Content Intelligence

For the Future of Work

Sponsored by: ABBYY

Holly Muscolino

November 2019

IDC
ANALYZE THE FUTURE

IN THIS WHITE PAPER

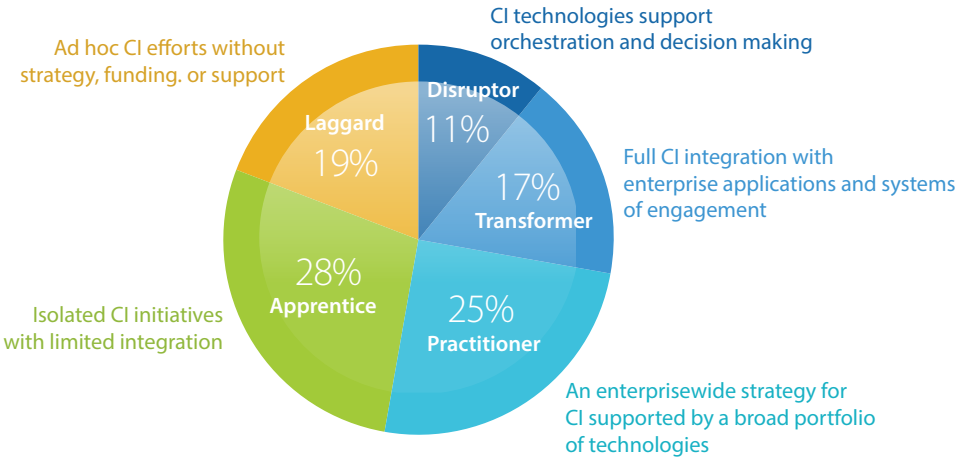
This IDC White Paper **explores content intelligence** (CI or "content IQ") and **the role that it plays** in driving and scaling digital transformation. The document **investigates the priorities, challenges, and benefits** associated with content intelligence and reviews how organizations are executing digital transformation initiatives today and planning for the future. Finally, this document delves into what the future holds for automation and the new age of the digital worker.

METHODOLOGY

To assess the CI trends, IDC conducted a global survey of 500 senior influencers and/or decision makers who were knowledgeable about how their organization digitized, automated, and/or optimized content and document workflows with information technology products and services. All organizations had deployed, or were planning to deploy, content intelligence technologies. Survey participants were from organizations with 500+ employees and were located in North America, Europe, and Asia/Pacific. Respondents represented five industries: banking/investment, insurance, transportation and logistics, manufacturing, and business process outsourcing (BPO).

IDC asked survey participants to evaluate the content intelligence maturity of their organization. Respondents were divided into five levels of content intelligence maturity (for more details, see the Definitions section). Almost half of the survey respondents (47%) were at the earliest stages of content intelligence maturity, indicating significant room for growth, even among organizations that are embracing these technologies. Only one-quarter of respondents were content intelligence "practitioners," and 28% were at the highest stages of content intelligence maturity (see Figure 1).

FIGURE 1 Content Intelligence Maturity



n = 500 Source: Content IQ Thought Leadership Survey 2019, IDC, May, 2019

SITUATION OVERVIEW

Introduction

Today, organizations are rethinking work as we know it. We are seeing a fundamental shift in the work model to one that fosters human-machine collaboration, enables new skills and worker experiences, and supports an environment unbounded by time or physical space. Many companies, including IDC, call this the "future of work." However, the reality is that we are seeing many of these changes occurring in the present day.

"Digital workers" are making up a growing share of the workforce. We define a "digital worker" as technology — including artificial intelligence (AI), intelligent process automation (IPA), augmented reality/virtual reality (AR/VR), and software robotics — that automates and augments work previously accomplished by humans. IDC research shows that the contribution of digital workers will increase by over 50% in the next two years. However, it is important to note that, in most cases, these technologies enhance, rather than replace, human capabilities, thus driving productivity and enabling human workers to focus on higher-value tasks.

In the future workplace, organizations will maximize value through automation and augmentation. These new approaches also *involve maximizing the value of content and information assets.* However, in the short term at least, a new generation of required

skill sets are in short supply. In a recent IDC survey, 75% of respondents said that their organization was finding it difficult to recruit digital skills, at least in some areas.

This IDC White Paper explores the concept of content intelligence as a set of technologies and services that leverage artificial intelligence to carry out important skilled tasks in this new era.

Maximizing the Value of Enterprise Content

Future-focused enterprises understand the value inherent in their organization's vast stores of unstructured content. Improving business operations and enabling better decision making are the top priorities for transforming document workflows (see Figure 2). Survey respondents also noted revenue growth, improved customer experience, cost reduction, and mitigating security risk as top priorities for managing unstructured content.

It is worth noting that the top priority for banking/ investment respondents was improved customer experience (34%). Manufacturing respondents highlighted improved decision making (31%) as the top priority.

Q.

What are your organization's top priorities for transforming document-intensive workflows and how unstructured content is managed?

FIGURE 2 Priorities for Transforming Document Workflows

n = 500 Source: Content IQ Thought Leadership Survey 2019, IDC, May, 2019

Content Workflow Transformation Drives Business Transformation

Organizations are struggling with fully realizing the integral value of enterprise content. When survey respondents were asked about the factors that drove them to deploy (or plan to deploy) content intelligence technologies, they indicated manual sorting and classification of documents, manual data extraction from documents, inadequate compliance with security/privacy regulations, and poor data, errors, and inaccuracy of information as top pain points (see Figure 3).

Organizations also need data for enterprise applications, and rekeying information to bridge disparate systems was a common problem.

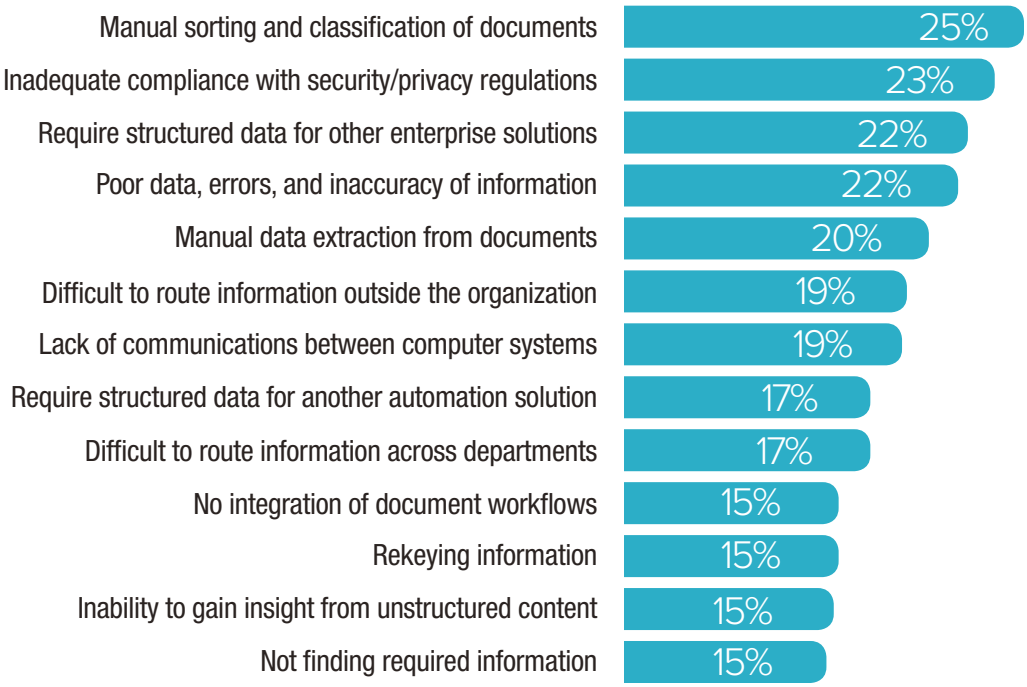
There are some important differences in these factors depending on industry. Requiring structured data for other enterprise solutions was the most important factor for both manufacturers (19%) and BPO organizations (19%). Inadequate compliance with mandated security/privacy regulations was the most important factor for insurance organizations (28%).

These inefficiencies and pain points impact business outcomes in a number of ways. About one-quarter of survey respondents told us that the most challenging business impacts of poorly managed, unstructured content were the lack of customer information, inadequate data to support decision making, and the lack of control over content, including challenges with identifying and managing sensitive data.

Q.

What were the factors that drove you to deploy (or plan to deploy) content intelligence technologies?

FIGURE 3 Factors Driving Content Intelligence



n = 500 Source: Content IQ Thought Leadership Survey 2019, IDC, May, 2019

Optimizing Content-Centric Workflows with Content Intelligence

Clearly, a new portfolio of content intelligence technologies is required to overcome the previously mentioned business impacts and other challenges inherent with enterprise content. We define content intelligence as a set of technologies and services that leverage artificial intelligence to carry out tasks such as reading and categorizing a document, routing a document, extracting and validating data from documents, and other tasks related to understanding and processing unstructured content to enable quicker, more accurate decision making and deliver greater business value. Relevant content IQ technologies include OCR, computer vision, machine learning, natural language processing, and content analytics.

Survey respondents experienced a number of business benefits by deploying content intelligence technologies. Almost 40% of respondents noted increased employee productivity or increased customer satisfaction. Over one-third of respondents saw an improvement in responsiveness to customers, new product or revenue opportunities, increased visibility and/or accountability, or increased customer engagement. Respondents agreed that they were able to redirect resources to higher-value tasks. Content IQ disruptors and transformers (i.e., companies at higher levels of CI maturity) were more likely to experience these benefits.

Past IDC research has shown that organizations deploying technology related to digitizing, automating, and optimizing document workflows reduced costs by over 35%. They also reduced the amount of time spent on document-related tasks by 17%. And perhaps the most compelling benefit was that they reduced errors by almost 52%.

Over half of the survey respondents also indicated technical benefits from deploying content intelligence technologies, including greater automation of repetitive processes, better decision making, and ease of finding required information.

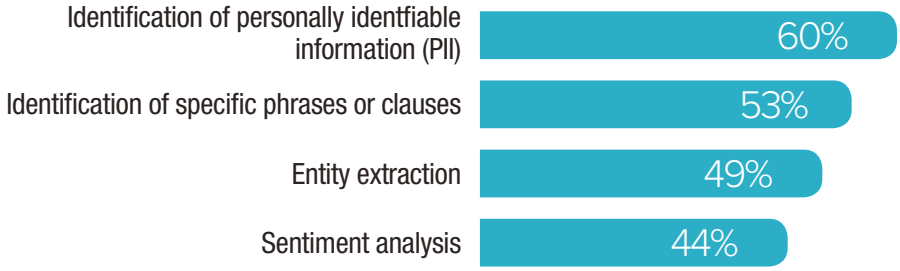
When respondents were asked what type of analysis is (or will be) conducted on the data extracted via content intelligence technologies, 60% indicated that CI technology is used to identify personally identifiable information (PII). Other key use cases are identification of specific phrases or clauses, entity extraction, and sentiment analysis (see Figure 4).

Respondents across all industries said content intelligence technologies enabled several key corporate initiatives, including employee engagement, customer engagement, work transformation, and overall digital transformation (see Figure 5). Banking/investment respondents cited greater need for shareholder engagement initiatives (44%) than other industries.

Q.

What type of analysis is (or will be) conducted on the data extracted via content intelligence technologies?

FIGURE 4 Data Analysis via Content Intelligence

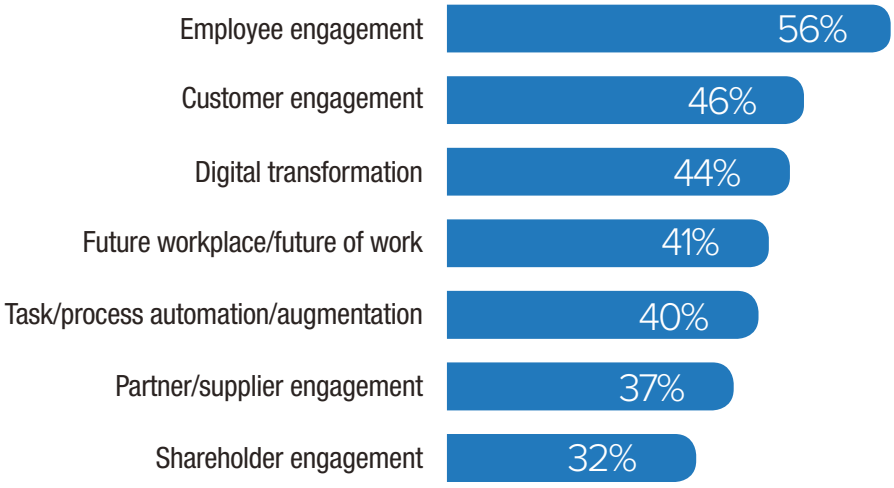


n = 500 Source: Content IQ Thought Leadership Survey 2019, IDC, May, 2019

Q.

In which of the following corporate initiatives does (or will) content intelligence technologies play a major role? In other words, which initiatives will content intelligence technologies significantly enable and/or impact?

FIGURE 5 Corporate Initiatives Enabled by Content Intelligence



n = 500 Source: Content IQ Thought Leadership Survey 2019, IDC, May, 2019

Executing Content Intelligence

Since there are clear business benefits to deploying content intelligence technologies, why are half of the organizations still at the earliest stages of content intelligence maturity? Almost one-third of respondents named cost as a barrier to deployment of CI technologies; however, over 20% cited inadequate worker skills and/or training as a leading challenge. Inadequate employee skill sets was the leading barrier to an organization becoming a content IQ disruptor or transformer, indicating a clear need for easily consumable training.

Siloed business units, disparate legacy systems, and lack of upper management support were other barriers mentioned by respondents for CI technology deployment, all pointing to the requirements for integration, standardization, and collaboration — in addition to leadership — as critical for success.

In fact, our respondents told us that the deployment of content intelligence technologies is an initiative led by senior management. 78% of respondents cited the chief executive officer (CEO) as a top decision maker. Half of the respondents also included the chief information officer (CIO), 41% cited the chief technology officer (CTO), and 37% cited the chief digital officer (CDO) as decision makers. On the other hand, budget holders were spread more evenly across the organization, with 30–37% of respondents mentioning the CEO, COO, CIO, CTO, and CFO. Line-of-business executives were cited by 27% of respondents as decision makers but were included as budget holders by only 8% of respondents.

Two-thirds of respondents told us that their organization had a center of excellence (COE) for task and process automation. This number rose to 71% for transformers and 78% for disruptors. Banking/investment firms are also more likely to have established a COE.

The average expected increase in spending for content intelligence technologies over the next year will be 31%. Not surprisingly, the biggest expected increase in spending among all maturity groups is with laggards, whose average expected increase will be 44%. The largest expected increase in CI technology spending among business sectors will be with banking/investment firms at 39%.

Content Intelligence and Intelligent Process Automation

IDC defines intelligent process automation as the group of software technologies, including content intelligence, that individually or collectively manage, automate, and integrate processes. IDC forecasts that as a group, the IPA software market will grow from $13.1 billion in 2019 to $20.7 billion in 2023 at a compound annual growth rate (GAGR) of 12.3%.

CI is an important part of the IPA portfolio because it automates the classification of documents and the extraction and validation of data from those documents. That structured data can then be consumed by other automation technology in the portfolio, or other enterprise applications, reducing processing time, minimizing errors, and freeing up human resources from repetitive tasks.

One of the fast-growing segments within IPA is robotic process automation (RPA). RPA software is designed to automate or augment manual repetitive tasks. Those tasks may be individual tasks executed by a knowledge worker outside the context of a business process or may replace a task that is currently manually performed by process participants in an enterprise application. The RPA market is projected to grow from $1.3 billion in 2019 to $3.9 billion in 2023 at a CAGR of 36%. Since many of the processes and tasks that are well suited for automation by RPA are document- and content centric, content intelligence technologies frequently go hand in hand with RPA in intelligent process automation use cases.

Customer relationship management (CRM), business intelligence/analytics solutions, and risk management systems are the most common enterprise applications that are integrated with content intelligence technology (see Figure 6).

Insurance organizations cite compliance or eDiscovery solutions (38%) more than their peers, and BPO organizations cite business process management (44%) far more than other industries.

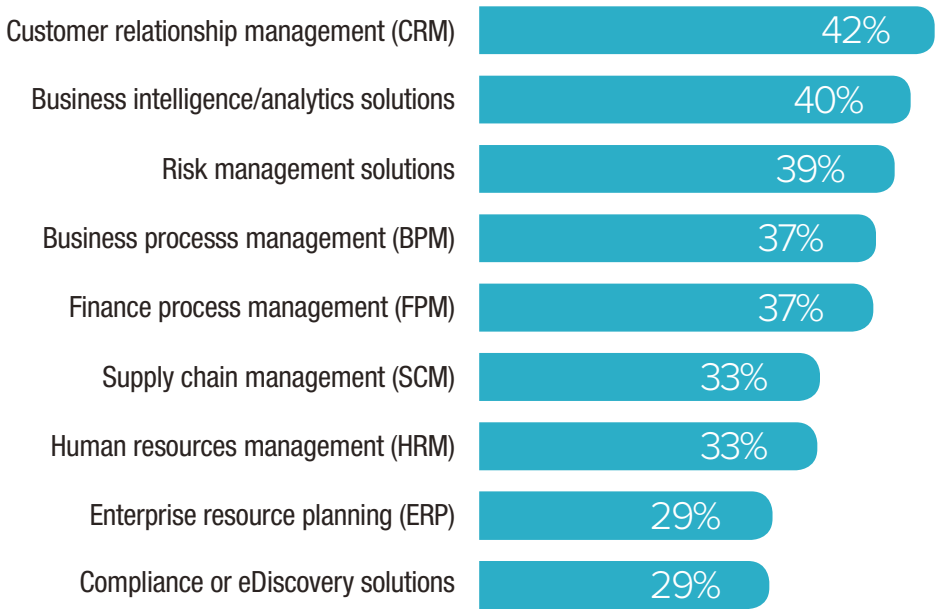
Similarly, a variety of different horizontal and vertical business applications have been automated (i.e., the processes run without human intervention unless there are exceptions) or augmented (the technology assists humans in completing the process) via integration with content intelligence technologies. The most frequently mentioned processes by survey participants were Remittance processing, customer onboarding, and employee benefit selection (see Figure 7).

There are industry differences here as well. Transportation/logistics organizations have automated sales order processing (48%) more than their peers, manufacturing has focused more on contract management (47%), and insurance is naturally highly focused on insurance underwriting (43%).

Q.

What are the top applications that consume (or will consume) the structured data generated by content intelligence technologies?

FIGURE 6 Enterprise Apps That Consume CI Output

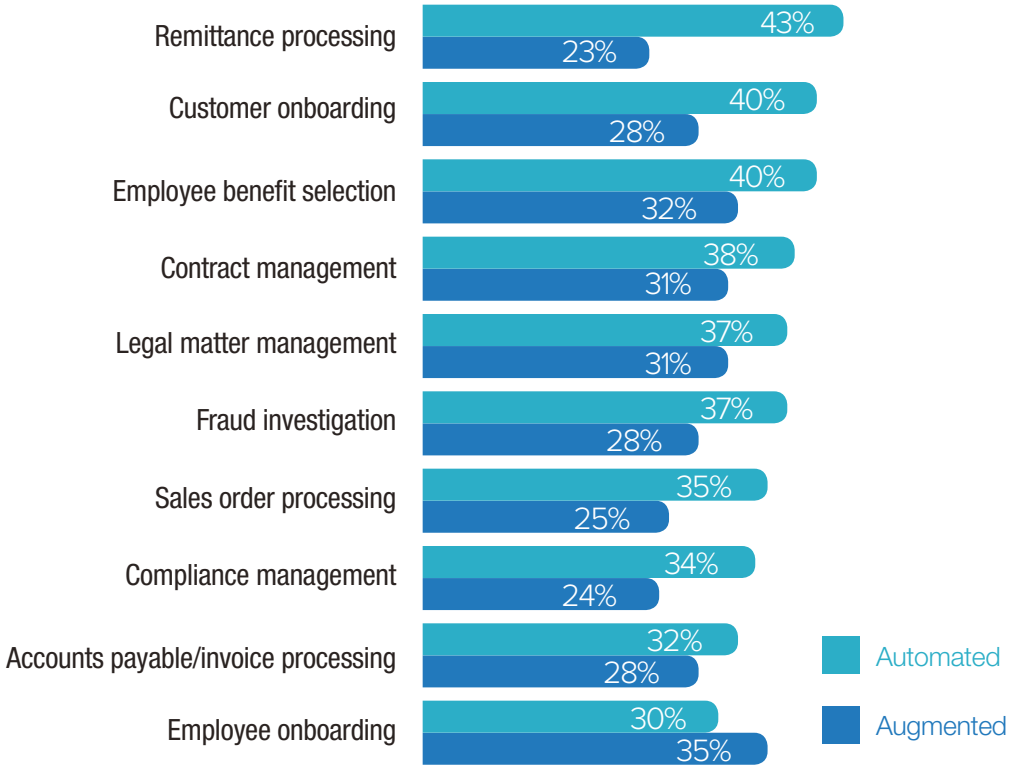


n = 425 Source: Content IQ Thought Leadership Survey 2019, IDC, May, 2019

Q.

What specific document-intensive business processes has your organization automated by deploying content intelligence technologies? Which processes have been augmented by deploying content intelligence technologies?

FIGURE 7 Processes Automated and Augmented by CI



n = 425 Source: Content IQ Thought Leadership Survey 2019, IDC, May, 2019

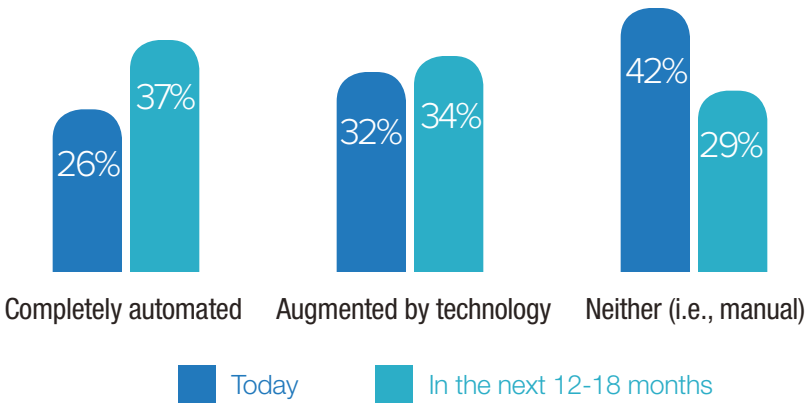
FUTURE OUTLOOK

What does the future of work look like? According to our survey respondents, a significantly higher percentage of processes are expected to be completely automated as well as augmented in the next 12–18 months compared with today, and manual processes are expected to significantly decline (see Figure 8).

Q.

What percentage of the processes in your organization are completely automated versus augmented today? What percentage in 12–18 months?

FIGURE 8 Process Automation: Today and In the Next 12–18 Months



n = 500 Source: Content IQ Thought Leadership Survey 2019, IDC, May, 2019

Perhaps more compelling is growth of the digital workforce. A growing number of employees will find themselves working side by side with a digital coworker in the future as technology automates many work activities. For example, the percentage of activities related to evaluating information performed by technology will grow by 28% in just two years, and 18% of activities related to reasoning and decision making will be performed by machines (see Table 1).

It is critical to understand that this transition is not "one and done." Digital transformation and work transformation will be ongoing initiatives, and intelligent process automation will continue to evolve, becoming even more "intelligent," driving growth of the digital workforce.

Successful leaders will guide empowered, agile organizations that embrace change and enable a dynamic, collaborative workforce that includes and empowers both human and digital workers and that provides a modern, integrated intelligent platform that can be scaled across the enterprise and enables the workforce to contribute effectively productively and securely.

TABLE 1 Activities Performed by Technology: Now and in the Next Two Years

Q. *For each of the following work-related activities, what portion (%) is accomplished by a machine or computer (versus a human worker) in your organization today? And in two years?*

	Today	In the Next Two Years	Change (%)
Evaluating information	13.87	17.71	27.7
Reasoning and decision making	15.01	18.07	20.4
Administering	13.85	16.45	18.7
Performing physical work activities	15.28	18.06	18.2
Communicating and interacting	15.09	17.75	17.7
Finding and identifying information	15.42	18.13	17.6
Performing complex technical activities	17.97	21.13	17.5
Performing physical work activities	15.88	18.63	17.3
Data processing	17.67	19.95	12.9
Authoring-related activities	15.15	16.50	8.9

n = 500 Source: Content IQ Thought Leadership Survey 2019, IDC, May, 2019

CONCLUSION

Work transformation — including the way that organizations digitize, automate, and transform content-centric workflows — is a foundational component of an organization's overall digital transformation strategy. This is particularly true as these initiatives progress from tactical and siloed improvements in productivity and worker experience to more strategic operational improvements that scale across the enterprise to improved decision making and insight— where disruptive transformation occurs.

To fully leverage the content intelligence opportunity, organizations must:

- Invest in modular, enterprisewide digital platforms that can integrate with existing enterprise applications and support a broad range of use cases.
- Develop a strategy for ongoing training and development of employees to ensure that they have the required digital skill sets.
- Think human and machine. The human-machine collaboration is not just the future of work, but it is the new normal for today's high-performing enterprises.

DEFINITIONS

- **Digital worker:** A digital worker refers to technology — including artificial intelligence, intelligent process automation (including content intelligence), robotics, and augmented reality and virtual reality — that performs tasks, jobs, and activities previously accomplished by a human worker.
- **Content intelligence (or "content IQ"):** Content intelligence is a set of technologies and services that leverage artificial intelligence to carry out tasks such as reading and categorizing a document, routing a document,

extracting and validating data from documents, and other tasks related to understanding and processing unstructured content to enable quicker, more accurate decision making and deliver greater business value. Relevant technologies include optical character recognition (OCR), computer vision, machine learning, natural language processing, and content analytics:

- **Artificial intelligence** is the science and engineering of making intelligent machines (especially intelligent computer programs); sometimes AI is about simulating human intelligence.
- **Computer vision** is the automatic extraction, analysis, and understanding of information from images, including scanned documents.
- **Content analytics** is software that analyzes, extracts, and organizes structured data from unstructured information and uses that structured data to find, locate, and provide answers and pertinent information to knowledge workers. Content includes text and rich media (such as audio, video, and images).
- **Machine learning** is software that enables machines to "learn" in real time or over time, improving accuracy and performance. In a process involving capturing documents, the technology learns from potentially thousands of variations of documents such as processing invoices or handling vendor orders.
- **Natural language processing** is the understanding of human language that enables software to read, interpret, and create structured data around unstructured content.
- **OCR** is technology that converts an image to readable text. This technology may be rules based or may employ machine learning and other cognitive technologies for better accuracy.
- **Unstructured content** is a content that is not contained in a database or some other type of data structure. Unstructured content can be text or rich media. Examples include Word documents, email messages, PowerPoint presentations, instant messages, JPEG images, MP3 audio files, and Flash video files.

- **Content IQ skills:** A Content IQ skill represents AI technology that provides the vision, understanding, and insight for processing unstructured content that can be quickly consumed as a service giving a digital worker (see the digital worker definition in the previous bullet point) the ability to carry out content-related tasks.

  We have defined five levels of content intelligence maturity:

  - **Content IQ laggard.** There are ad hoc efforts to deploy content intelligence, but there is no overarching strategy, funding, or support.
  - **Content IQ apprentice.** There are isolated content intelligence initiatives in some functional areas with limited integration, collaboration, or standardizations.
  - **Content IQ practitioner.** There is an enterprisewide strategy for managing unstructured content, and a broad portfolio of content intelligence technologies is deployed.
  - **Content IQ transformer.** There is full integration with enterprise applications and systems of engagement, extending to external stakeholders.
  - **Content IQ disruptor.** Content intelligence technologies support orchestration and decisioning within content-centric workflows, including information governance.

About IDC

International Data Corporation (IDC) is the premier global provider of market intelligence, advisory services, and events for the information technology, telecommunications and consumer technology markets. IDC helps IT professionals, business executives, and the investment community make fact-based decisions on technology purchases and business strategy. More than 1,100 IDC analysts provide global, regional, and local expertise on technology and industry opportunities and trends in over 110 countries worldwide. For 50 years, IDC has provided strategic insights to help our clients achieve their key business objectives. IDC is a subsidiary of IDG, the world's leading technology media, research, and events company.

IDC
ANALYZE THE FUTURE

Global Headquarters

5 Speen Street
Framingham, MA 01701
USA
508.872.8200
Twitter: @IDC
idc-community.com
www.idc.com

Copyright Notice

Document #US45582119