

# Document Imaging Report

Business Trends on Converting Paper Processes to Electronic Format

## Table Of Contents

AI and the Pace of Change	2
ibml Is Now THE Company for Ultra High-Volume Scanning	3
Security Lessons From MFPs – What Can Scanner Manufacturers Learn?	7
AI Regulations, Ethics, and the Delicate Balancing Act Ahead: A Sneak Peek Interview With Andrew Pery	11
Two Question Tuesdays	16
PDF Turns 30	17
AIIM Conference Snapshot via Ralph Gammon	17
News and In Brief	18

# AI and the Pace of Change

We're near the half-way point of 2023 already. With the pace of change and deluge of news these days, it feels like the world is speeding up.

Part of the reason, as least for me, is the daily barrage of generative AI news. This technology is having its moment in the spotlight – and we'll begin to see those tools wrapped into IDP products more frequently.

That's easy enough to see, even if it's impossible to predict how this will all shake out in reality.

One thing that does strike me, as always, when the breathless panting about new tools begins to reach a crescendo: how do we best use these new technologies and techniques to improve our customers' lives? I've always thought that what was once the ECM industry, including the IDP piece, has always done a too-thorough job of bouncing from new thing to new thing while there are still people discovering daily and are amazed that the ability to digitize a piece of paper exists.

Dan Lucarini, long-time IDP vendor, marketing executive, and now analyst, wrote an interesting post; [Documents, Content, Files, Records, Semi-Structured, or Unstructured Data: Do Labels Matter Anymore?](#) Dan points out that the IDP industry never settled on a single word to describe what the software manages (this always bugged me too).

At any rate, he makes two points worth sharing:

1. AI almost certainly renders the "what do we call this stuff" moot. AI doesn't care.
2. Figure out what your customer calls their content/documents/files/forms/etc. and use that when talking to them (yes, it's possible to modify those sales PowerPoints!)

While these are exciting – even revolutionary changes – let's not get so caught up in the hype that we confuse the audience for IDP.

Thanks for reading,

Bryant Duhon  
Editor-in-Chief, Document Imaging Report  
bdu@info-source.com

Comments, criticisms, and witticisms welcomed.



# ibml Is Now THE Company for Ultra High-Volume Scanning

On June 9, 2023, **ibml** announced the acquisition of **Exela Technologies'** IntelliScan Smart Scanning Solutions. ibml will acquire the assets of the high-volume scanner and North America hardware maintenance businesses of Exela Technologies.

We had the chance to quickly catch up with [Susheel John](#), ibml's CMO. Here's what he had to say about the news.



**Let's start with an easy one, what's this mean?**

**SJ:** We were already leaders in the ultra high-volume scanning market itself. With the acquisition of Exela Intelliscan Smart Scanning Solutions business (formerly BancTec Intelliscan), we are now definitely the dominant player here in the US and Europe.

BancTec has a good base in the Asia Pacific region. This now allows us more paths and partners into the Asia Pacific region.

It also helps us consolidate our end of the market and serve customers better, consolidation makes sense and allows us to combine the resources of both companies to make better decisions in terms of investment – both for products as well as the world-class service we've always provided.

We think we'll be able to improve our portfolio and service while allowing sustained investment in innovation.

**The first question in any acquisition is what's the plan for product consolidation, which you alluded to already. So what's the plan for the Intelliscan line?**

**SJ:** No decision has been made to consolidate products. All that I can say for now is that we are definitely going to continue to carry both the product ranges. The future is TBD.

Keep in mind that both product sets have some unique capabilities that do not overlap. As you know, these products are customized for each customer. We can't afford to turn off any of these unique capabilities because customers depend on them for mission-critical processes.

One important aspect of this for ibml is that the Intelliscan Raptor is a net addition to our portfolio, positioned below the ibml Fusion. We've not had a product in that category. This gives us a nice product for us to compete for high-volume scanner customers – the segment of the market below where ibml has historically competed with our lineup. We'll be a more effective competitor to Opex, Innotec, or Kodak high volume scanners offering our customers better choice.

**You mentioned the unique capabilities of each product line, what does this acquisition bring to you that you didn't have before?**

SJ: BancTec lines have the ability to do vertical pocketing. That decreases the length of the scanner while providing more capacity. For larger capacity, ibml Fusion scanners can go up to 24 pockets. Most of our customers have high volume capacity pocket requirements. However, in the future if a customer needs only a few pockets or only collect 100 to 200 documents in a tray, then those vertical pockets are a good solution.

There are certain parts of the world where you need to print on the reverse side of the document. We currently don't do that. We only print on the top side of the document. That's a unique capability BancTec brings to our portfolio.

Their software is different from our software, not necessarily to say one is better or the other. There are some unique value propositions of ibml software because we are network-based software, we can manage fleet of scanners while Intelliscan USC software is a simpler application that does flat file output. There are subtle nuances for scanning needs when one could be more useful for a company than the other.

As ibml continues to invest in both hardware and software, over time there may be synergies in these products that we can bring together. There are so many nuances in how both product lines address customer needs that it's simply too early to predict.

**From a market sector, I'll say domination, standpoint, how's the "huge honking scanner" segment now look?**

SJ: We already had 80% market share for mailrooms. There are three, well, now two, manufacturers in this segment: us, Intelliscan, and Primagest. Primagest originally licensed from ibml and manufactured in Japan. The licensing agreement is over, and they exist as a separate brand by themselves in Japan. They sell mostly in Japan and Korea.

**You mentioned the bigger footprint in Asia, but is there potential conflict between the ibml and Intelliscan lines in the US/Europe?**

SJ: Not really. It's more additional for sure. We have almost doubled our install base with this acquisition. We also will continue the partnership with Exela, which is a large BPO obviously already using Intelliscan and ibml products as part of their customer solutions.

As part of this acquisition, we only acquired their hardware, software, and maintenance services business. We will continue to partner with Exela to offer their customers world-class high-volume digitization solutions..

**So you consolidated your hold on the market segment and bought into a BPO customer base?**

**SJ:** BPOs are roughly one-third of our customer base (banking and government are the other two “thirds”). Exela has a global presence. Their Asian presence, especially, allows us access into a market we haven’t heavily focused on. Also, the Intelliscan business has existing partners within Asia, which could potentially extended market access for ibml.

**Can you share any information on install numbers and manufacturing direction?**

**SJ:** Too soon to say with manufacturing. Both product lines are manufactured in the US; Birmingham for ibml, Dallas for Intelliscan. Regarding the install base, you need to remember that BancTec Intelliscan was primarily a check reader-sorter company before they got into full page scanning. There’s a long history of scanners in the field. We aren’t just supporting the ultra-high-speed scanners now.

**While there will be an ongoing need for high-speed scanning for a good while, eventually (and theoretically), this market will shrink. Over time, do you see a shift to a larger focus on software as your business?**

**SJ:** First, our hardware business isn’t under any kind of threat, especially with this acquisition now and doubling the install base itself. There are plenty of legacy products out there that need to be replaced. Regarding the Intelliscan product line, they haven’t launched new products for some time now and those customers are looking for new technologies.

That said, let’s tackle your question. We recognize generally from a trend perspective document volumes in business are reducing, but we can all agree that paper is long from gone.

**Yes, I can recall a poster in the Inform Magazine editor’s cube from February of 1995 that said something like “The Paperless Office Is Here!”**

**SJ:** Yes, well, I’ve said and sold that idea too. Today, there are specific initiatives helping our business, like the NARA initiative here in the US to convert everything to digital. There are similar initiatives around the globe.

For digital transformation to take place completely digitization needs to happen, there’s still a lot of backlogs and back file conversions to get through. Also, most organizations still have a multi-generational customer base – not everyone is comfortable with digital.

There are still 15 billion checks processed in the US every year as an example. So paper is not gone. Concern about a disappearing market isn’t why we made this move.

We will also continue to invest in our software. It continues to be one of the strongest products in the market for batch processing applications and for structured forms processing, best-in class in terms of high-volume mission critical capture applications. We also have OEM partnerships with other software technology vendors to offer extended unstructured forms processing.. That's an intentional decision to work with partner products today. It helps us wet our feet in the market build a go-to-market first and then think about more investments in that area.

**You're really locked as the leader for any application that needs super-fast scanning. As FADGI trickles down, you're in a great spot to grow.**

**SJ:** I think we've been in advance of the market on our products here and have a good strategy. I think we have the best FADGI scanning solution on the market. These projects will need image quality, high speed, and document handling. That's where we excel.

For FADGI, our solution is essentially a self-contained part of the regular production process.

**What about current leadership – and you're hiring a new salesperson, correct?**

**SJ:** Current IBML management continues: Martin Birch, CEO; Russell Smith, CFO; Pete Rudak, CTO; and me as chief marketing officer.

We will announce our new CRO soon, which is an investment in our ongoing go-to-market focus. Once we have an even better strategy, focus & execution there, I think we can sell anything. Products and technologies come and go, but the ability to understand and serve customers by continuing to solve different problems with technologies is the foundation of everything that we are building.

Roughly 100 BancTec employees will move over to ibml – these are people in field service, engineering, and manufacturing mostly; also three people in sales.

**I jokingly commented that as a marketing guy, Susheel didn't want to do sales anyway. He quickly reminded me I had forgotten his previous role!**

**SJ:** Absolutely wrong, I did sales before marketing. In fact, I have additional responsibilities for sales in APAC as well as the sales arm of our IDP business.

**My mistake! So will the new CRO take those responsibilities off your plate?**

**SJ:** I don't know the answer right now, but I think that would be the ideal outcome for us. To begin, as the new CRO onboards, he'll focus on our core business. Also the reason I am handling APAC now is because of my prior experience there.

My last note was to check in about your ibm! Cloud Capture efforts. As we discussed, paper will be around for a while, but I think it's equally important to point out that you are dipping your toes into selling software.

SJ: We're still trying to make inroads. We do have a pipeline building up and a few interesting projects in various stages, but it's still very much a work in progress.

## Security Lessons From MFPs – What Can Scanner Manufacturers Learn?

**As MFPs have become more sophisticated, it's become essential to include them in a company's security strategy.**

Article contributed by Oda Shinichiro. [Click here to connect with him on LinkedIn.](#)

Looking at MFPs from the perspective of preventing information leaks, from a functional and operational standpoint, we know that there are significant risks associated with the operation of these devices. If the MFP is accessible from the Internet, it could be a target for hackers.

Even MFPs in a LAN protected by a firewall can have information that's been copied or scanned surreptitiously extracted if the appropriate network security isn't implemented.

To emphasize this point, Fig. 1 below shows the results of a search for the keyword "MFP" on SHODAN, a hacker-friendly website that allows users to search for devices on the Internet. It shows how many MFPs are exposed to the public Internet. One can imagine that if the administrator password is left as the default, it could be easy to view or download the scanned or copied images from the history information.

In addition, MFPs usually have a built-in Apache-based Web server that allows the user to remotely control the MFP, so that if the user ID and password are known, the data in the MFP can be viewed from anywhere in the world.

[Editor's note: HP created a series of videos starring Christian Slater as "the Wolf" that did a fantastic job of spotlighting printer/copier security holes. Here's a link to the first one: <https://www.youtube.com/watch?v=2HQ4kwCm8U4>.]

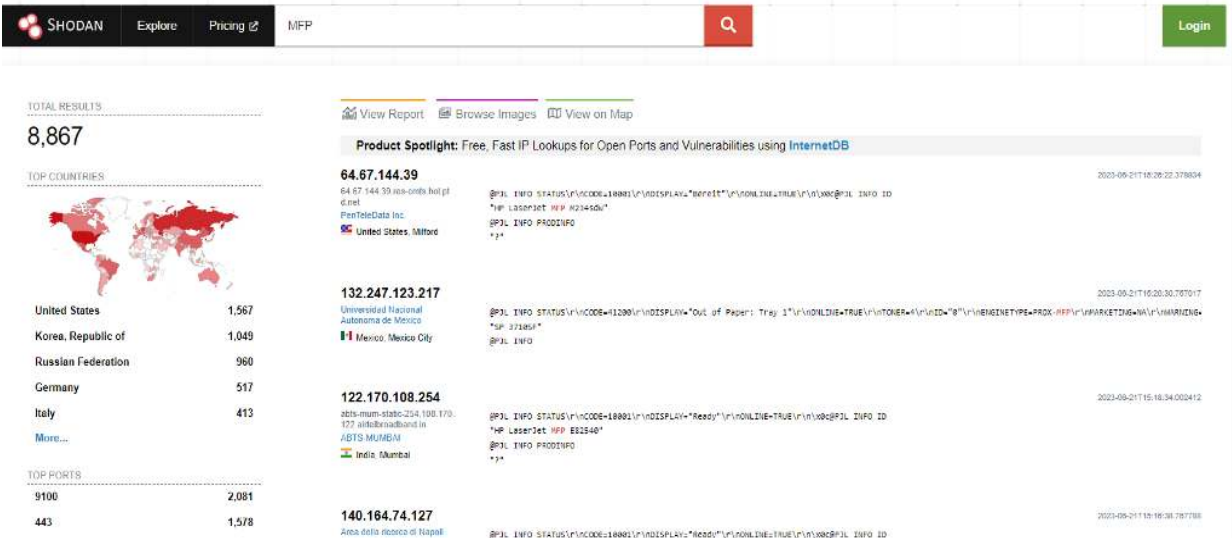
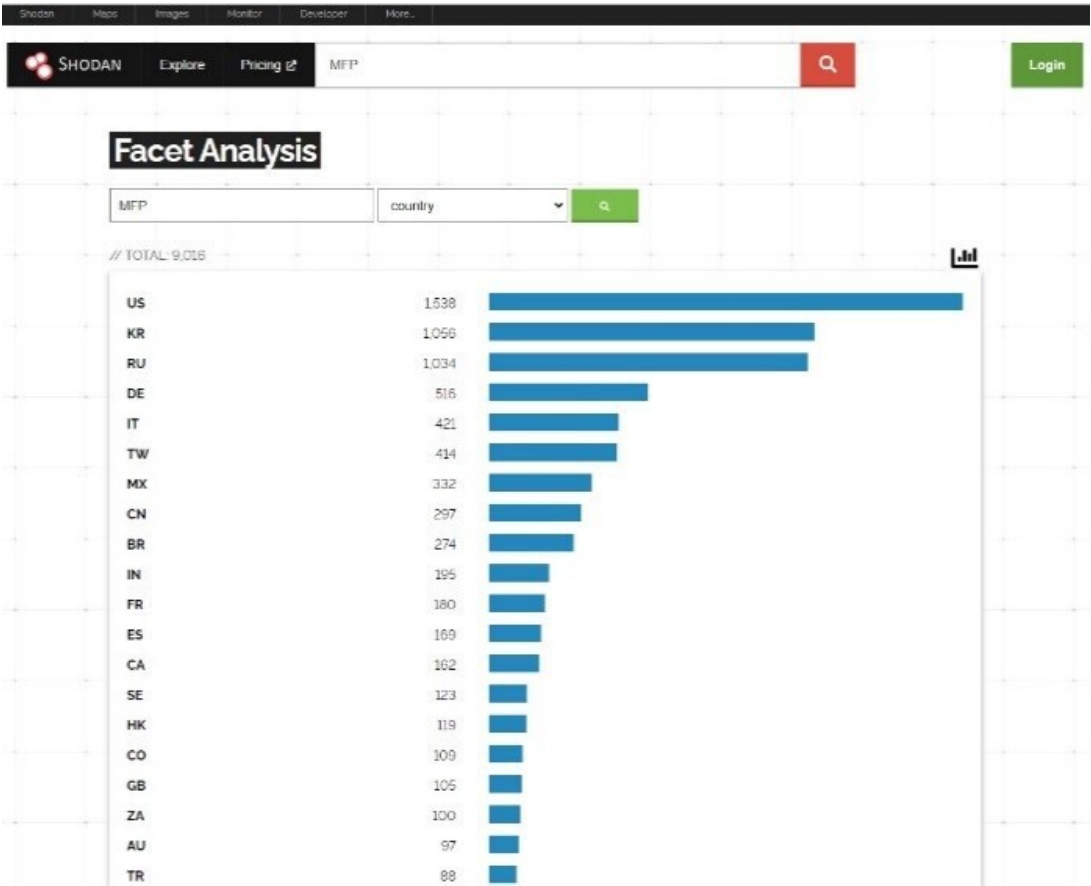


Fig. 2 by country shows that the U.S. has the overwhelming number of addresses directly connected to the Internet, followed by South Korea and Russia.

This situation is unlikely to be the result of an intentional act by the network administrator of an organization, and more likely to be the result of universities and research institutes that have been allocated an abundance of global addresses.





Although MFP vendors incorporate a variety of measures and features to prevent information leaks, information leaks cannot be prevented unless the MFP is installed, configured, and managed by the organization's network administrator using appropriate built-in design features.

The security features implemented by most MFPs today include:

- Restriction of operable functions by administrator password
- Restrict operable functions and protect documents by user authentication
- Setting access rights/restricting and controlling web server functions
- Restricting/controlling server functions
- Restrictions through IP filtering
- Firmware updates
- Other security features (e.g., *preventing fax misdirection, setting passwords for scanned documents, encrypting hard disks and other storage devices*)

These implement the same or better functionality of general network work servers and are considered as security measures from both technical and human aspects, such as hacking, spoofing, eavesdropping, mishandling, and forgetting, but proper configuration and thorough organizational operation rules are essential.

### Will Scanners Require the Same Security Measures?

MFPs have evolved as office equipment devices that provide document solutions, such as duplicating various documents generated in office work, sending/receiving faxes and printing out, network printer functions, copy functions, scanning and other functions that store input/output images internally and send them to designate locations or print. As information is held within the 'machine' functions/ applications to protect stored information from 'leakage' have also evolved in the form of security functions and operational design.

Standalone document scanners, on the other hand, have internal buffer memories which enable continuous scanning and communication and are erased as soon as the job is finished. Since there is no need to retrieve and print images asynchronously as with an MFP, the security features required for document scanners are not as sophisticated as those of MFPs.

Network scanners, which operate without a PC, tend to have limited security features such as user authentication functions (ID, password, IC card) and separation of data and control signals. The reason is that document scanners specialize in digitizing documents, and scanned images are processed by dedicated software running on a separate server. In PFU's example, PaperStream NX Manager is responsible for configuring, managing, and executing JOBS. For Kodak Alaris, INFUSE Management Software and other software performs these tasks. Since these software products are run on a server on a corporate LAN security, measures are therefore implemented by the network administration of the organization. The scanner itself only needs to be able to guarantee user management and secure communication.

In cases where scanned image files are transferred to another system using RPA, etc., the RPA robot itself needs to be configured and the security management of the PC on which the robot is running needs to be well protected.

### Conclusion

As the paperless society progresses, the volume of paper documents themselves will decrease. MFPs will be used more often as scanners. The security challenges described above will remain and will need to be continuously updated.

Scanners, on the other hand, will likely continue to improve in function and performance in the direction of greater convenience as input end-devices. For example, when scanned images are transferred to a designated server or a folder in the cloud – or encrypted and sent as an attachment to a designated e-mail address – security functions that can be set only by the administrator will be necessary.

However, MFP-like security features may increase the cost of the scanner and the time required for its operation. Therefore, Infosource believes that the functionality level of network scanners currently on the market is sufficient from a security perspective, and we do not believe that the market will demand more functionality than that and the security measures would be better handled on the RPA robot and management server side, which would allow for a freer choice of scanners and preserve economy and convenience.

## Register Now!

Learn more about the upcoming Capture & IDP  
Conference on Sept 6 & 7.

We look forward to seeing you there!

[Learn more](#)

# AI Regulations, Ethics, and the Delicate Balancing Act Ahead: A Sneak Peak Interview With Andrew Pery

The proliferation of AI technologies is reaching a critical inflection point. Goldman Sachs suggests that by 2030 up to 18% of the global workforce, representing some 300 million people could be displaced by advanced technologies, such as conversational AI. On a more positive note, the World Economic Forum projects that AI will add 58 million net new jobs and will contribute \$ 15.7 trillion to the global GDP. While AI is still in a nascent stage, yet it is way too big and potentially dangerous to be left for self-regulation.

Our second Capture & IDP Conference Sneak Peek, long-time IDP industry expert Andrew Pery, Ethical AI Evangelist, raises a range of ethical concerns around artificial intelligence that organizations need to be thinking about NOW! From training bias to the potential societal disruptions from job loss (and the need to train for a new category of careers), Pery pinpoints the issues you need to understand today to set a foundation for an AI-enabled tomorrow.

He'll be exploring how to navigate a complex AI regulatory framework in his session at the upcoming 2023 Capture & IDP Conference on September 6&7.

Click here for more conference details. <https://www.info-source.com/home-capture>

You can connect directly with [Andrew here on LinkedIn](#).



Here's what we asked him:

- Can you explain what AI is – and isn't?
- What are you most concerned about regarding AI?
- Most optimistic about?
- Is AI a job creator or destroyer? Not to be dismissive, but there are always “buggy whip makers” who go out of business as technology advances.
- As someone who's made a career out of doing things with words, what are the copyright implications (and other ethical concerns) of generative AI tools like ChatGPT?
- In your opinion, what's the most important ethical concern with AI?
- You're going to be talking about the looming regulatory environment for AI. Give us a snippet of what that could look like.

**Let's start off with a basic question; can you explain what AI is and what it isn't?**

**AP:** At the risk of oversimplification, AI is essentially an algorithm or a smart machine that's designed to perform specific tasks that kind of mimic human intelligence.

I want to qualify that statement because there are basically 2 fundamental classifications of AI systems. The first is “narrow AI.” These are AI applications that are designed to perform very specific tasks. For OCR technology and intelligent document processing, they use machine learning technology, like supervised machine learning that's based on training data or convolutional neural nets for image classification and extraction. These technologies have been around for a while, and they perform these tasks much better and much more accurately than humans can with a high level of recognition, extraction, and classification rates.

The other class of AI systems is defined as general AI. An example of this is a large language model such as ChatGPT. There's a lot of buzz and conversation around these generative AI applications that come much closer to human cognitive capabilities.

They use much more advanced deep neural networks and analytical capabilities based on analysis of large volumes of text and then use natural language to produce output which sounds like a human conversation. At times they produce very accurate results, at other times, questionable results. There's potential risk of harm and disinformation. There are discussions about how close generative AI comes to human cognitive intelligence.

We've seen just recently some of the subject matter experts like Jeffrey Hinton from Google have a dystopian view of the potential impact of ChatGPT, calling for a moratorium for any further development. He's predicting that without significant guardrails around conversational AI, it can potentially be an existential threat to humankind.

There is an evolution of AI into the realm of general AI. I think we have to keep in mind that even general AI applications like GPT don't really have human reasoning capabilities or human cognitive capabilities or compassion or analysis capabilities.

I think of it this way: the human brain consists of 100 billion neurons so it's going to be a while before general AI applications will supplant, if you will, human cognitive intelligence. But there is this underlying skepticism and concern. This is why you're seeing significant momentum towards regulating AI; providing for some guardrails to ensure that AI applications are used for things that have real socioeconomic utility.

**As you were talking about the lack of true cognition for ChatGPT, I couldn't help but think how they've begun putting up a notice that it only includes information through X date. That wasn't there a few months ago (to the best of my recollection).**

**Carrying on the dystopian theme, what are you most concerned about regarding artificial intelligence?**

**AP:** There are many different dimensions to that question. First of all, it has a tendency to amplify bias because it's based on training data upon which the algorithm delivers its output. As David Weinberger of Harvard University (one of the leading commentators of AI technology – and knowledge management as well) put it, “bias is machine learning's original sin.”

There are different types of bias, historical bias, aggregation bias, confirmation bias. And so it's important to ensure that effective data governance best practices are put in place to ensure that the data quality provide accurate results.

Secondly, AI itself is somewhat opaque, right? It's hard to understand exactly how they work and the outputs that they generate, sometimes even the developers themselves – ChatGPT is one example of this.

The explainability of AI systems will become important to reassure consumers and to provide a level of transparency to determine how the outcomes were produced and to mitigate potential harmful impacts to create consumer trust.

The second is the issue of enforceability – how do consumers demonstrate the risks of harmful AI outputs. I think it's important for legislators and the AI community to ensure that it's easier for consumers to demonstrate potential harm and to shift the onus from consumers to the developers. There's a presumption there'd be a presumption of harm and would be a rebuttable presumption.

So inherent bias; opaqueness and explainability; and mitigating harmful impacts from application of the technology.

**So let's flip that question around, what are you, what are you optimistic about?**

**AP:** First, I should note that AI applications are becoming pervasive and is impacting literally all dimensions of our lives: IoT devices, driverless cars, making people more efficient by reducing repetitive, labor-intensive tasks. There's an obvious need to harmonize AI regulation so that we develop specific standards before introducing AI technology to the marketplace. There needs to be legal certainty that balances innovation between AI technology and preserving and safeguarding consumer rights.

We're seeing momentum in the marketplace towards this regulatory framework. The European Union has proposed the EU Artificial Intelligence Act. After several amendments, a compromise text has been voted on by the European Parliament voted on. There's an expectation that by the end of this summer there will be a first sort of comprehensive regulation. Vendors themselves are also implementing rigorous voluntary codes of conduct. As disruptive as AI is, it makes good business sense to create customer trust.

There's also a lot of positive social utility of AI. It could potentially emancipate people from tedious, labor-intensive tasks. In healthcare, AI could help introduce drugs into the marketplace faster. There are significant opportunities for us to improve the quality of life on the one hand.

The real question is will we be able to control the more nefarious potential uses of AI and also to limit the kinds of applications of AI that may create disparate treatment of or negative impact on consumers.

So we're still looking a visions of a dystopian Skynet and AI overloads and the true intelligence of Asimov's robots in *i, Robot*. And, as usual with technology, we'll most likely end up somewhere in the middle.

**Do you think AI is a job creator or destroyer? Not to be dismissive, there are always "buggy whip makers" whose businesses are overtaken by technology advances.**

**AP:** I think this is the \$64 million question. There is no doubt that in the short term there could be significant adverse impacts with the introduction of AI technology. A few data points on this. Goldman Sachs has done analysis on this. They project that virtually two-thirds of jobs in the marketplace today could be somewhat adversely impacted by the introduction of AI technology. That translates to some 300 million jobs.

They don't say that all of these jobs will be fully impacted, but there's some concern about that. There's sort of a presumption that the introduction of destructive technologies represents a zero-sum game: jobs go away but aren't replaced with new jobs. On the other side, the World Economic Forum projects that by the end of 2026, , AI technology will create something in the nature of 95 million net new jobs.

We have to think in terms of looking at the history as well. We've seen these disruptive events occur with the Industrial Revolution, with the introduction of computers. There's always that level of, I would say healthy skepticism that Oh my God, the jobs will be taken away. I think the key thing here is no question in the short run there will be negative impacts.

I think we need a societal effort in retraining and upskilling. That's going to be important and there has to be an investment in that regard. We also have to think in terms of the jobs of the future. Somebody will have to monitor these AI systems, do QA, etc. There's real potential for new careers too.

In the writing community within the entertainment industry, there's concern about how ChatGPT could supplant original works. I think we'll see a proliferation of legal challenges and issues arising here.

**I know the marketing community is in a near state of panic (some intentional fearmongering as well) over ChatGPT. That's a nice segue into ethics. What is the most important ethical concern with AI?**

**AP:** That's the crux of the issue with generative AI. There are two dimensions here: the ethical component and responsible AI.

Ethics is from the Greek word "ethos." That's a normative value of human dignity, autonomy, justice, equity, etc. On the other hand, responsible AI is a set of tools that enable organizations to ensure that it produces equitable results. Progress has been made towards responsible AI in terms of transparency, accountability, explainability, confidentiality, and security.

Significant challenges remain. How do you ensure that even though the AI algorithm produces accurate results that the results are equitable? Is there a need to have some sort of a corrective component to AI technology, notwithstanding the accurate statistical results? For example, in lending, financial institutions typically use the FICO score to make a determination for eligibility. Many marginalized people have never applied for a loan before and don't have a FICO credit score, so the loan could be denied. One of the aspects of ethical AI is to make sure that there's human agency/oversight to look at the results to produce desirable outcomes for society, such as making credit available to a larger constituency. It applies to other areas as well; employment, criminal justice system, etc. I think we're going to see significant challenges and conversations as to how to make sure there's a balance between social justice and social equity, making AI available to a larger constituency of stakeholders.

This has to be balanced against AI that produces efficiency and economic value to businesses. That's going to be the \$64 million question: who has responsibility? What is ethical? And from whose perspective are we trying to enforce or impose ethical principles since that can be contextual.

It's also a governmental issue: What policies can be put in place to ensure that there is continuous monitoring of how AI impacts society as a whole? I think this is where we're going to see massive transformational initiatives arise. Hopefully, at the end of the day, the net impact will be positive.

**As you were talking about the brute force aspect of AI, to be able to deliver the efficiency and all the good things that it does, reminds me of what you said at the beginning of our conversation, that there's no cognition, there's no judgment on the part of AI. It's figuring out how to use AI to help make a decision as opposed to AI always making the decision, which could be unequal and not serving broader societal needs.**

**AP:** That's right. One of the key dimensions of the EU Artificial Intelligence Act, apart from other things in terms of transparency and accountability and risk management, is human oversight. What that means is that there needs to be an obligation on the providers of the AI system to analyze the outputs and to make sure that they have oversight and the ability to override decisions made by an AI system. Human judgement is needed to reflect on the potential harmful impacts, notwithstanding the accuracy of the algorithm.

I think this is the balancing act that we'll have to be aware of so we can recognize both the economic and societal value of AI.

**Let's wrap with a little preview of what you'll cover at the conference.**

**AP:** It's such an evolving process [regulations]. I mentioned the upcoming vote on the EU Artificial Intelligence Act. There are ongoing issues as AI regulation attempts to create and impose a risk-based framework for AI evaluation. This classification ranges from prohibited AI (social scoring) to high-risk (privacy rights) down to low risk. They're proposing a rigorous mechanism where even before you put an AI system on the market you have to go through rigorous conformance testing to make sure that an AI vendor has quality processes in place so that the AI system behaves in accordance with expectations and the desired outcomes. Once that's been certified then you have a continuous process of monitoring, transparency, accountability, human agency oversight of the AI system.

Much like with GDPR, I expect other jurisdictions to follow. The US is more sectional, and Canada has proposed legislation in place. Over the next 12 to 18, I think we'll see a more comprehensive regulative framework emerging that creates a heightened level of obligations on providers of AI systems. Generative AI has definitely raised awareness of the potentially broad and pervasive implications of AI.

## Two Question Tuesdays

If you've not seen our new YouTube series, Two Question Tuesday, head over to our YouTube channel and subscribe (or follow the Infosource LinkedIn page – we publish them in both places).

Here are the latest (click on the title to go to YouTube):

- [The Hottest Market for IDP](#), with Bob Fresneda, TCG Process
- [Hyperautomation and Digital Transformation](#), With Jason Burian, KnowledgeLake
- [InfoGov, Lip Service, and the Problem With AI](#), with Steve Weissman, Holly Group
- [Network Scanning Is Here to Save You Time and Money](#), with Kevin Neal, P3iD

Would you like to participate? Contact Bryant Duhon; [bdu@info-source.com](mailto:bdu@info-source.com), with ideas and suggestions.



## PDF Turns 30

I'm just young enough to take PDF for granted, but it really has made sharing information easier for the past 3 decades.

Duff Johnson, CEO of the PDF Association, wrote a short post on the anniversary. Here's a snippet:

From the vantage-point of 2023 we are positioned to recognize 1993 as a year of two key developments; the first specification of HTML, the language of the web, and the first specification of PDF, the language of documents. Today, both technologies predominate in their respective use cases. They coexist because they meet deeply related but distinct needs.

Thirty years after it was popularized, the World Wide Web has revolutionized the human experience of information. Meanwhile, PDF has replaced paper documents with digital analogues. Some expected the web to replace digital documents as well, but all indications are that [PDF continues to grow](#). Many websites are - let's face it - mostly navigation to help visitors find a specific PDF. Maybe that's why, after HTML, [PDF is the 2nd most common format on the internet](#). A recent [survey by Bitkom Research](#) in Germany showed that "PDF...has become indispensable for the vast majority of companies".

[You can read the entire post here.](#)

Permit me a short anecdote. AIIM was involved in the development of the PDF/A standard. At the time I reported to Betsy Fanning, who headed up AIIM's standards initiatives (and thanks to Betsy for alerting me to the anniversary). I tagged along to one of the early standards meetings near Union Station in DC.

After thirty minutes, I quickly realized how in over my head I was as the conversation went DEEP into the intricacies of standards, needs, etc. The words were English, but the way they were strung together, they may have well been Greek.

For anyone out there reading this who helps create/support the standards that make modern life livable, THANK YOU!

## AIIM Conference Snapshot via Ralph Gammon

Ralph Gammon, senior analyst, software for Infsource, shared the following note.

AI was the most popular topic of discussion at the AIIM Conference held April in New Orleans. I led a roundtable discussion titled, "Get to know Intelligent Document Processing (IDP): your first step in end-to-end business automation." There was a full room, but it was a challenge to get the attendees to talk about their use cases. Many were having trouble connecting the buzz around AI with their existing information management requirements.

Attendance at the conference was up significantly from 2022, but the primary end user group seemed to be records managers, many of whom are just starting to explore potential use cases for AI. While AI certainly has a place in Records Management, its initial use in the Capture & IDP market has been primarily for transaction-driven applications in areas like Case Management and Accounting, although the potential of using it for more advanced records management applications like data analytics is intriguing.

### Microsoft Vision for Combining Generative AI & IDP

At the conference Microsoft demoed how its market leading generative AI can be leveraged along with Capture and IDP to solve information management challenges for businesses. During its keynote presentation, Microsoft showed how users can leverage the recently released GPT-4 large language model to search internal content. This is a step toward further commercializing the technology.

To date, large language models have been primarily applied to publicly available content, which has led to great adoption, but is not really something that is going to help Information Management professionals too much. What Microsoft showed was the use of "co-pilot" technology to help narrow down searches utilizing GPT-4.

After the demo, Chris McNulty, Director of Product Marketing for Microsoft 365, explained how he envisions Capture & IDP playing a vital role in GPT-driven searches. If Capture and IDP is not used initially for indexing and tagging businesses' content before running it through large language models, McNulty explained that leveraging large language models in a search process is cost-prohibitive and will take too long to execute to be an effective tool.

## News and In Brief

**BIS** recently integrated Open AI's APIs into its Grooper Capture & IDP platform. According to Tim McMullin, VP of enterprise sales for BIS, GPT greatly increases the number of parameters that can be used for auto-classification within Grooper.

**NARA** has published new regulations for digitizing permanent federal records (US); 36 CFR 1236.Subpart E. [Read more here.](#)

**Digitech Systems**, a cloud information management provider, and **Contex**, a provider of large-format scanning, announced a summer bundle for the secure information management of large format scanned assets. Available until August 31, 2023, customers who purchase any Contex HD Ultra X or IQ Quattro X large format scanner can receive a license for Nextimage scanning software and three months access to the PaperVision.com Professional service, both free of charge. See more [here](#).

**ABBYY** appointed **Neil Murphy to Chief Sales Officer (CSO)**. He will be extending the company's focus on helping global enterprises put their information to work and accelerate business value from their digital transformation initiatives. Murphy has been at the forefront of leading customer-centered strategies with direct sales and partners in EMEA and APAC in a broad spectrum of industries, giving him deep insights into the challenges facing organizations today.

In early April, **Hyland Software** announced it would be laying off approximately 20% of its 5,000-person workforce. In a message explaining the lay-offs, President and CEO Bill Priemer wrote that Hyland, "has been navigating the global economic situation and shifts in our market. We are transforming into a cloud company.... Our new organizational structure will make us more effective at developing, delivering and supporting cloud-based solutions.... At the leadership level, we are removing layers of management to improve communication and accelerate decision-making." The company intends to continue to invest in all its platforms and doesn't have any plans to end-of-life product lines.

**TCG Process** has signed a partnership agreement with **Duck Creek Technologies**, which provides core systems to the P&C Insurance market. Duck Creek focuses on policy, rating, billing and claims solutions. It is based in Boston and reported approximately \$300M in revenue in 2022, with \$280M coming in the U.S. TCG Process is a Swiss-based Capture & IDP ISV with an established presence in the insurance market in the DACH region. The Duck Creek partnership represents a good avenue for increasing North American sales. Duck Creek relies on a range of partners to help it deliver complete solutions. It does not currently have any other IDP partners listed on its website.

**ShinyDocs**, which specializes in helping users uncover their dark unstructured data, recently raised \$16.25M in funding. ShinyDocs is based in Waterloo, ON, and the round was led by Toronto-based First Ascent Ventures and Export Development Canada. ShinyDocs was founded in 2013 and the funding is earmarked to help it expand sales and marketing.

**Kodak Alaris** recently introduced its new E1000 series, with two workgroup segment models, the E1030 and E1040. They feature Dual Illumination, which is designed to improve image quality and increase OCR accuracy by reducing shadows.

**Xerox Emirates** and **Kofax** recently expanded their partnership in the UAE. Xerox Emirates will now be able to market, resell, implement, and support multiple Kofax products including Kofax TotalAgility (KTA). Xerox Emirates will also provide implementation and professional services to joint customers.

I.R.I.S. recently appointed a new Managing Director for its Product and Technologies Group. **Gerald Rudiger** joined I.R.I.S. in February after spending seven years with EASY Software, a document management ISV focused on the DACH region. In addition, **Marcel Rosenbaum** has been promoted to CTO, while **Jean-Louis Amphoux** has been promoted to CFO.

Cloud IDP specialist **Rossum** recently named its first Chief Marketing Officer. **Elena Melnikova** joins Rossum after spending eight years as a marketing executive at Talkwalker, a consumer intelligence business, where she built a global marketing organization. Based in Prague and founded in 2017, Rossum made waves in 2021 when it raised a \$100M series A round, a record for an Eastern European company. To date, Rossum has focused primarily on invoice capture and is building out its AI-driven platform. The majority of its business has been in Europe, but it has global expansion plans.

**Canon** has introduced a new version of its DR-M1060 departmental document scanner. Like its predecessor, the DR-M1060II is rated at 60 ppm/120 ipm and features an 80-page ADF and both u-shaped and straight paper feeding options. The DR-M1060II features upgraded software and a three-year Advanced Exchange Warranty. The new scanner comes bundled with Canon's CaptureOnTouch v5, which includes bar code reading and batch separation functionality as well as output to Word and Excel formats. Canon has dropped the bundling of Kofax VRS in the new model.

## DOCUMENT IMAGING REPORT

### Business Trends on Converting Paper Processes to Electronic Format

DIR is the leading executive report on managing documents for e-business. Areas we cover include: Document Capture; OCR/ICR, AI and Machine Learning; RPA; ECM; Records Management; Document Output; and BPM.

DIR brings you the inside story behind the deals and decisions that affect your business.

Vol. 32, No. 5

Managing Editor: Ralph Gammon; [rg@info-source.com](mailto:rg@info-source.com)  
Editor-in-Chief: Bryant Duhon, [bdu@info-source.com](mailto:bdu@info-source.com) and +1 (301) 275-7496

DIR is published approximately 15 times per year by:

Infosource SA  
Avenues des Grande-Communes 8, 1213  
Petit-Lancy, Geneva, Switzerland  
<http://www.info-source.com>

Copyright © 2023 by Infosource SA. Federal copyright law prohibits unauthorized reproduction by any means including photocopying or facsimile distribution of this copyrighted newsletter. Such copyright infringement is subject to fines of up to \$25,000. Because subscriptions are our main source of income, newsletter publishers take copyright violations seriously. Some publishers have prosecuted and won enormous settlements for infringement. To encourage you to adhere to the law, we make multiple-copy subscriptions available at a substantially reduced price.

Subscriptions available from single-user to enterprise-wide. Contact [bdu@info-source.com](mailto:bdu@info-source.com).

