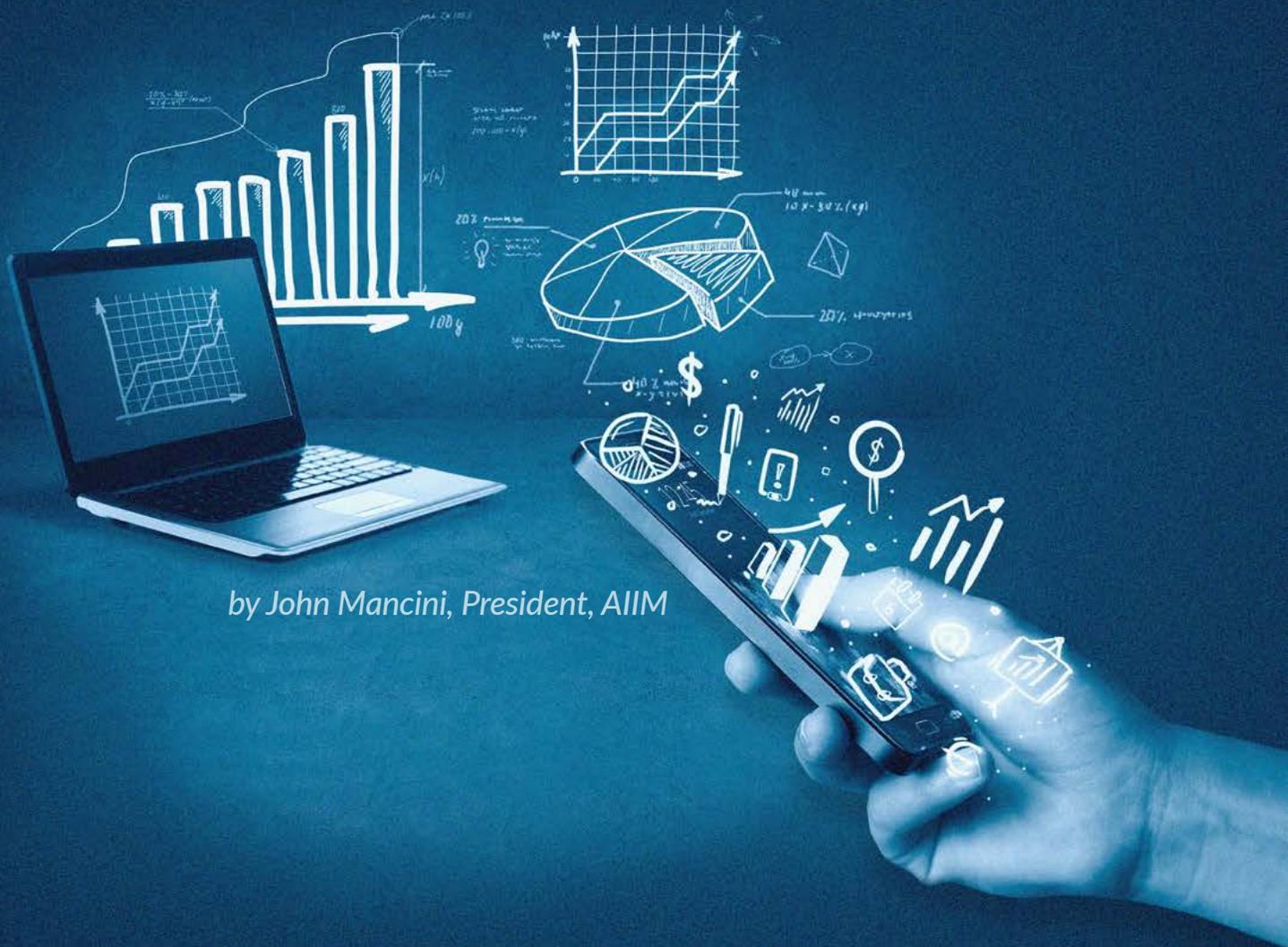




AIIM TRENDSCAPE:

# CONTENT ANALYTICS



*by John Mancini, President, AIIM*

## Table of Contents

<b>1) Executive Summary .....</b>	<b>3</b>
<b>2) Obstacles to content analytics adoption that need to be overcome.....</b>	<b>5</b>
<b>Obstacle #1</b> - Definitions of what content analytics actually is are poorly understood and inconsistent. ....	5
<b>Obstacle #2</b> - There is still too much focus on technology and not enough on business value; who truly are the buyers, influencers and key stakeholders and how are they different from those of yesteryear? .....	6
<b>Obstacle #3</b> - The pace of technology change is accelerating and running ahead of the ability of organizations to consume it.....	7
<b>Obstacle #4</b> - There is an on-going lack of specific use cases to really demonstrate value.....	8
<b>Obstacle #5</b> - There is a splintering of data -- and data responsibility -- in the enterprise.....	10
<b>5 Data Points You Need to Know About Content Analytics</b> .....	10
<b>3) Strategic Lessons Learned from Early Implementers.....</b>	<b>11</b>
<b>Lesson #1</b> - Address data structure, tagging and permissions as close as possible to the point of creation.....	11
<b>Lesson #2</b> - Define a clear set of corporate objectives for content analytics initiatives.....	12
<b>Lessons #3</b> - Always remember that the hardest part of a content analytics project is not the technology. ....	13
<b>Lesson #4</b> - Done right, content analytics should drive you to rethink the questions you ask about your organization. ....	13
<b>Lesson #5</b> - Position your content analytics initiatives as a means to both enhance customer value AND address information security -- and get the buy-in of business stakeholders EARLY. ....	14
<b>What You Should Do TODAY: Your Eight Point Content Analytics Action plan.....</b>	<b>15</b>
<b>About John Mancini .....</b>	<b>19</b>
<b>About The Research.....</b>	<b>20</b>



“ There are decades in which nothing happens and there are weeks in which a decade happens. ”

— Vladimir Lenin



“ The status quo is perpetually on trial. But that doesn't mean you can stand by and wait. It wasn't raining when Noah started building the ark. ”

— Thornton May

## 1. Executive Summary

Deriving meaning and insight from the different forms of content held and managed within most organizations is fast becoming one of the number one challenges and priorities for many in business. AIIM's June 2015 Industry Watch, [Content Analytics: automating processes and extracting knowledge](#), revealed that six in ten enterprises believe that content analytics will be essential within five years' time:

“ The capacity of computers to recognize meaning in text, sound or images has progressed slowly and steadily over many years, but with the constant arrival of faster computing resources, and the continual refinement of software algorithms, we are in a position where both the speed and the accuracy of recognition can support a wide range of applications. ”

“ In particular, when we add analysis to recognition, we can match up content with rules and policies, detect unusual behavior, spot patterns and trends, and infer emotions and sentiments. Content analytics is a key part of 'big data' business intelligence, but it is also driving auto-classification, content remediation and Information Governance, security correction, adaptive case management, and operations monitoring. ”

That smart content analytics can be of huge value to an organization is clear. But it is equally clear that there are a number of obstacles to content analytics adoption that need to be overcome. At the most recent meeting of AIIM's think-tank, the Executive Leadership Council (ELC), the following obstacles were identified as particularly pertinent:

- Definitions of what content analytics actually is are poorly understood and inconsistent.
- There is still too much focus on technology and not enough on business value; who truly are the buyers, influencers and key stakeholders and how are they different from those of yesteryear?
- The pace of technology change is accelerating and running ahead of the ability of organizations to consume it.
- There is an on-going lack of specific use cases to really demonstrate value.
- There is a splintering of data -- and data responsibility -- in the enterprise.



As with any emerging technology, there has been a wave of content analytics early adopters. The ELC agreed that it was imperative to learn strategic lessons from these early implementers, particularly focusing on:

- Addressing data structure, tagging and permissions as close as possible to the point of creation.
- Defining a clear set of corporate objectives for content analytics initiatives.
- Always remembering that the hardest part of a content analytics project is not the technology.
- Done right, content analytics should drive a rethinking of the questions you ask about your organization.
- Positioning your content analytics initiatives as a means to both enhance customer value AND address information security.



In conclusion, a lot is happening in the world of content analytics, and it is happening very quickly indeed. But organizations must embrace the changes and the pace at which they are happening, as content analytics takes its place at the heart of many organizations.





## 2. Obstacles to content analytics adoption that need to be overcome.

### Obstacle #1 - Definitions of what content analytics actually is are poorly understood and inconsistent.

AIIM has defined content analytics in the following way:

“ Content analytics is the act of applying business intelligence (BI) and business analytics (BA) practices to digital content. Companies use content analytics software to provide visibility into the amount of content that is being created, the nature of that content and how it is used. ”

Yet as with so many areas of technology, there is an on-going debate about the “true” meaning of content analytics. There is a lack of common language and concepts that can be used across organizations and industries and often, if you asked people from different teams in the same organization about content analytics, you would get a different answer.



#### John Mancini, AIIM:

“ Ask an ‘ECM’ person what they mean by ‘unstructured information’ and you will get an answer focused on PDFs and application files and TIFF images. Ask a ‘Big Data’ person, and you will get an answer focused on the data exhaust associated with the web and the internet of things. ”

This is something that must be addressed as a priority, but progress is undoubtedly being made as organizations focus less on the distinction between data and content and instead on outcomes.

#### John Mancini, AIIM:

“ Traditional lines between data and content are blurring. No one cares about the distinction; they care about how information intersects with a business outcome. ”



**Obstacle #2 - There is still too much focus on technology and not enough on business value; who truly are the buyers, influencers and key stakeholders and how are they different from those of yesteryear?**

Business outcomes from content analytics should be a higher priority than getting bogged down in technological jargon and slightly different interpretations of what content analytics is. Content analytics needs to be repositioned. It not only drives business decisions -- it is critical to business transformation. Organizations should be asking such questions as:

- Are our customers leaving? If so, why?
- If customers are staying but assets are leaving, why?
- Where are the customers or assets going?
- What are the drivers?
- How do we get early and accurate detection?
- What do we do about it?



This is something that ELC participants felt particularly strongly about, and also an area where most agreement was found.

**Thornton May:**

“ We are in a period of ‘competitive victimhood.’ Every human population parses as follows: 17% are early adopters, 70% are mass adopters with a ‘show me attitude,’ and 13% are refuseniks. We need to find and celebrate the early adopters. ”

**Monica Tolson, Assistant Director, BC Pension:**

“ Use content analytics to inform your customer experience plan, keeping the focus on ease of use and seamless flow. The technology or platform doesn’t matter to users. ”

**Pamela Click, Divisional CIO - CSEB Technology, TIAA-CREF:**

“ We don’t want a scanner, a database, or an application. People just want a button that does what they need. ROI and business case is critical in order to move forward and it is often unclear. ”



**Jean Louis Arsenault, EVP of Sales and Marketing, Haystac:**

“ Focus the value proposition on helping organizations understand in a way that is repeatable, quantifiable, and defensible. No one asks for information governance, but they are realizing there is too much information and that they must decide what is really needed. ”

**Thornton May:**

“ The 3 Vs (Volume, Velocity, Variety) are being replaced by the 4 Ms (Make Me More Money). ”



**Dennis Devlin, Co-Founder and CISO, SAVANTURE:**

“ Content analytics is a critical tool that will enable us to begin to address the digital landfill and to stay current with acquisition, creation and discovery. We will be able to do a much better job managing information risk when we can illuminate dark data, classify consistently, reduce our attack surface and protect data commensurate with risk. ”

### Obstacle #3 - The pace of technology change is accelerating and running ahead of the ability of organizations to consume it.

The consumerization of IT makes it increasingly difficult for organizations to keep pace with that change. Knowledge workers are used to a level of technology performance in their home technology that is often not yet present in the workplace. We are in an era of “bimodal Content Management.” Mode 1 is focused on assuring information and process reliability. The real value is in Mode 2, creating business agility. It’s the difference between a marathon runner and a sprinter.



**Pamela Click, Divisional CIO - CSEB Technology, TIAA-CREF:**

“ Most of our customers don’t use our website or login to their online account - instead they rely on paper based communication and phone. The technology is growing faster than our current participants are ready to absorb can absorb. ”



## Thornton May:

“ One of THE most important outputs of content analytics might be matching content tools to content uses/users. ”

Dennis Devlin, Co-Founder and CISO, SAVANTURE outlined his vision of the content analytics reality right now, and what questions must be addressed:

### The Content Analytics Reality

- Bits have replaced atoms
- Digital technology is ubiquitous
- Storage is unlimited and “free”
- Everything is networked
- Everyone publishes
- Copies are originals
- Digital is forever

### The questions that must be addressed:

- What data do we have?
- Which data can we trust?
- How much of our data is at risk?
- Are we regulatory compliant?
- Who can I trust with my data?
- Who respects my privacy?



### Obstacle #4 - We lack specific use cases.

Demonstrating how and in what ways a technology can help an organization is of course a key element in encouraging adoption. If organization A has achieved great success with content analytics, then organization B will want in on that. Unfortunately, content analytics does not inherently have a “killer app” in the way that other technologies does -- it is a catalyst to resolving business problems.



**Theo Priestley, Technology Evangelist:**

“ Organizations are not applying content analytics due to industry creating segmented offerings. There are no practical examples of content driven process using content analytics as the automated trigger for decision making. ”

**Jeremy Phillips, COO, EditorEye:**

“ Collaboration is driven by content and is a key use case for content analytics. Content analytics is key to understanding the connection between topics (what people care about) and communities (how they interact with others). ”



**John Mancini:**

“ Defensible disposition is often viewed as a key use case for content analytics. But are we clear on the business case for spending money to solve the ROT problem (Redundant, Obsolete, and Trivial data and content)? Is ROT mission critical? ”

**Mark Roboff, Executive Evangelist, IBM Watson:**

“ We need to understand not just the problems that content analytics solves, but how it happens. Why is unstructured content better at answering questions for decision making than structured data? ”



## 5 Data Points You Need to Know About Content Analytics

Source: AIIM -- Content Analytics: automating processes and extracting knowledge

**53%** of organizations agree that auto-classification using content analytics is the only way to get content chaos under control.

**54%** feel that their organization is exposed to considerable risk due to stored content that is not correctly identified.

**73%** consider that there is real business insight to be gained if they can get the analytics right.

**63%** are being held back by a lack of analytic skills and an absence of allocated responsibilities.

**17%** consider content analytics to be "essential" now for their organization, growing to **59%** in 5 years' time.

### Obstacle #5 - There is a splintering of data -- and data responsibility -- in the enterprise.

In AIIM's June 2015 Content Analytics Industry watch, 80% of the research respondents said they were not yet able to allocate a senior role to initiate and coordinate analytics applications. This, combined with the ever-increasing volumes of data, provides a major obstacle to content analytics adoption.

**Jeremy Phillips, COO, EditorEye:**

*“Content is fragmented across many places and applications; users have to search across these and are frustrated which in turn limits their desire to do so.”*

**Theo Priestley, Technology Evangelist:**

*“Chief Data Officers are critical to implementing analytics. CDOs must have both change management skills AND communication skills. They must drive short term improvements in data quality. And most importantly, they are key to the evolution and disruption of business models.”*

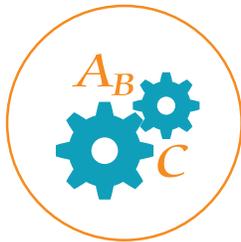




## 3. Strategic Lessons Learned from Early Implementers.

### Lesson #1 - Address data structure, tagging and permissions as close as possible to the point of creation.

Organizations often take the easy option when it comes to content tagging and permissions, figuring that they will get to it “sometime.” However, it really helps to address these issues as soon as you possibly can.



#### Thornton May:

“ CSIS (the Canadian NSA) has categories and rules to label data as it comes in -- every paragraph is labeled separately. Labeling data on the way in makes information management more difficult to initially implement, but ultimately easier to consume and manage. ”

#### Russ Stalters, CEO, Clear Path Solutions Inc.:

“ You need a defined information and data on-ramp so that you have a deliberate and controlled data ingestion process. This is critical to ensuring the quality of the underlying data used for content analytics. ”



Theo Priestly believes that organizations need to adopt a “Blockchain” approach to security. Per Wikipedia, [Blockchain](#) is a “permissionless distributed database based on the bitcoin protocol that maintains a continuously growing list of transactional data records hardened against tampering and revision, even by operators of the data store’s nodes. The initial and most widely known application of the blockchain technology is the public ledger of transactions for bitcoin which has been the inspiration for similar implementations often known as altchains.”

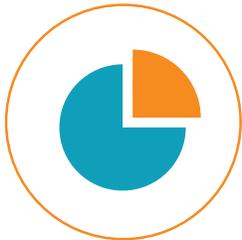


Priestley believes that a Blockchain approach has these advantages:

- Storage Cost - An autonomous decentralized storage system would reduce the cost of storage
- Security - Client-side encryption will always beat server-side
- Speed - The scale of a true decentralized P2P network would be vastly superior to a centralized server

## Lesson #2 - Define a clear set of corporate objectives for content analytics initiatives.

Deciding what you want to achieve before embarking on ANY project is basic business, yet it doesn't always follow that that will happen. Two of our ELC participants had strong advice on defining corporate objectives for your content analytics objectives.



**Jane Alexander, CIO, Cleveland Museum of Art:**

- Ensure your backend systems are flexible and interoperable
- Your key data for artwork must be clean, current, and represent one version of the truth
- No one-off projects
- Think of everything in terms of the BIG PICTURE
- Content reuse - repurpose content created for other purposes

**Pamela Click, Divisional CIO - CSEB Technology, TIAA-CREF:**

- Customer Perspective: Identify the "why" vs. the "what"
- Employee Perspective: Improve productivity; increase job satisfaction
- Generate and retain customers
- Generate revenue
- Brand recognition
- Not just metrics - - behaviors



## Lesson #3 - Remember that the hardest part of a content analytics project is not the technology.

Sourcing the right technology to achieve your goals is actually one of the easier elements to content analytics. What you must really never forget is the people element.

**Jane Alexander, CIO, Cleveland Museum of Art:**

“ The hardest part of linking silos was the people aspect. The pace was slow, reluctance was hard to deal with and unrealistic expectations that things should not/could not move that fast. ”

**Jeremy Phillips, COO, EditorEye:**

“ Content nirvana takes different forms depending on your role in the organization, and content analytics is key to delivering this value. Power users seek to share the flow of content, help colleagues out most efficiently, and control costs. The average end user, though, prefers to view content analytics as magic, mysteriously delivering the right stuff to them when they need it without needing to do anything. 90% of the users in any group fit in this latter category. ”



## Lesson #4 - Done right, content analytics should drive you to rethink the questions you ask about your organization.

The insight that content analytics can generate will not only provide you with the answers to the questions you ask of your organization, but it will also help make the questions even more pertinent, making the next set of answers even more insightful.

**Steve Lohr, New York Times:**

“ Defining problems more expansively is part of the design-thinking ethos. At a course in New York recently, a group of IBM managers were given pads and felt-tip pens and told to sketch designs for ‘the thing that holds flowers on a table’ in two minutes. The results, predictably, were vases of different sizes and shapes. Next, they were given two minutes to design ‘a better way for people to enjoy flowers in their home.’ In Round 2, the ideas included wall placements, a rotating flower pot run by solar power and a software app for displaying images of flowers on a home TV screen. ”



**Lesson #5 - Position your content analytics initiatives as a means to both enhance customer value AND address information security -- and get the buy-in of business stakeholders EARLY.**

Content analytics is not necessarily a silver bullet for all of an organization's key challenges, but it does have value beyond enhanced customer value. The ELC meeting focused a lot on the value and insight to be derived, but content analytics also has a significant role to play when addressing information security.

**Mark Roboff, Executive Evangelist, IBM Watson:**

“ My view is that it is a mistake to position content analytics as an ECM tool. Analytics needs to be understood in the context of a business problem and with a clear understanding of how and why they want/need analytics. Content analytics should not be part of an ECM sale alone; it requires analytics expertise and should be part of a larger analytics implementation. ”



**Dennis Devlin, Co-Founder and CISO, SAVANTURE:**

“ I am a recovering information risk management professional. Information risk used to be a lonely profession. But now I have new friends -- in Business Intelligence, Compliance and Audit, Privacy, Boards of Directors, Corporate Communications, Investor Relations, Office of General Counsel, Information Management, Physical Security. Collectively we are losing the security battle. There are more security breaches than ever before, more unknowns than knowns, growing digital landfills that no one seems to be responsible for, and I am still being asked to 'protect our digital assets.' ”



## What You Should Do TODAY: Your 8 Point Content Analytics Action Plan

01

If your content or records management deployment is stalled due to poor decisions early on regarding classification, metadata and taxonomies, or if you are migrating content from multiple repositories to a single system, take a look at metadata correction agents that can sort ROT from valuable content, and align content types and metadata.



02

If you have access to contextual search, ensure that it is properly tuned, and that staff knows how to use it. If you are reliant on more basic search, consider improving the searchability, and therefore the value of your content, by correcting and enhancing the metadata using analytic agents.



03

Unless your staff is diligent and consistent at declaring, classifying and tagging records, consider providing auto-classification assistance or full auto-classification. Be aware that your information governance policies need to be updated and consistent as they will provide the rules for automated agents.

03



04

Take control of your emails. If you have no archive, or the archive is "file and forget" you are losing potential corporate knowledge, but are also exposing the business to risk, and creating a potential e-discovery nightmare.

04

05

Look at your retention policies as a way to control increasing storage requirements. Accurate metadata and enforced retention policies are the only way to limit storage, but will also improve your compliance and risk exposure.



06

Inbound content handling can rapidly overload process staff, and reduce speed of response to customers. Implement a digital mailroom philosophy, and use automated recognition, routing and data extraction.



Look across the range of your business activities to see where content analytics could provide business insight to understand customer needs, improve competitive advantage, help to solve cases and investigations, or prevent non-compliance and fraud.

07



Check out AIIM's Content Analytics Resource Center and Training Course.

[www.aiim.org](http://www.aiim.org)

08

Thank you to our US Executive Leadership Council Companies who underwrote this research:



Thank you to our EU Executive Leadership Council Companies who underwrote this research:



**Canon**

**CAPITA**

  
ccubesolutions

**CROWN**  
WORLDWIDE GROUP

*DataSolutions*  
from  Group plc

**Deloitte.**

**DocuWare**  
*From Documents to Value*

  
**EASY SOFTWARE**

**FUJITSU**

**IBM.**

 **ibml**

**KELVIN  
TOP SET**  
Professional Incident Investigation

**Kodak alaris**

  
KONICA MINOLTA

**OPEN TEXT**

 **Software  
Innovation**

**spigraph**  
network 

**SPS**  
a Swiss Post company

 **Workshare**

**xerox** 



## About John Mancini

**John Mancini** is an author, speaker and respected leader of the AIIM global community of information professionals. He believes that in the next 5 years, a wave of Digital Transformation will sweep through businesses and organizations, and that organizations now face a fundamental choice between Information Opportunity and Information Chaos.

As a frequent keynote speaker, John offers his expertise on Digital Transformation and the struggle to overcome Information Chaos. He blogs under the title “Digital Landfill” and has almost 10,000 Twitter followers and a Klout score in the 60s. He has published seven e-book titles including “Information Chaos v. Information Opportunity: The Business Challenge for the Next Decade” (<http://www.aiim.org/infochaos>), “#OccupyIT – A Technology Manifesto for Cloud, Mobile and Social Era” and the popular “8 Things You Need to Know About” e-book series.

### He is the author of “Mancini’s Law”:

- Organizations are systems of information networks. They only operate effectively when there are clear and predictable information flows within and between these networks.
- 50% annual growth in the volume of digital information means that these networks – and especially the points of connection between them – will become increasingly unstable.
- Without intervention, the resulting #infochaos will threaten the viability of the entire system.

John can be found on Twitter, LinkedIn and Facebook as [\*jmancini77\*](#)

## About The Research

As the non-profit association dedicated to nurturing, growing, and supporting the information management and social business community, AIIM is proud to provide this research at no charge. In this way, the entire community can leverage the education, thought leadership, and direction provided by our work.

We would like this research to be as widely distributed as possible. Feel free to use this research in presentations and publications with the attribution – © AIIM 2016, [www.aiim.org](http://www.aiim.org)

Rather than redistribute a copy of this report to your colleagues, we would prefer that you direct them to <http://info.aiim.org/aiimtrendscape> for a free download of their own.



© 2016

### AIIM

1100 Wayne Avenue, Suite 1100  
Silver Spring, MD 20910  
+1 301.587.8202  
[www.aiim.org](http://www.aiim.org)

### AIIM Europe

Broomhall Business Centre  
Worcester, WR5 2NT, UK  
+44 (0)1905.727.600  
[www.aiim.org](http://www.aiim.org)

## About AIIM's Executive Leadership Council

In 2012, AIIM formed a think tank to define, discuss and offer directives on today's emerging issues in information management. This think tank is the Executive Leadership Council (ELC).

The ELC brings together top thinkers, high performance practitioners and leaders in information management for two theme-centric summits annually. Each summit creates a shared space for dynamic conversations to determine the role of the information management industry in a new era of business.

### *Want to Participate?*

Details of the 2016 summit themes can be found at [www.aiim.org/elc](http://www.aiim.org/elc). Should you be interested in learning more about participating in the Executive Leadership Council, please contact Jessica Lombardo at [jlombardo@aiim.org](mailto:jlombardo@aiim.org)