it curated? How do you get post-live broadcast value from it? How do you

correlate what is said in the video and how do you search for certain words across thousands of hours? And those same technologies that will be utilized in security, home automation, CRM, and so forth, will have enormous applicability to content creation.

Now, I realize that there are going to be skeptics regarding this post. I think that's natural. But if you've been following previous posts, these technologies are being developed by companies that have enormous total available market revenue potential. In other words, technologies that will be combined to create these new content creation methods must be developed in order for those large companies to continue to grow. Content creation will benefit because of that R&D spend.

And that's the sixth operational characteristic that will impact the digital media supply chain and resulting content creation application functionality—a completely different approach. And that's how we are going to get to hundreds of millions (and a billion) people editing.

More on the media industry and supply chain in the next post.

Automation, and the Economy
Executive Office of the President

December 2016

On the State of Digital Nonlinear Editing
Number 11

This is post number 11 on the State of Digital Nonlinear Editing and Digital Media. In Posts 9 and 10, I covered the six distinct operational characteristics that will impact the digital media supply chain and resulting content creation application functionality. To review, they are:

1. Craft Editorial.
2. Specialized Content Editorial and Packaging for Platforms.
3. Automated Content Packaging.
4. Artificial Intelligence and Machine Learning Applied to Automated Content Creation.
5. Interactive and Game Engine Applied to Content Creation.
6. A Completely Different Approach.

We then explored the significant technology advancements being made under #4 and which will be applied to audio and video recordings:

Cognitive Metadata Extraction
Tonal Analysis
Visual Recognition
Speech to Text
Language Recognition
Near Human Voice Quality Dubbing

The constant of these posts has been: "What will it take for a billion people to be editing?" But, really, when I ask that question, think about replacing the word "editing" with "communicating"—and communicating with context-aware video and audio. And, apart from the ability to empower individuals to create, the amount of growth in the use of video—not only for entertainment, but for so many other industries—requires that automation and technology be applied.

It is no longer a point of conjecture that Automation, AI, and Machine Learning will be applied to content creation. There are various examples of developments being driven by research institutions as well as the Hypergiants that I have discussed in prior posts.

I will be covering many of these issues during my paper presentation at this year's SMPTE Technical Conference. Among them:

Computational Video Editing for Dialogue-Driven Scenes
Real-time Speech Recognition
Automatic Editing of Footage from Multiple Cameras
Learning Lip Sync from Audio

And that is only a start. Just one of those developments, in and of itself, would be interesting and academic, but as per the above sampling, there are multiple initiatives underway. Further, the amount of money, time, and research that are being applied to create breakthroughs from the status quo of content creation are considerable. The dependencies on these innovations are coming from industries with far larger potential total available markets than anything comparable in the M&E market.

After viewing the research that is being done, we can clearly see that the content creation process is on the verge of undergoing amazing changes. And, therefore, an entire new set of applications that will be used for creating content is not only in the pipeline, but will eventually find their way into the marketplace.

When we look at the types of research being done via those examples, let's for a moment, not think about the applications for the media and entertainment industry. Automatic creation of content for the financial services, education, state / local / federal sectors?

Daily public meetings for state agencies which are covered live for the public and which must also be made available for video on demand. Add three voice—and motion-activated robotic cameras. Then, the ability to take the output of those cameras and automatically create the required VOD assets based on specific selections related to keywords. All the coverage, automatically edited of legislators and public speakers as they discuss "Bill #2112. Legislation relating to…"

And that is just one application. YouTube is citing 1B hours of video viewed each day. Facebook is citing 100M hours of video viewed each day divided among 500M people. And each day the number of hours and the number of viewers have consistently been growing. But the intelligence surrounding that content and providing context to that content—beyond simple metadata tagging—requires the type of automation that results from these considerable R&D efforts.

Crowd-sourced content from fans of car racing, music concerts, college and high school events? Today, what typically happens, post event? Usually, the content is uploaded by individuals to a social network or a private network (e.g. the high school's website). And, at that point, you search on the social network platform and see possible matches (as long as the metadata that you used for the search criteria has some form of match). Unfortunately, what mostly occurs is that you see the same section of material covered by different cameras and from very similar angles. In other words, there's no narrative and no beginning, middle, and end.

And so, we ask, why isn't all this crowd-sourced material automatically sequenced, edited, and presented with all the alternative angles—which the viewer can choose—resulting in more content being searchable? And, by extension, adding more value to content will have an impact on the value that can be derived from the content in some form of currency—whether outright purchase, subscription, or ad revenue. The issue, of course, as I've been outlining in previous posts, is that the total cost associated with preparing content so that it is ready for consumption (and not free) has to be reflective of the content itself. A feature film with a $100M budget has historically had an associated 10%-20% cost associated with the post-production and preparation of that content. But the cost of preparing content has to dramatically scale downward when both budget numbers and unknown audience numbers are present.

Applying automation and the technologies that are outlined above will continue to form solutions to bringing more video to all types of industries.

We are experiencing unprecedented acceleration in the advancement of technologies that are designed to increase the automating of manual tasks. In December 2016, the Executive Office of the President of the United States published a report entitled, "Artificial Intelligence, Automation, and the Economy". It was signed by individuals representing the Council of Economic Advisors, Office of Science and Technology, and Domestic Policy. It is a detailed report on the effects of technical innovation—in this case AI and Automation—on the economy, jobs, and other factors.

Among the highlights (I'm paraphrasing): "A 2015 study of robot usage in 17 countries added an average of 0.4 percentage points to annual GDP growth between 1993 and 2007. That accounted for one-tenth of overall GDP for those countries. Technology implementations increase productivity as a result of decreasing the number of labor hours required to create a unit of output."

And this: "The current wave of progress for AI… driven by three mutually reinforcing factors: the availability of big data from sources including e-commerce, businesses, social media, science, and government; which provided raw material for dramatically improved machine learning approaches and algorithms; which in turn relied on the capabilities of more powerful computers."

And, finally: "During this period, the pace of improvement surprised AI experts. For example, on a popular image recognition challenge that has a 5% human error rate according to one error measure, the best AI result improved from a 26 percent error rate in 2011 to 3.5 percent in 2015. This progress may enable a range of workplace tasks that require image understanding to be automated, and will also enable new types of work and jobs. Progress on other AI challenges will drive similar economic changes."

I think, it's fair to say that "may enable" is now beyond a shadow of a doubt. And, when we step back and look at all this progress—we can see the positive applications for content creation. This is why it is inevitable that what I posited in Post #10 will become a mainstream methodology of video and audio content creation and consumption for the masses. Now, again, I realize that there may be some skepticism here, but consider:

If content from multiple sourced consumer cameras can be automatically parsed by facial, object, content, and audio recognition systems… If content can then be analyzed for contextual and syntactic intents and meanings… If content can then be manufactured (with later human tweaking and approval when necessary and appropriate) … Then we have all the means by which we can truly scale not only how content is created but how people communicate.

And these "ifs" are now rapidly becoming reality. We're seeing the research, we're seeing the prototype applications, and we will eventually see the commercially available applications. And it will be very, very exciting as these products roll out.

And from Post #10, I wrote about how words and sentences—and manipulating those words and sentences—will dramatically expand opportunities for people to not only interact and rearrange content but to significantly change how they use video and audio. Certainly, professionally created content, workflows, and the artists who create that content will greatly benefit from these technological introductions. Whether the content is destined for theatrical, broadcast, or OTT, content creators will benefit.

But it is for the non-theatrical, non-broadcast, non-OTT (and so on!) types of content where these benefits will represent breakthroughs in communication. All that content—as live streams or as recordings of any type of event—will be able to be manipulated by individuals who are not classically trained at video and audio editing.

Consider: As best as I can ascertain, there are a minimum of 750M Microsoft Office users worldwide. There are at least 800M Google Drive users worldwide, but the number of Google Docs users is unclear. Is it 10% of the Drive users? 30%? Any way we slice it, that's a lot of people who are using applications for document creation and editing.

And when that number of users begins to manipulate video and audio with the same ease with which they manipulate words, that is the moment when a billion people are able to create and interact with content in ways that they never before could.