Running at the Speed of Digital: Hyper-Digital Information Management

As digital becomes more ingrained in everyday life, information services providers will need to fundamentally change how they manage and deliver content to increase operational flexibility, reduce costs, and closely align with customers’ needs.
EXECUTIVE SUMMARY

Today’s customers expect information to be readily available through whatever channel or channels they prefer, at a time and place of their choosing. To meet these expectations and remain relevant and profitable, information services (IS) companies must run at the speed of digital.

At the same time, the industry is hindered by the lack of an automated, integrated platform that can support organizations’ requirements across the information value chain. This makes it difficult to effect digital transformation – leaving many companies struggling as they attempt to control and streamline their business processes.

This white paper charts the current state of the IS industry, the common challenges faced by enterprises handling large volumes of information, and the impact on the end customer. We also present a platform-based strategy for building a robust, automated IS management ecosystem, as well as an approach for transforming companies’ existing systems into an integrated, hyper-digital environment that improves and synchronizes processes and optimizes costs.
MANAGING INFORMATION IN THE DIGITAL AGE

Consumers’ growing need for instantaneous information through multiple channels is driving IS organizations to overhaul their information management systems. This isn’t an easy task, since information formats can be structured or unstructured. Plus, the volume and velocity of information generated on a daily basis both in digital and non-digital forms is huge (hence the term “big data”).

These dynamics create a variety of challenges IS companies must overcome in order to make sense of and act on information that customers require and consider relevant. The focus is on unlocking the value in information and delivering it to consumers as quickly as possible, in a format and through the channel they choose.

The global IS market was valued at roughly USD 1.5 trillion in 2016, and is expected to grow at a compound annual growth rate (CAGR) of 5.2%.

Even so, industry players are hindered by the lack of an integrated, automated platform for end-to-end data processing, information delivery, and digital transformation. While several companies in specific IS industry segments offer extremely powerful capabilities, none provide the full range of technologies and resources needed to support a cohesive, all-inclusive, digitally-equipped environment for analyzing, ingesting, managing, and delivering content across the value chain (see Figure 1).

In today’s increasingly digital world, deconstructing high-volume collated information manually is virtually impossible. Add to this challenge continually changing consumer behaviors, which influence enterprise strategies and decisions, and make it harder to ensure that consumers get the content they want when and where they want it. In a hyper-digital environment, companies can deliver content at breakthrough speeds – anytime, anywhere.
INDUSTRY OVERVIEW & OUTLOOK

As noted earlier, data flowing through a processing system can be structured or unstructured.

Structured information can include financial data, company information, and processed documents. Unstructured data can be in the form of e-mails, voice recordings, and images. Figure 2 provides a snapshot of some of these data types. Figure 3 highlights high-volume information categories across various industries.

Consumer expectations — including instant delivery of relevant information — generated the term “hyper-digital.” Likewise, bringing information to digitally inclined, increasingly demanding consumers at the right time and place is at the heart of hyper-digital processes and applications (see Figure 4, next page).

A Range of Information Sources

<table>
<thead>
<tr>
<th>Traditional/ Legacy Content</th>
<th>New Age/ Digital Content</th>
</tr>
</thead>
<tbody>
<tr>
<td>Documents</td>
<td>IOT</td>
</tr>
<tr>
<td>Paper-Based</td>
<td>Biosensors Data</td>
</tr>
<tr>
<td>Photos</td>
<td>Social (Blogs, Tweets, Posts)</td>
</tr>
<tr>
<td>Published Data &amp; Statistics</td>
<td>Web Logs</td>
</tr>
<tr>
<td>Published Research</td>
<td>Internet</td>
</tr>
<tr>
<td>Corporate Systems &amp; information</td>
<td>Apps</td>
</tr>
<tr>
<td>Archives</td>
<td>E-Mail</td>
</tr>
<tr>
<td>Transactional</td>
<td>Audio / Video</td>
</tr>
<tr>
<td></td>
<td>Mobile</td>
</tr>
</tbody>
</table>

Figure 2

The Information Glut Across Industries

Figure 3
For example, publishing articles in science, technology and medical (STM) journals involves a number of steps (see Figure 4). Peer reviews alone can take between 100 to 150 days depending on the subject area. However, the overall time it takes to produce and publish content is less, since the process consists primarily of repeatable tasks. Over the last decade, publishers have been able to reduce production time from 60 days to around 20 days due to modern digital content management software and platforms (see Figure 5).

The STM Publishing Process

For example, publishing articles in science, technology and medical (STM) journals involves a number of steps (see Figure 4). Peer reviews alone can take between 100 to 150 days depending on the subject area. However, the overall time it takes to produce and publish content is less, since the process consists primarily of repeatable tasks. Over the last decade, publishers have been able to reduce production time from 60 days to around 20 days due to modern digital content management software and platforms (see Figure 5).

The amount of cross-border bandwidth used has grown 45 times since 2005. It is projected to increase by an additional nine times over the next five years as waves of information, searches, communications, videos, transactions, and intra-company traffic continue to surge.

In the U.S. healthcare industry, for example, it is estimated that 6.4 billion pieces of data are present in static images, paper, and document repositories, and that on average, it costs $1.34 to process a form or document.
In our view, it is now possible to trim the publication process even more through an integrated information management platform.

INDUSTRY PAIN POINTS & KEY TRENDS

Content-intensive industries such as publishing, information services, insurance and healthcare face some tough challenges:

- **Outdated technology and organizational silos:** Across industries, many organizations are challenged by how to make technology work effectively and seamlessly on the large content sets on which they sit. Enterprises are often reluctant to embrace technology changes, since doing so restricts their ability to innovate and increases costs. Organizational and technology silos present other bottlenecks that hinder progress in the digital age.

- **Business/operating model changes:** Business and operational modifications in information-intensive industries typically result in increased work volumes, correspondingly higher fees, and more content to manage due to:
  - The growing number of content sources and data providers, including the Internet of Things (IoT) and rich media.
  - Increased segmentation in the nature and delivery of information services.

- **Multiple delivery channels:** The number of channels through which information is delivered also complicates where and how information must be tailored to fit each and every delivery touchpoint.
• **Free content:** The availability of free information online has impacted the number of consumers willing to pay for it.

These issues can create difficult challenges for consumers. For example:

• Increased information management costs lead to higher subscription fees for consumers. In the legal services industry, which is primarily information-driven, average charges range from $200 to $600 – sometimes more per hour for services rendered.\(^6\)

• An exorbitant amount of time and effort is required to identify relevant information due to the large volumes of uncategorized data.

• Analysts, on average, spend nine hours\(^7\) per week breaking down information (see Figure 6).
  » Expectations with respect to speed are rarely met.
  » Identifying relevant information is difficult.

These problems circle back to affect enterprise information management activities:

• **Systems work in isolation.** Building a connected digital ecosystem is key to profitable growth. Information-led business prosperity can only be achieved through a well orchestrated, highly efficient information supply chain and associated systems. Otherwise, operational costs increase – a burden that customers must also bear.

• **Information consumption exhausts information sources.** Information can be digital and non-digital in nature – structured, semi-structured, and unstructured. Collating these assets in a timely, accurate manner helps avoid obsolescence, but takes more time and manpower than many businesses can afford.

• **Information delivery has become more complex.** Multiple delivery channels underscore the need for customized information across customer touchpoints.

• **The volume and speed of data is increasing exponentially.** Information management platforms should be capable of handling large amounts of data at ever-higher speeds – efficiently, and at the lowest possible cost.

### Analyzing Information

![Analyzing Information Diagram](image)

**Figure 6**
Information-led business growth requires a well orchestrated, highly efficient information supply chain and associated ecosystems.

**The Information Value Chain**

<table>
<thead>
<tr>
<th>INGESTION</th>
<th>ENRICHMENT</th>
<th>DELIVERY</th>
</tr>
</thead>
<tbody>
<tr>
<td>Authoring</td>
<td>Entity Abstraction &amp; Recognition</td>
<td>Assembly &amp; Packaging</td>
</tr>
<tr>
<td>Content Extraction</td>
<td>Validation</td>
<td>Format Conversion</td>
</tr>
<tr>
<td>Collation</td>
<td>Editorial Process</td>
<td>Template Management</td>
</tr>
<tr>
<td>Cleansing</td>
<td>Ontology/Taxonomy Management</td>
<td>Multichannel Publishing</td>
</tr>
<tr>
<td>Preliminary Categorization</td>
<td>Reference Management</td>
<td>Reporting &amp; Analytics</td>
</tr>
<tr>
<td>Format Management</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Figure 7**

- **Discovering relevant information is increasingly difficult.** Every day, the amount of information available for consumption grows. This makes it all the more necessary to properly identify, categorize and store data managed by enterprises in a media-neutral form that makes it easy to identify and retrieve.

**DEVELOPING AN INFORMATION MANAGEMENT PLATFORM**

An integrated, comprehensive information management platform should offer capabilities that apply across the value chain (see Figure 7). The platform should also be supported by technologies and tools throughout the information lifecycle (see Figure 8, next page).

As shown in Figure 9 on the next page, a hyper-digital enterprise uses automated tools, techniques, and technologies to support crucial capabilities:

- **Ingestion:** Using template-based systems to read, parse, and collate incoming information from multiple providers can significantly reduce the time it takes to process, store, and ingest data flowing through these “data factories.” Image and entity recognition play a major role here – enabling systems to automatically identify and categorize pictures and videos. The ingested content is typically converted into a machine-readable format – identifying issues and flagging them for resolution. This task can also be automated. Artificial Intelligence is increasingly used for information selection across multiple sources, at a far faster rate.

- **Enrichment:** Natural language processing (NLP) and machine learning are being used for operational enrichment, which involves capturing, analyzing, categorizing, and enhancing information to
make it easily retrievable. Based on linguistic indicators (phrases and their meanings) and context cues, the system can be taught to recognize and categorize textual units automatically. It can also identify topics and extract relationships automatically. Enterprises can build their own ontologies to suit their business domains while concurrently experimenting with various levels of granularity. The results of automated categorization are validated to further strengthen the system. This can be applied to various information types, such as written, oral, audio-visual, and imagery.

Data lakes provide a single source for storing enterprise data — from raw to transformed — and converting it from one format to another as needed. Ontology reduces the complexity of organizing information. It is just one of many methods for analyzing data; new tools and techniques further refine how information is collected, dissected, and acted upon.

Support Tools Across the Information Value Chain

<table>
<thead>
<tr>
<th>INGESTION</th>
<th>ENRICHMENT</th>
<th>DELIVERY</th>
</tr>
</thead>
</table>
| • Web scraping & spidering tools  
• Automatic monitoring tools and alert mechanisms  
• OCR and translation tools  
• Authoring tools | • Information extraction engine  
• Editorial platforms  
• Semantic text analytics toolkit  
• Master data management solutions | • Integrated workflow tools and APIs  
• Publishing platforms  
• Multi-platform and format delivery tools  
• KPIs for outcome measurement |

Hyper-Digital Enterprise Platform Modules

**COLLECTION & CONVERSION (ETL)**
Authoring, ingestion from various sources, and conversion with the help of other supporting modules. Rules management with input rules for content flowing in and output templates for information published.

**MDM SERVICES**
Information about entities handled and the ontologies, taxonomy models, and schema used for identifying and categorizing the information.

**CONTENT ENRICHMENT**
Content modifications and updates to ensure relevance, and a touchpoint for maintaining the domain model and schema.

**CONTENT REPOSITORY**
Heart of the system storing the final converted content in media-neutral form. Used to deliver information to consumers after conversion to appropriate format.

**INFORMATION DELIVERY**
Content conversion on the fly to suit the needs of the consumer based on the platform, delivery channel, device, and needs of the entity consuming the information.

**ANALYTICS & REPORTING**
Support consumers in analyzing and identifying insights from the stored data. Also assist search and retrieval through appropriate techniques.
• **Storage & Delivery:** As additional data is collected, costs to maintain crucial infrastructure, back-ups and archives increase. Hence, enterprises are moving to the cloud to take advantage of the cost benefits provided by services such as Amazon Web Services (AWS), Google, and Microsoft Azure. Currently, original content is stored as is, then converted to a neutral form that is both machine- and human-readable. This has multiple advantages:
  » Information can be generated on the fly in an appropriate format to be distributed across multiple delivery channels. Newer formats can be added just by defining the output template, helping to future-proof the platform. Data need not be stored in different formats – reducing storage costs.
  » The use of detailed, multi-layer metadata makes it easier to search for and retrieve information, and improve its relevance.
  » Information in this form easily supports analytics and reporting.

**QUICK TAKE**

Observations from the Field

Based on our client engagements and research on information management, we see the following scenarios playing out across industries:

• Enterprises agree that auto-classification using content analytics is the only way to bring content clutter under control.

• Workflow, content analytics and enterprise search capabilities have a strong net demand; at the same time, users find the convergence of collaboration and social tools confusing.

• Roughly 27% of users who do not utilize the cloud are likely to deploy cloud content systems within the next two years, another 21% within the next four years.

• Approximately 52% of enterprises are looking to build a company-wide ECM system.

• Applications such as records management, scanning and capture, HR, and AP/AR already reside on the cloud.
KEY PERFORMANCE INDICATORS

Enterprises that are developing an information management strategy need to measure their performance (see Figure 10). The following KPIs should be considered when evaluating any solution or product for this purpose.

- Time and effort spent on information processing operations.
- The volume of information processed in a specific time period.
- Wait times for data processing without human intervention.
- Information maintenance costs and efforts.
- The ability to deliver the required information to consumers when and where needed.

These metrics can help identify areas of improvement and opportunities for increasing process efficiency. They can be further sub-categorized into metrics that cover the main stages of the information value chain – from ingestion through delivery.

BENEFITS TO THE ENTERPRISE

Figure 11 highlight the benefits of a hyper-digital enterprise information management platform.

Enterprise Platform Performance Modules
Hyper-Digital Information Management

**Platform Approach to Information Management**
Eliminate silos and bottlenecks.
Ensure seamless integration of business processes.

**Automation of Ingestion, Conversion & Categorization**
Increase volume of information processed.
Identify and resolve errors quicker.

**NLP and Machine Learning for Enrichment**
Reduce operational costs due to less manual intervention.
Quickly discover information patterns.
Improve entity recognition and resolution.

**Cloud Storage in Media Neutral Form**
Low cost, high availability, and better performance.
Rapid delivery of information to consumers.
Quick onboarding of new formats and templates.

**ENTERPRISE BENEFITS**

**LOOKING AHEAD**
For organizations considering an integrated enterprise information management platform, we suggest either of the following approaches:

- **Build-integrate:** Enterprises can build individual modules to support their specific business processes, then integrate them with current systems and applications. This approach gives organizations more flexibility when determining the level of functionality needed, but can be more costly and take more time to build.

- **Buy-integrate:** With this approach, enterprises have the option to buy off-the-shelf products that support some portions of the value chain, customize them to suit the organization’s needs, then integrate them with existing systems. This approach costs less but may not provide all of the features required to support key processes.

There may not be a “one-size-fits-all” solution for information management – compelling companies to create “homegrown” ecosystems over time to suit their evolving business needs. The strategies highlighted here can serve as a good starting point for any organization taking its first steps to becoming a hyper-digital enterprise.
Digital Business

FOOTNOTES


3 Ibid.


9 Ibid.
ABOUT THE AUTHORS

Shyamkrishnan Balasubramanian
Consultant – Communications, Media & Technology Consulting Practice, Cognizant

Shyamkrishnan Balasubramanian is a consultant within Cognizant’s Communications, Media & Technology Consulting Practice. He has eight years of experience providing business solutions to clients in the IS industry. His areas of expertise include information services, data and content operations, business platforms and enablers. His experience spans domains – from information services, publishing, and broadcasting, to media and entertainment. He has an engineering degree from VIT, Vellore, and received an MBA from BIM, Trichy. He is also a certified business analyst and helps train and mentor junior business analysts as part of workshops and training sessions. He can be reached at Shyamkrishnan.Balasubramanian@cognizant.com | www.linkedin.com/in/shyamkrishnanb/.

KS Venkataramanan
Manager – Communications, Media & Technology Consulting Practice, Cognizant

KS Venkataramanan is a consulting manager within Cognizant’s Communications, Media & Technology Consulting Practice. He has 13 years of business consulting experience across the information services and publishing industries in multiple geographies. His other areas of expertise include portfolio analysis, new product development, digital strategy, solution envisioning, program management, process consulting, and business development. He is an electronics graduate and has completed his MBA from Bharathidasan University, Trichy. A certified scrum master (CSM) and SAFe Agile certified, he can be reached at Venkataramanan.ks@cognizant.com | www.linkedin.com/in/venkataramanan-ks-b737648/.

Krishna Durbha
Senior Director – Communications, Media & Technology Consulting Practice, Cognizant

Krishna Durbha is the Senior Director of Cognizant’s Communications, Media & Technology Consulting Practice. He has over 24 years of experience, has built business consulting practices in South East Asia, and led large business transformations programs for B2B and B2C engagements. He has set up profitable digital properties in Internet and mobile value-added services that involved product management, platforms, applications, content procurement, digital content management platforms, and mobile advertising. An engineer from NIT, Surat, he holds an MBA from IIM, Ahmedabad, an executive MBA from Leeds University Business School, and is PMP certified. He can be reached at Krishna.Durbha@cognizant.com | www.linkedin.com/in/kdurbha/.
ABOUT COGNIZANT

Cognizant (NASDAQ-100: CTSH) is one of the world’s leading professional services companies, transforming clients’ business, operating and technology models for the digital era. Our unique industry-based, consultative approach helps clients envision, build and run more innovative and efficient businesses. Headquartered in the U.S., Cognizant is ranked 205 on the Fortune 500 and is consistently listed among the most admired companies in the world. Learn how Cognizant helps clients lead with digital at www.cognizant.com or follow us @Cognizant.