

View Point



Toward a Unified Enterprise Intelligence

Daljinder Singh Uberoi

Executive Summary

The evolution of business intelligence has moved from data mining of structured sources toward enterprise data mining across structured and unstructured data sources within an organization.

While most businesses have sizeable investments in Business Intelligence tools and applications, these traditionally address only the structured data that resides in databases. Recent efforts to address unstructured data and build intelligence around it have created pockets of intelligence. Organizations need to put a strategy in place to leverage intelligent information access and analytical capabilities on enterprise data – structured as well as unstructured.

This Infosys paper traces this evolution from the pockets of intelligence to Enterprise Intelligence. It also provides insights into two key strategies available to achieve this objective.

Enterprise Intelligence: The Evolution

Enterprises have come a long way in their quest to better manage information, bringing in efficiencies and intelligence into businesses for competitive advantage. The journey began with integrating islands of information from a multitude of diverse applications to introduce operational efficiencies, creating what are known as Enterprise Applications. The next step was to analyze the loads of information and data made available through these applications bringing intelligence to strategic as well as operational business decisions.

Although there is a realization of the need along with common understanding on the issues involved and the steps needed to address them, many enterprises end up trying to hammer square pegs into round holes in the implementation.

Conventional business intelligence (BI) tools enable multidimensional views of **structured** data stored in the databases, allowing users to slice or dice data in the offline mode without concern for the impact on the transaction data and performance. However this focus has been on structured data only.

Unstructured data a.k.a 'content' is constantly being generated within the organization in multiple forms such as:

Documents

- Word processor documents, spreadsheets, reports

Web content

- Content shared internally and with the outside world through various web sites

Digital content generated through various communication channels

- Emails, text and voice chats, discussion threads, etc.

Other digital assets

- Text, audio clips, video clips etc.

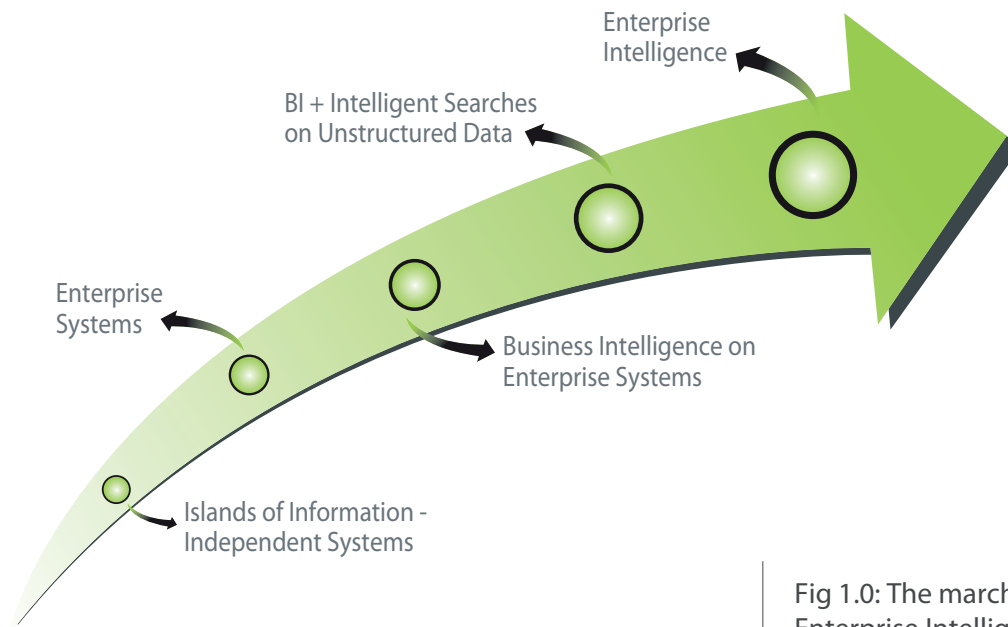


Fig 1.0: The march toward Enterprise Intelligence

Unstructured data, spread across the organization, despite its importance and value, is untapped by most BI tools. To mine the potential of this huge resource base, it is essential to:

- Recognize the existence of this unstructured data and the fact that everything cannot be moved to databases
- Understand its importance and the need to loop such unstructured data/content sources into the cycle of drawing information for business or operational intelligence

Lately, there are visible indicators of progress on this front. Some organizations are taking cognitive steps to provide shelter to the homeless entity called unstructured data and build intelligence around it. While this has generated significant benefits for these organizations, in the process it has created [pockets of intelligence](#). With this step in the evolutionary process, the march toward Enterprise Intelligence is now truly under way (Fig 1.0).

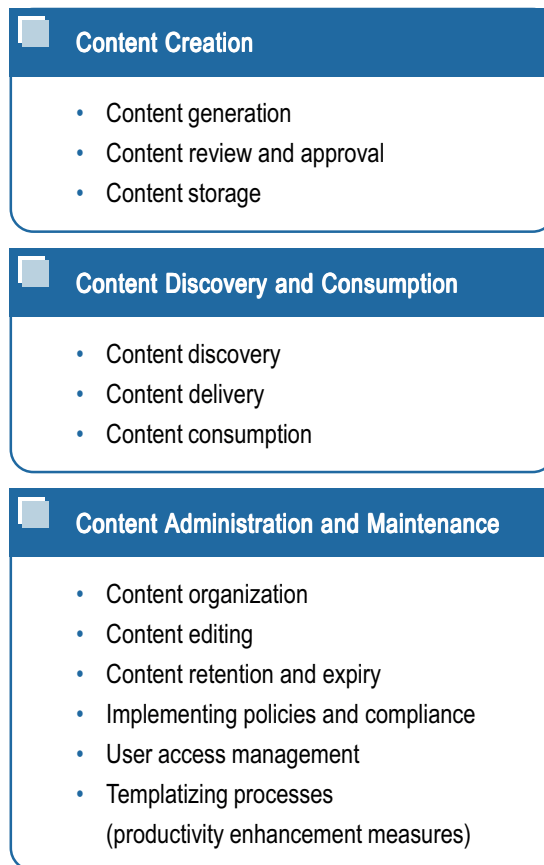
The evolution of Enterprise Content Management (ECM): Shelter for the homeless entity called Unstructured Data

Enterprise Data = Structured Data + Unstructured Data

Businesses now recognize the importance of both types of data – structured as well as unstructured – and the need to search and consume that information for competitive advantage. The traction around content management solutions and the overwhelming success of search engines stand testimony to this fact.

[Content Management Solutions \(CMS\)](#) have come a long way in addressing unstructured data as have [search technologies](#) that now offer flexibility to bring relative intelligence into the results. CMSs now offer a home to unstructured data and build a structure around it. Intelligent search engines, meanwhile, are striving to provide the route map and a vehicle to reach that home address.

[Content Management](#) includes the technologies and tools to manage content and different work-actions associated with content pertaining to organizational needs and processes. Work-actions associated with content across its lifecycle stages are*:



* Please see Appendix I for definitions of the terms used above

Need for effective management of unstructured data at enterprise level: An important aspect to consider here is that content/ unstructured data and the associated workactions are not limited to the individual user level. Being scattered across the enterprise (worker desktops, various applications, shared repositories, etc) makes it imperative to effectively manage this unstructured data across the enterprise.

Viewing each of these work-actions at the enterprise level, the number of touch points across various stakeholders grows exponentially. Since enterprises no longer work in pockets, information/ data needs to be available in real time to different parties involved. This makes the need for effective enterprise level collaboration evident at each stage of content generation, review and approval, consumption, etc.

Compliance brings in another dimension: Legal and regulatory compliance needs have brought in a new dimension to data handling and associated work-actions on this enterprise data. Both structured and unstructured data need to be tracked, audited, retained, retired and mined at the enterprise level in a holistic manner to be able to comply with regulatory needs.

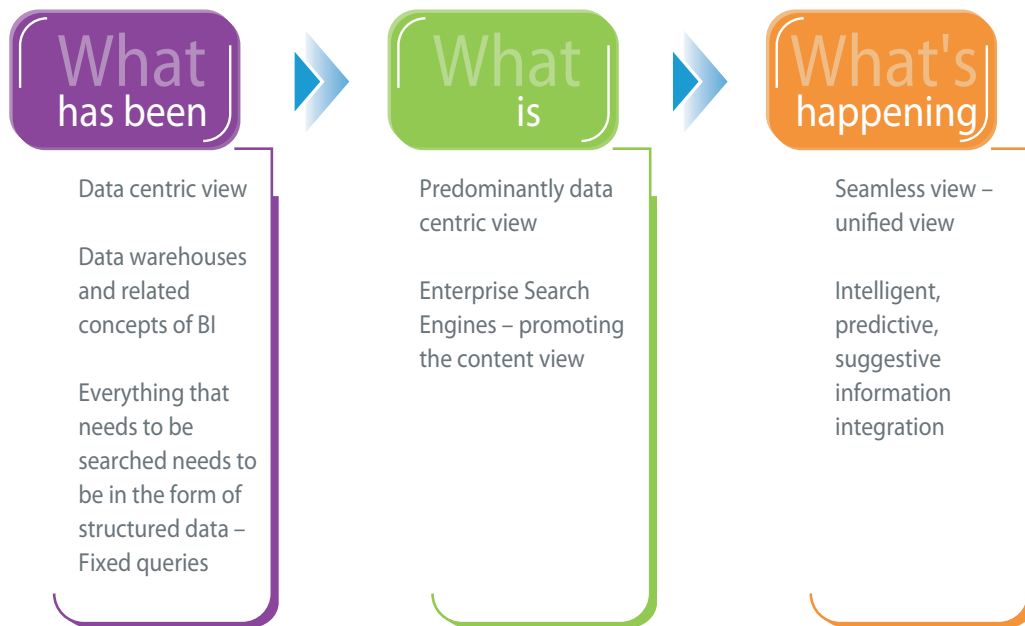


Fig 2.0: The evolution of Enterprise Intelligence

Enterprise Content Management Capabilities (ECM): Enterprise Content Management solutions available today provide capabilities for effective management and collaboration of unstructured content across the enterprise. This can, however, create two unconnected managed worlds:

- Managed world of structured data
- Managed world of unstructured data

Both these ‘worlds’ have their respective solution offerings.

Enterprise Information Management and Enterprise Intelligence

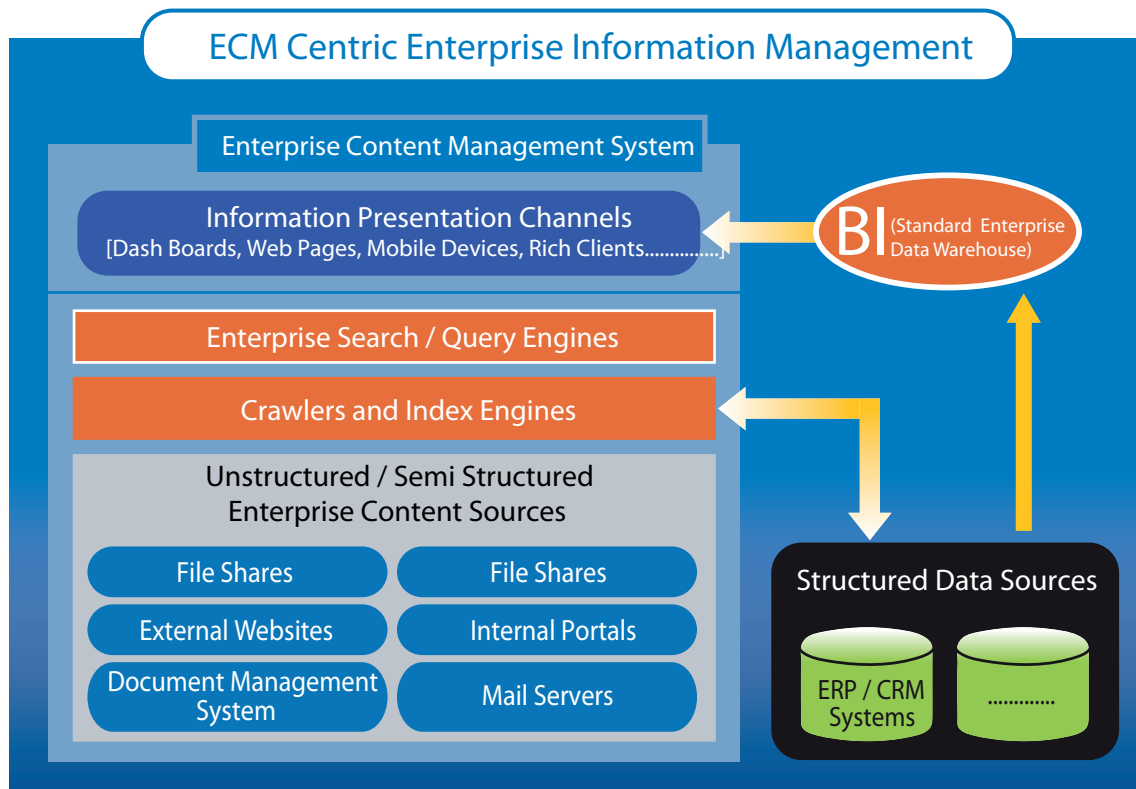
The world of managing structured data has been around for some time now and has fairly matured solution offerings. The other world, i.e., Enterprise Content Management solution offerings, is maturing rapidly.

What organizations now need is an enterprise view of information management, across both structured and unstructured information. Enterprise Information Management and Enterprise Intelligence are the efforts to bring these two worlds together to provide a consolidated view.

Enterprise Intelligence Strategy Options

Driven by the historic offerings of each vendor, there are different approaches available to achieving Enterprise Intelligence. Structured data world vendors are investing in providing intelligence on non-data content, whereas content management vendors are providing features to plug in contextual information and intelligence from the different data sources and applications into content management portals and dashboards. Predominantly, the landscape can be segmented into two approaches. These are:

- Unstructured data / enterprise content management driven approach
- Structured data driven approach

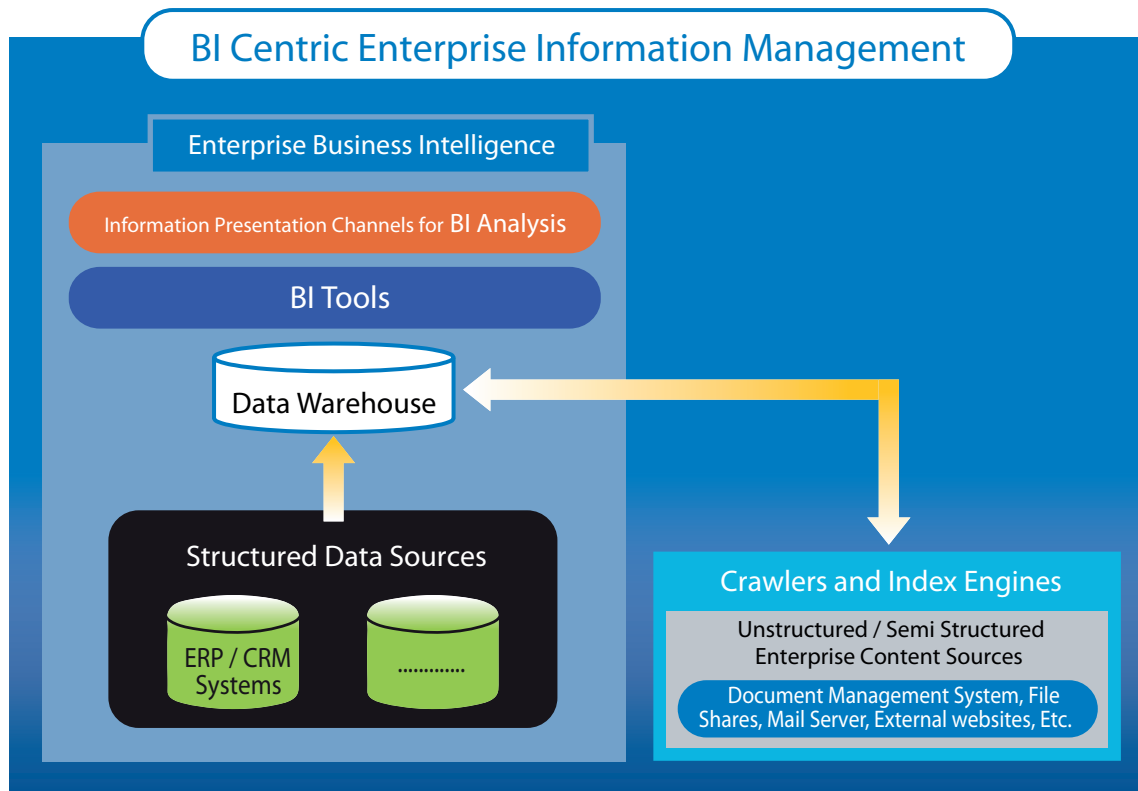


Unstructured Data / Enterprise Content Management Driven Approach

Vendors taking the content-driven approach have leveraged their Enterprise Content Management platform as the core offering and attempted to plug structured data into it. In the world of content management, besides the numerous content management features discussed above, content/information discovery plays a critical role. This approach uses 'crawlers' for unstructured content sources which leverage different proprietary algorithms. The content crawled is indexed and stored in index databases. These indexes are leveraged by query/search engines to intelligently collate, sort, filter, and arrange information depending on a user's search query. Search engines, through increasingly sophisticated algorithms, are now capable of highly complex pattern recognition.

Through the ECM-centric Enterprise Information Management approach, vendors are trying to leverage their existing investments and provide extensibility to achieve intelligent mining of structured data sources such as ERP, CRM or LoB systems. They have effected this by using crawlers to crawl and index these sources of information. When a user submits search criteria for the search engine, the engine leverages its intelligence to work not only on the unstructured and semi-structured data sources but also on the business application data, thus providing users appropriate information across the enterprise, easily and efficiently.

Another level of consolidated information management that ECM vendors are trying to achieve is providing a unified view of information in a contextual manner. This entails an ability to fetch information from unstructured data sources as well as the business applications based on a context and provided in a structured format. This can help improve productivity significantly. Further, some vendors are working on presenting information from conventional Business Intelligence and data warehouse on the same dashboard.



Structured Data Driven Approach

Business Intelligence solutions traditionally provide a local view to the management along with intelligence around the structured data in various business applications.

In the data driven approach, the BI platform is the core and the view on the unstructured data is plugged in to fulfill the enterprise level business intelligence needs.

In this approach, crawlers and index engines with inbuilt intelligent algorithms and pattern recognition capabilities are leveraged to bring structure to the unstructured data. This newly structured data is then loaded into the central data warehouse. This approach provides an opportunity to tie up the data in database-driven business applications, documents and other data stored in managed repositories. The BI tools can then be leveraged on this central data warehouse to obtain an enterprise view of the business intelligence.

The choice of approach depends on the organization

Both approaches are strong efforts to provide capabilities for an enterprise view of information management across structured and unstructured data. Both aim to bridge the pockets of intelligence within the enterprise. However, until these approaches converge for a seamless view of the worlds of structured and unstructured data, organizations will need to choose between the two.

Infosys believes that neither approach is better than the other. Both are equally important and effective. The Enterprise Information Intelligence strategy should be primarily driven by:

1. The predominant structure of the enterprise data/content within the organization – existing as well as in the future
2. How the organization intends to consume the intelligence generated on top of enterprise information

In addition, other factors such as turnaround time, current portfolio of investments in IT applications [operational, analytical and strategic], etc, may influence the choice of approach.

Next Steps

Based on the factors mentioned in the last section, organizations need to identify and align their Enterprise Information Management realization strategy. As per the key direction, i.e. ECM-centric or BI-centric, the next logical step would be to evaluate vendors and system integrators in the space.

For some organizations, the strategy alignment could be quite obvious – for others the question of choice could be fairly complex.

Take a leading law firm with multiple offices across the globe. For users predominantly working on documents and emails along with other channels of communication, unstructured data is the lifeline. However, for each of their clients, users also have billing and other information stored in databases and managed through respective applications. They need a consolidated view of clients while being able to draw on intelligence across cases.

This, then, is a clear case for implementing an ECM driven Enterprise Information Management Strategy.

Most other organizations, however, are likely to have a substantial mix of both structured and unstructured data. They may have certain processes working entirely on (and creating) documents and other unstructured data. At the same time, they may have sizeable investments in typical operational and business intelligence capabilities built on their systems.

Within the same organization, there can be different business functions with varying data or content structures. There could also be different expectations form Enterprise Intelligence capabilities.

For such organizations, it need not necessarily be either an ECM or a BI-driven approach. Specific functions could be ECM-driven, while others could be BI-driven.

The critical success factor for these organizations is to invest upfront in formulating a robust and comprehensive Enterprise Information Management and Enterprise Intelligence strategy. The implementation could be rolled out in a phased, evolutionary method prioritizing maximum business value areas or minimal risk zones.

Summary

It is evident that Enterprise Intelligence is evolving rapidly with new developments in both camps of Enterprise Content Management as well as Business Intelligence. Infosys sees both approaches maturing and possibly converging in the near future, and thus providing Enterprise Intelligence in its true sense. Until then, organizations can yet realize immense value from the current stack of options and capabilities available. A well-developed strategy to leverage structured and unstructured enterprise data into the cycle of drawing information for business or operational intelligence can help businesses stay ahead of the competitive curve.

Appendix I

Definitions of terms used

Content Creation

- Content generation: Generation of documents, spreadsheets, reports, multimedia content, etc., or web content that may be part of the internal or external facing web sites
- Content review and approval: Content may have to go through multiple cycles of review and approval to ensure that it is well written, usable, addresses IP issues, etc., before it is deemed fit for consumption
- Content storage: The content created needs to be stored in a repository from where it can be accessed and consumed

Content Discovery and Consumption

- Content discovery: Searching for the content that one needs to complete one's task
- Content delivery: Content can be delivered to the end user through both push and pull models and across multiple channels, such as email, mobile devices, publishing on online portal, etc.
- Content consumption: Using the content for the generation of further content (reports, proposals, references, etc.), as it is or in derived format

Content Administration and Maintenance

- Content organization: Assessment and organization of content relative to the audiences' leveraging taxonomy and categorization techniques
- Content editing: Tuning the content and also the style of delivery
- Content retention and expiry: Duration for which the content needs to be retained before it expires
- Implementing policies and compliance: Creation, maintenance and application of review standards in compliance with internal policies or regulatory needs
- User access management: Managing users, and user access rights to create, view, edit, and consume content
- Implementing authentication and authorization mechanisms
- 'Templatizing' processes [productivity enhancement measures]: Standard templates that can be applied to new and existing content

About the Author:

Daljinder Singh Uberoi is a Technical Architect at Infosys with over nine years of experience helping customers bridge the gap between technology and business needs. He has managed product development, worked on program management involving large software engineering groups, and has been responsible for delivering custom solutions on .NET besides implementation of eCRM solutions.

Questions? Opinions?

Discuss Enterprise Intelligence at <http://blogs.infosys.com/microsoft>

For more information, contact catalyticit@infosys.com



For more information, contact askus@infosys.com

About Infosys

Many of the world's most successful organizations rely on Infosys to deliver measurable business value. Infosys provides business consulting, technology, engineering and outsourcing services to help clients in over 30 countries build tomorrow's enterprise.

For more information about Infosys (NASDAQ:INFY), visit www.infosys.com.