Enterprise Content Management (ECM) Overview

A White Paper

Process, Power & Marine, a division of Intergraph

Abstract: Explores the extended market space traditionally covered by document management system vendors. Identifies the differentiating features between generalized ‘back-office’ systems (ECM) and specialized ‘engineering’ systems (EngCM). Indicates potential beneficial integration points between an ECM and an EngCM product such as SmartPlant Foundation.

Table of Contents

1. Introduction 3

1.1 "Making the case for Integration" 3

2. ECM Defined 4

2.1 ECM Capabilities 4

2.1.1 Content at Work 5

2.1.2 Compliance 5

2.1.3 Collaboration 5

2.1.4 Cost 6

2.1.5 Continuity 6

Paperless or Paper-efficient 6

2.1.6 Business Process Management/Workflow 7

2.1.7 Content and Documents 7

2.1.8 Scanning 7

2.1.9 Document Imaging 8

2.1.10 Forms Processing 8

2.1.11 Recognition 8

2.1.12 Categorization, Taxonomy 8

2.1.13 Indexing 8

2.1.14 Document Management 8

2.1.15 Records Management 8

2.1.16 Email Management 9

2.1.17 Web Content Management 9

2.1.18 Digital Asset Management 9

2.1.19 Repositories 9

2.1.20 Storage 9

2.1.21 Content Integration 9

2.1.22 Migration 9

2.1.23 Backup/Recovery 10

2.1.24 Search/Retrieval 10

2.1.25 Syndication 10

2.1.26 Localization 10

2.1.27 Personalization 10

2.1.28 Publish 10

2.1.29 Paper Electronic 10

2.1.30 Security 10
2.1.31 Collaboration 11
2.1.32 Long-Term Archival 11

3. The "Other" ECM: Engineering Content Management (EngCM) 12

3.1 Differences Between ECM and EngCM 12
3.1.1 Unstructured vs. Structured Information 12
3.1.2 Hierarchical vs. Molecular 12
3.1.3 Non-document Objects 13
3.1.4 Structure Management (WBS, PBS etc.) 13
3.1.5 Configuration Management 13
3.1.6 Termination 14
3.1.7 Concurrent Engineering 14
3.1.8 Compound and Composite Elements 14
3.1.9 Search/Find/Query 15
3.1.10 Document Control 16
3.1.11 Engineering Data Warehousing (integration of granular data) 17
3.1.12 Comparison 17
3.1.13 Engineering Application Integration 18
3.1.14 Design Basis Re-use for Collaborative Engineering 19
3.1.15 Change Impact Analysis (CIA) and Management of Change (MoC) 19
3.1.16 Engineering Application of Documents (such as Mechanical Completions) 19

4. Capability Matrix for ECM and EngCM 21

5. Integration Between ECM and EngCM 23

5.1 Federated Search and Retrieval 23
5.1.1 Query-By-Example – A Common Taxonomy 23
5.1.2 Inter-relationship Navigation 24
5.1.3 Search Proxy – 'Google Everything' 24
5.2 Federated Workflow Execution (including task management and calendaring) 24
5.2.1 Overarching Cross-functional Workflow Processes 24
5.2.2 Shared Folder Routing (You use mine, I'll use yours.) 24
5.3 Collaboration "Team Spaces" 25
5.3.1 Project Dashboards and Rollups 25
5.3.2 Project and Supplier Portals 25
5.4 Common or Shared Services 26
5.5 User Profile Management (including access, security, and IPR) 26
5.6 Presentation: Rendering, Streaming, Viewing, and Markup Services 26

1. Introduction

Enterprise Content Management (ECM) systems and Engineering Content Management (EngCM) systems evolved from the same fundamental requirements – to categorize/classify and index documents on a network for easy retrieval as part of a business workflow process. These documents, or "containers," represent all types of files that need to be managed for access, security, revision, distribution, disposition, and eventual destruction (over and above the capabilities provided by the operating system). Corresponding to businesses drive for productivity, information re-use, data integrity, information provenance, auditability, and regulatory compliance, these systems evolved to manage the "content" (what is ***in*** the files, databases, and models), as well as the containers. However, the requirements for managing Web content, for example, are completely different than the requirements for managing the engineering design basis of a power plant. So, the functional and system infrastructural requirements for these two types of system diverged, fuelling separate software markets for ECM and EngCM. Recent business initiatives (such as Sarbanes-Oxley), environmental and safety controls, and global partnering opportunities have highlighted for many CIOs the business benefits and financial need to integrate these two diverse information systems. This paper summarizes some of the key functional capabilities of both ECM and EngCM systems and illustrates some possible beneficial integration points between them.

1.1 "Making the case for Integration"

The lifecycle of a typical process plant, power generation plant, marine vessel or other 'unique' asset involves many partners;- engineering companies, suppliers, regulators, operators and of course the owner. Masses of data and documents are produced through the life of the asset by these entities which can be broadly split into two categories un-structured and structured, or rather business and engineering. This 'information asset' is constantly manipulated by hundreds of applications throughout the 'value-chain' of the 'physical' asset;- many point solutions, legacy and in-house applications, as well as modern data-centric intelligent applications. In fact many individual applications have been tuned and automated to such an extent that further improvement can only be gained from improving the interface points between them (which includes disciplines, departments, and companies) – i.e. via integration.

Integration of the applications though information management hubs would lead to;- potential reduction in the number of point solutions used, the definition and implementation of exchange and sharing standards for the 'value-chain' to collaborate with rich content reducing the time to market, the lower total cost of ownership by reducing the cost of content that is continuously updated throughout the asset life with the least possible information degradation/reduction.

2. ECM Defined

Abbreviated as ECM, ***E**nterprise **C**ontent **M**anagement* has been adopted as the market definition formerly occupied by vendors of Document Management Systems (DMS). Why the subtle change in name? Initially, DMS vendors concentrated on extensions to basic file management capabilities provided by their operating systems. As such, they were interested in the file (irrespective of its source) and storage/indexing/retrieval mechanisms to allow the user to classify and retrieve documents in an organized, library-like manner. In essence, they were initially concerned only with the file as a container.

As market needs changed, the DMS vendors' focus shifted from "file" management to "content" management. For instance a FDA drug submission is a substantial document that is made up of many sub-documents – the contents – which, in turn, is composed of pages (perhaps formatted XML documents). Content is also manifest in many other application types. A Web site, for instance, is composed of HTML, XML or ASP pages that need to be managed. Electronic mail forms part of the "document record" for a project. So instead of simply managing the containers – the documents – DMS vendors focused their attention to the management of the contents. Hence, we have Content Management.

What is a Document?
Documents are more than just files. They are files that are uniquely characterized and named such that they may be unambiguously identified, referred to, and retrieved. Consider the term "documentary evidence" as a clue to their uniqueness and status.

File systems *are* rudimentary document management systems – providing, that is, if everyone knows the filing structure! File system hierarchies, though, are not always intuitive to each end user, which is why they require additional search and retrieval techniques (one of the initial market drivers for DMS). For example MOSS2007 overlays document management capabilities on top of the basic Microsoft® Windows® operating system.

2.1 ECM Capabilities

ECM already has some pre-defined and recognized capabilities. The organization AIIM[1] is an internationally recognized authority in what constitutes ECM. In fact, the organization is branded as "AIIM - The Enterprise Content Management Association." Therefore, rather than creating new definitions, the following section (in italics) is an extract from the AIIM Web site (http://www.aiim.org/article-aiim.asp?ID=18274), which was written in 2005 by Bryant Duhon, AIIM - The ECM Association, Jeetu Patel, and Rick Tucker (Doculabs). This Web site neatly describes the capabilities of an ECM system.

Enterprise Content Management (ECM) comprises the technologies used to capture, manage, store, preserve, and deliver content, and documents related to organizational processes. ECM tools and strategies enable the management of an organization's unstructured information, wherever that information exists.

---

[1] AIIM was founded in 1943 as the National Microfilm Association and later became the Association for Information and Image Management. AIIM is now known as AIIM - The Enterprise Content Management Association

2.1.1 Content at Work

It's not enough to "manage" content. Of course, the ability to access the correct version of a document or record is important, but companies must go further. Content must be managed so that it is used to achieve business goals. Central to this strategy are the tools and technologies of ECM, which manage the complete lifecycle of content, birth to death.

While there are ECM technologies, more importantly, ECM is an ongoing and evolving strategy for maximizing how your content is to be used. Use the information below as a starting point to review a common content lifecycle. Map a current process to see where you may find overlap and room for improvement for the applications and strategies that your business is developing. The information below only hints at the complexity inherent in any process that deals with managing an organization's content. As always, you must match up the technology tools to address YOUR businesses needs. Technology can enable streamlined management of content, but the underlying strategy must come first.

2.1.2 Compliance

The key to a successful compliance strategy is integrating the idea of compliance success into your business-not viewing compliance as a project that can be completed and then considered "finished." While painful, complying with regulations should be viewed as an opportunity to improve common business processes and not just an ongoing cost to the business. It is no secret that there can be high costs associated with your compliance initiatives for both technology and employees. Only securing compliance for one regulation such as Sarbanes-Oxley or HIPAA will cause your costs to continue to grow as each new regulation is delivered over the years. To help limit the risk and cost, proactive ECM strategies must be developed within key areas, such as records management and business process management. Ensuring that the proper business practices are followed and that content is properly captured, stored, managed, and disposed of at the appropriate and legal time in its lifecycle. Developing a compliance initiative properly will tap many areas of expertise, particularly legal, IT, and records management; all in support of the overall business objectives of the organization. Individuals from each of these areas must contribute their knowledge and perspectives to ensure the benefits of a sound compliance program. While compliance is not always a technology problem, information technology, and the massive growth of unstructured content, contributes to corporate exposure. The tools of ECM, properly used, can help reduce the overall cost of compliance to the business.

2.1.3 Collaboration

Collaboration is the art of working together. The key to strong collaboration is utilizing the set of technologies-instant messaging, whiteboards, online meetings, email, etc.-that allow work to take place wherever and whenever needed. It's good business; groups can accomplish more than individuals. Collaboration allows individuals with complementary, or overlapping, areas of expertise to create better results faster than before. With today's collaborative tools, business units and teams can work together anytime-whether in adjoining offices or a world apart. The technology can now address operational objectives like saving time, streamlining processes, cutting costs, and improving time to market. With the many different types of collaborative tools available, companies must be sure they select the correct tool for their business need. Functionality can be broadly grouped into (1) communication channel facilitation, which enables short-lived interaction such as chat, instant messaging, white boarding, etc.; (2) content lifecycle management,

which manages content objects involved in a business process; and (3) project facilitation, which organizes and simplifies the way that people work toward a common goal. However, there is a catch with collaboration. When using collaborative tools, you must be aware of records management, knowledge capture, and compliance requirements. For some industries, all customer communications must be kept. And, for a collaborative product design process, companies must be sure that the results are kept as business records.

2.1.4 Cost

While ECM can be a costly initiative, what are the costs of not properly managing your content? The cost of not implementing ECM tools is too often left unmeasured until too late. Things like the cost of long legal proceedings, the loss of repeat business through the inability to perform simple customer service interactions, and the cost of typical business process delays are easy to measure after the fact-lawyers' time, the cost to acquire new customers, and FTE salaries. Understanding the cost of these potential losses will allow you to see that ECM investments have valuable benefits that often can be measured, but not always. The key is to set your key metrics for success up front and measure your success based on those expectations. Measuring the revenue based on improved information in the call center can be done as well as measuring the cost benefits of improvements in process speed for a loan application, claim process, or FDA drug approval (to name a few). The improvements will not always show on the final balance sheet but they are out there. While identifying a direct ROI can be difficult, it is not impossible to see the impacts of the improved process efficiency on the business. ECM tools can make your organization more efficient and drive down the cost of doing business. These technologies provide value to your organization by more efficiently organizing information for its subsequent retrieval, use, and, ultimately, disposition. Plus, as these tools are used by more organizations, it becomes part of how you work. What's the ROI on a telephone? Yet, you wouldn't think of doing business without one, would you?

Paperless or Paper-efficient

Many companies are seeking to deploy ECM to support a reduction of paper end-to-end.

In some industries this is expressed as a desire to become ‘paperless’. Clearly in industries where the paper document is the information transport medium or method of interacting with the end customer, e.g. insurance claims processing, then parallel processing of the claim cannot occur while there is still paper in the loop. In this mode the document interface on the front and the back end may be hardcopy, but the intermediate processing may be entirely electronic.

In other industries this is expressed as a desire to become ‘paper-efficient’. Many functional workers still prefer the paper rendition to the electronic rendition. Indeed it is still a challenge in difficult or hazardous areas to take an electronic rendition. Additionally for auditing and insurance purposes many organizations still require the ‘wet’ signature or stamp on the paper record. In this mode it is transported electronically but printed (with the necessary watermarks or banners) at the point-of-use, and later re-captured as a record of the time when it was in hardcopy form.

2.1.5 Continuity

Keeping a business going 24x7 is the task of business continuity planning. While often mentioned with disaster recovery, business continuity planning is the overall strategy for ensuring that operations continue in the event of any disruption-natural or man-made. Disaster recovery is more

narrowly focused on getting an organization's IT infrastructure going again, a subset of business continuity. Because the lifeblood of most businesses today is represented by electronic documents, ECM has a key role to play in continuity. After all, without access to the most vital electronic documents, a business is dead in the water. ECM technologies allow the creation of centralized repositories where all vital corporate information can reside. The method of storage will vary depending on how critical the content is to the company-from off-site back up tapes to redundant, mirrored sites separated by geography and on different power grids. A strong continuity plan will show you that not all content is critical. Companies must prioritize their content to determine how quickly content needs to be back online in the event of a disaster. Business continuity begins with a sound plan and high-level executive support. Next, mission-critical processes and the entities on which they are dependent must be determined, followed by a business impact assessment to determine the impact of a disruption, or losing, those processes. Defining what a business considers a disaster and explaining how key processes will be recovered are the next steps in the plan. A crisis operations center should also be established with procedures for chain of command and other roles. Finally, don't forget to update and test the plan annually or as business needs change. Effectively delivering on a continuity plan will enhance your ability not only to recover during a system failure but will enable you to better define the priority of your business content and improve your overall ECM strategy.

2.1.6 Business Process Management/Workflow

The tools that move content throughout an identified business process, such as claims processing. BPM solutions are frameworks that can be used to develop, deploy, monitor, and optimize multiple types of process automation applications-including processes that involve both systems and people. Consider which processes are candidates for automation, and whether they require some degree of ad hoc processing or manual intervention. Workflow is now commonly associated with the manual processes of managing documents. Workflow handles approvals and prioritizes the order documents are presented. In the case of exceptions, workflow also escalates decisions to the next person in the hierarchy. These decisions are based on pre-defined rules developed by system owners.

2.1.7 Content and Documents

Unstructured content enters an organization's IT infrastructure from a variety of sources. Regardless of how a piece of content enters, it has a lifecycle. Follow a document through its lifecycle as viewed through the use of ECM technology.

1. *Electronic Unstructured Data: email, instant message, text document, spreadsheet, etc.*
2. *Electronic Forms*
3. *Paper Documents/Forms*

2.1.8 Scanning

Paper generally enters the organization through a scanner, or sometimes, a multifunction device. In centralized scan operations, large volumes of paper are put into the system by dedicated workers. In distributed operations, smaller volumes of documents are captured with lower volume scanners or multifunction devices closer to their point of creation.

2.1.9 Document Imaging

Software captures the image of the paper document. Increasingly, electronic document images have the same legal status as a paper document.

2.1.10 Forms Processing

Business forms are ingested into the system. Most forms today are "structured" - the location of the form elements are known. The ability to process unstructured forms, those without a pre-defined form template, is improving

2.1.11 Recognition

Technologies that allow paper information to be translated to electronic data without manual data input. Recognition technologies have progressive capabilities from optical character recognition (OCR) to intelligent character recognitions (ICR) and are important for converting large amounts of forms or unstructured data to usable information in a content management system.

2.1.12 Categorization, Taxonomy

A taxonomy provides a formal structure for information, based on the individual needs of a business. Categorization tools automate the placement of content (document images, email, text documents, i.e., all electronic content) for future retrieval based on the taxonomy. Users can also manually categorize documents. Critical step to ensure that content is properly stored.

2.1.13 Indexing

An essential part of the capture process, creates metadata from scanned documents (customer ID number, for example) so the document can be found. Indexing can be based on keywords or full-text.

2.1.14 Document Management

Document management technology helps organizations better manage the creation, revision, approval, and consumption of electronic documents. It provides key features such as library services, document profiling, searching, check-in, check-out, version control, revision history, and document security.

2.1.15 Records Management

Content of long-term business value are deemed records and managed according to a retention schedule that determines how long a record is kept based on either outside regulations or internal business practices. Any piece of content can be designated a record.

2.1.16 Email Management

As the de facto standard for business communication, removing emails from the server and saving them to a repository isn't enough. Email must be classified, stored, and destroyed consistent with business standards-just as any other document or record.

2.1.17 Web Content Management

Web content management technology addresses the content creation, review, approval, and publishing processes of Web-based content. Key features include creation and authoring tools or integrations, input and presentation template design and management, content re-use management, and dynamic publishing capabilities.

2.1.18 Digital Asset Management

Similar in functionality to document management, DAM is focused on the storage, tracking, and use of rich media documents (video, logos, photographs, etc.). Roots of the technology are in the media and entertainment industry, currently experiencing growth, especially in marketing departments. Digital assets typically have high intellectual property value.

2.1.19 Repositories

Structured and unstructured - the core of many ECM systems. This is where the data resides and where much of a company's investment in ECM resides. A repository can be a sophisticated system that costs hundreds of thousands of dollars, or as simple as a file folder system in a smaller company. The key is to have information that can be found once it is placed in the system.

2.1.20 Storage

Content needs to "live" somewhere. Storage technology (optical disks, magnetic, tape, microfilm, RAID, paper) provide options for storing content online for rapid access or near- or off-line for content that isn't needed often.

2.1.21 Content Integration

Enables disparate content sources to look and act as a single repository.

2.1.22 Migration

As storage media ages, content must be moved to new media for continued accessibility.

2.1.23 Backup/Recovery

Backing up content in various formats and/or locations helps to ensure business viability in the face of a disaster.

2.1.24 Search/Retrieval

One of the greatest benefits of a strong ECM system is the ability to get out what you put in. By having strong indexing, taxonomy, and repository services, locating the information in your system should be a snap.

2.1.25 Syndication

Distribution of content for reuse and integration into other content.

2.1.26 Localization

Recasting content based on the needs and cultural mores of different global markets.

2.1.27 Personalization

Drawing on a taxonomy and based on established user preferences, various types and subjects of content can be delivered via user-defined preferences.

2.1.28 Publish

Content gets where and to whom it needs to go through a number of tools. Content can be delivered via print, email, websites, portals, text messages, RSS feeds.

2.1.29 Paper Electronic

Portal, Intranet, Extranet, Email, Fax

2.1.30 Security

Restricts access to content, both during its creation and management as well as when delivered.

1. *Digital Rights Management - prevents the illegal distribution of rights-managed content by restricting access to content down to the sentence level as well as granting/restricting permissions for forwarding and accessing content.*
2. *Digital Signatures - ensures the identity of a document sender, and the authenticity of the message.*
3. *PKI - uses a public and private key pair held by a trusted third party to transact business over the public Internet.*

2.1.31 Collaboration

Collaboration technologies enable individual users, such as employees or business partners to easily create and maintain project teams, regardless of geographic location. These technologies facilitate collaborative, team-based content creation and decision-making.

2.1.32 Long-Term Archival

Content that must be preserved over decades must be saved to media, such as paper and film-based imaging, with longevity to match.

3. The “Other” ECM: Engineering Content Management (EngCM)

Known as the “other” ECM, Engineering Document Management Systems (EDMS) were primarily conceived to fulfill the same “container management” requirements as the more traditional office document management systems, except that they specialized in the unique requirements of engineering users. In the same way that DMS systems evolved to ECM, EDMS systems have also evolved over time to manage the engineering content, not just the engineering documents. SmartPlant® Foundation is an example of such a system that manages engineering content as well as engineering documents. To differentiate Enterprise Content Management and Engineering Content Management, we will refer to the latter as EngCM.

3.1 Differences Between ECM and EngCM

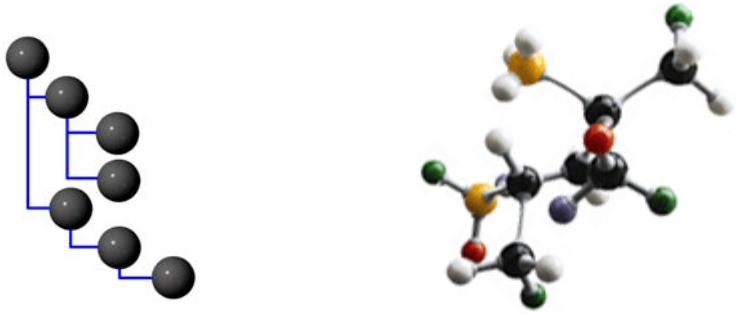
While the primary goal of both ECM and EngCM may be that of managing information, there are business and technical requirements that have resulted in functional differences between the two types of systems.

3.1.1 Unstructured vs. Structured Information

One of the primary and most obvious differences between an ECM and an EngCM is that the ECM is geared towards supporting *unstructured information* and EngCM is geared towards *structured information*. Of course, each of these types of systems may, to a greater or lesser extent, support both unstructured and structured information, but it is their primary focus that determines the features/functions that they support.

3.1.2 Hierarchical vs. Molecular

Most ECM indexing/classification systems focus on the hierarchical model, as represented by the familiar cabinet>folder>file paradigm. Engineering information, however, is massively interrelated and interdependent, which does not lend itself to such neat, flat hierarchies. A drill-down for a plant engineer expecting to find the data for a pump at the bottom of a plant>area>system hierarchy may not be as intuitive for a purchasing engineer who would expect to find the same pump data at the bottom of a supplier>requisition>contract hierarchy (without obviously duplicating the data). Therefore, the engineering indexing/classification system can be more likened to that of a molecular model, albeit ultimately represented as flattened hierarchies for ease of navigation. Figure 1 illustrates hierarchical vs. molecular structure.



Hierarchical Molecular

Figure 1: *Most ECM indexing/classification structures are hierarchical vs. molecular.*

3.1.3 Non-document Objects

In EngCM systems, there are many objects required to characterize the configuration, structure, topology/connectivity, and process conditions of the plant, as well as to provide an aid to finding documents. These objects typically represent the logical and physical aspects of the plant. They include, but are not limited to, plant, system, tag, and asset. These non-document objects not only carry meta-data for identification, but also meta-data that characterizes the "engineering design basis" or "engineering record." These non-document objects form the navigational nodes in the molecular classification system indicated above. Documents (including drawings, schematics, 3D models, and others.) may hang off relationships to these non-document objects. Documents may also be the presentation mechanism for metadata on these non-document objects. For example, a mechanical equipment datasheet is a presentation of the metadata on the equipment object generated on-the-fly from the database, which may or may not be stored as a traditional document (e.g. a spreadsheet) as well.

3.1.4 Structure Management (WBS, PBS etc.)

By using the combination of the molecular model and the non-document objects indicated above, an EngCM can simultaneously represent many different structures or views that depict the plant, a work breakdown structure, a plant breakdown structure, a commissioning structure, and so on.

3.1.5 Configuration Management

As the plant evolves, so do the structures that represent the plant. To support regulatory, safety, and design alternative analysis, there must be the option to turn back the clock and examine the state of the data – the structures and content – that were in place at that point in time. This involves effectivity setting – most notably date-level effectivity. These date settings are carried on every object in the EngCM to determine when they came into effect, and when they ceased to be effective.

3.1.6 Termination

To support configuration management, objects (document or otherwise, and the relationships between them) are not deleted, but terminated – setting the date when they cease to be in effect. This is a subtly different concept from the records management and migration concept of ECM systems.

3.1.7 Concurrent Engineering

Plants are in a constant state of change, with many projects being executed to maintain a safe and highly productive operational envelope. But plants also need to have a high operational uptime, and therefore cannot be taken offline for each and every change project. Thus, these projects need to be "stacked up" for the next window of opportunity when the plant can be shut down for these changes. This means that the engineering projects must run concurrently – involving the same objects existing in multiple, differing states (for each project) simultaneously (concurrently).

3.1.8 Compound and Composite Elements

Many engineering documents are composed as either compound or composite documents. Compound documents are made up of many pages from multiple sources, but are still referred to as a single document, as in the example on the left in Figure 2. In this example, the relationships managed by the EngCM are external between the files. On the other hand, in the example on the right below, a single document may be composed of many other references. For example, a drawing may have the drawing border as a reference, or other symbology from graphics libraries, or other embedded objects (e.g. Microsoft OLE). This is referred to as a composite document, and the internally related/reference objects are managed by the EngCM.

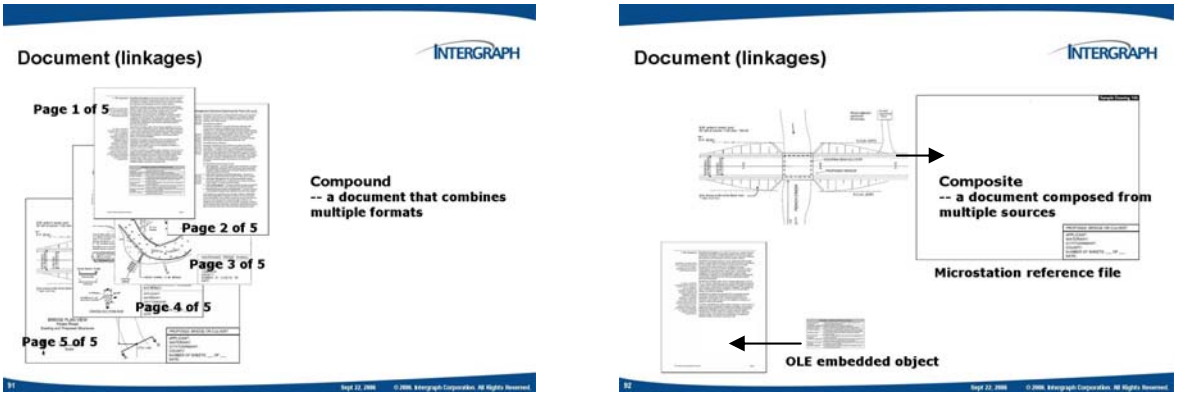


Figure 2: *Many engineering documents are either compound or composite.*

3.1.9 Search/Find/Query

If one thing is for certain, users performing a wide variety of roles during the life of a facility will require a wide spectrum of methods for organizing, classifying, and retrieving data. It must be recognized that different classes of users will have different skill sets and experience when it comes to interacting with an EngCM. For example, information/knowledge workers, such as design engineers, have a primary role of creating, modifying, and approving information. Typically these "power users" demand creative, flexible, and configurable styles of interface. However, maintenance engineers, who use the information system only as a tool to fulfill their primary role, demand a simple, intuitive, and discoverable style of interface.

Therefore, an EngCM needs to support a wide range of search and retrieval paradigms. Traditional Query-By-Example (QBE) dialogs include fields that may be filled with as much text known about an attribute (meta-data) associated with the requested object, using familiar plant terminology such as Classification (pump, centrifugal pump, etc.), Discipline (mechanical, electrical, etc.), System (firewater, process line, etc.), Geographical Location (sector, area, floor, etc.), Specific Tag Number (pump P-101), Drawing Type (P&ID, plot plan, installation manual, etc.), Work Packages (procurement, construction, testing, etc.) and including wild cards (*).

If the plant object being requested is a document or drawing (containing text) and none of the attributes (meta-data) are known, a Query-By-Content (QBC) or Full-Text-Retrieval (FTR) technique can be used. For example, find all the objects/documents containing the phrase "pump P101" within the contents.

Another technique for plant object location is by exploiting the relationships inherent in the plant model as indicated above, or Query-By-Relation (QBR). This explorer-style navigation and drill-down through data (e.g., plant breakdown structure) is a dramatic aid to decision support, cause-and-effect analysis, where-used and used-on searching. For example, if a certain type of seal is rejected during an inspection, simple queries can identify all other plant locations for similar seals, enabling a quicker fix and reduced production shutdown time. Or, a user searching for a P&ID may not know the drawing number, but they probably they know the name of one of the tags represented on the drawing. In this case, the user can find the tag of interest and then follow the relationship directly to the P&ID.

An extension of the QBR methodology is that of exploiting the plant data model relationships graphically using links such as web URLs and graphical hotspots within the graphical file or model. For example, select a graphical object representing a pump on a P&ID, navigate to the same plant object within a 3D model, and then navigate to the same objects data sheet, and so on.

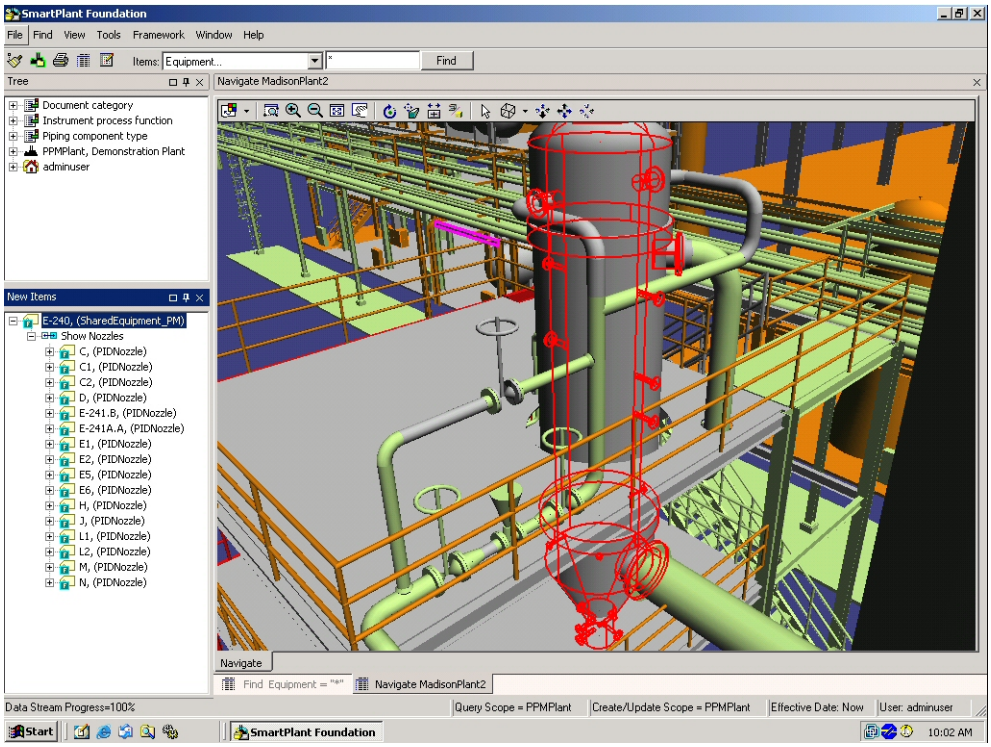


Figure 3: *A 3D model can depict the results of a complex query.*

The 2D and 3D files and models can also be used to represent the results of complex queries. (See Figure 3.) For example, highlight plant objects in the 3D model all items that have been scheduled within the project planning system, and that have not yet been ordered within the purchasing system. As such, 2D and 3D files/models provide excellent visualization vehicles for evaluating the status and reporting of the plant as it develops.

This “find what you want by what you know” philosophy allows users to spend more time performing engineering tasks and/or operating the plant and less time looking for information, with the constant assurance that the information is accurate and approved for use.

3.1.10 Document Control

EngCM systems that communicate documents between project partners typically implement a document control function beyond basic document management capabilities for the purpose of handling transmittals. A transmittal is a controlled package of documents with an accompanying report used to perform, record, and track the distribution of project documentation between the different design teams, subcontractors, vendors, and construction sites involved in a project. The transmittal is a critical record that is often referred to for status of work completion and contractual satisfaction.

The term “transmittal” often refers to the top sheet or report that lists how many copies of each document are sent to whom, and for what reason. In a paper-based system, the transmittal is used by the print room to prepare physical prints. In an EngCM, the transmittal becomes a collection object that links a set of documents to a set of recipients with reasons for receipt. The transmittal report is attached to the collection object and is emailed to the recipients upon issue of the

transmittal. In many cases, transmittals are used to distribute documents between different companies, which would also need to be recorded on the transmittal.

Transmittals play an important part in tracking the progress of a project. Milestone payments from the client are often linked to the distribution of particular documentation on transmittals, the most notable of which is the site construction transmittal when documents are released to the site for construction. Transmittals are also useful for internal recipients, such as for the distribution of control documentation, safety procedures, and other documents between departments/disciplines.

3.1.11 Engineering Data Warehousing (integration of granular data)

During a project, engineering data will come from many sources, applications and partners. This data forms a critical part of the handover record from an engineering company to the plant operator, and is typically used during construction, commissioning, and loading operational systems such as maintenance management. Therefore, the data needs to be valid, correct, and of high integrity; however, since it comes from multiple sources, it could conceivably contain duplicate or erroneous records. As part of the capture process it should be subject to a validity, consistency and integrity check – and quarantined if not clean!

Since the data will come from many different data models, have different metadata, and even different numbering systems, the requirement from an EngCM is to map this data from the source systems to a common data model, aggregate it, and consolidate it. This process includes not just the documents, but also drawings, models, databases, and records from many (often 400+) applications used during the engineering and construction phases of a project.

3.1.12 Comparison

Since the EngCM contains all types of engineering content – such as granular data and document records sourced from many applications over the life of the asset it – is a natural step to provide document and data comparison capabilities. This is important, as indicated in the concurrent engineering requirement as discussed above, and it is also important from a change management perspective. Figure 4 compares two versions of a P&ID, and the result indicates on the righthand table that the pressures and temperatures have been changed for the highlighted pipeline.

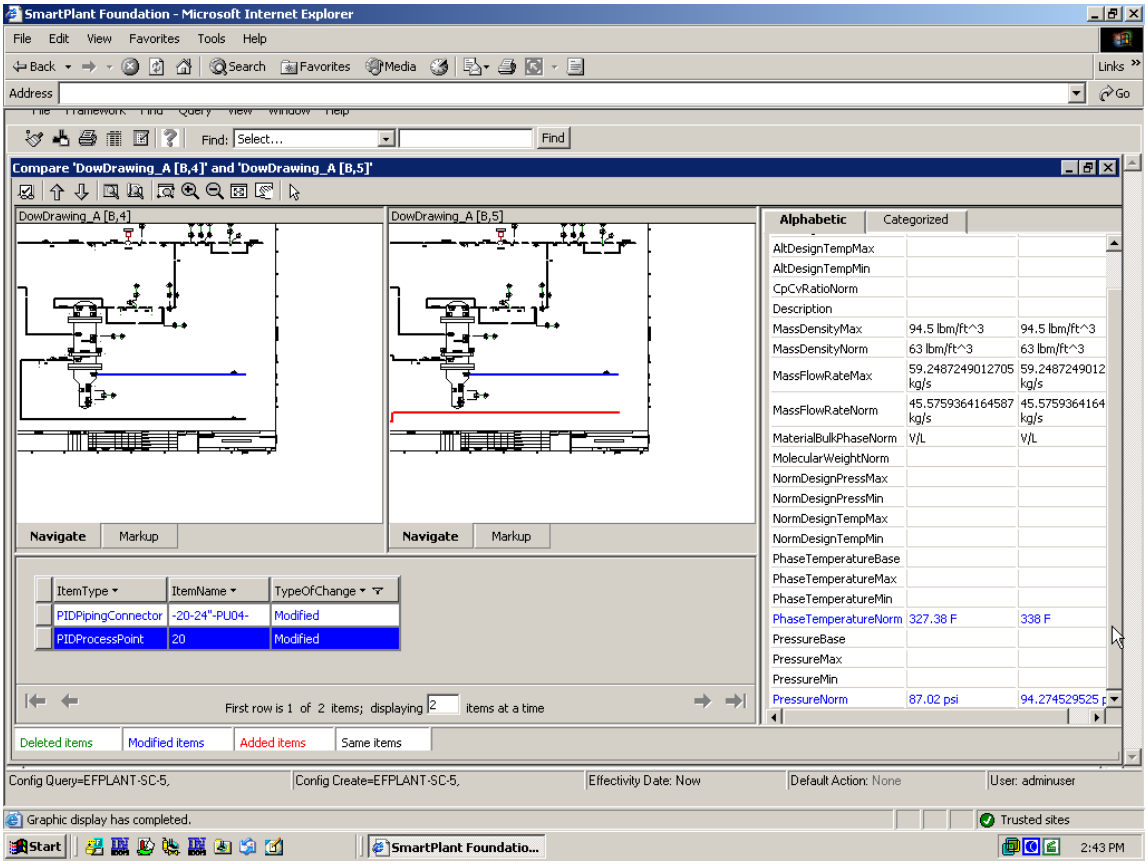


Figure 4: *SmartPlant Foundation compares two versions of a P&ID.*

3.1.13 Engineering Application Integration

Many organizations today are looking to integration to dramatically reduce execution time and improve productivity. They have pushed the envelope of task automation and see “write once read many” as the next logical improvement. Integration and interoperability of data between heterogeneous applications is the goal, since no one can deliver the one-size-fits all application. However, while technologies exist for point-to-point integration, heavy reliance on it has two primary effects: a degree of fragility in the interchanges (if one breaks they all potentially break), and no overarching control or visibility of data changes (the integrity, or lack thereof, may not be spotted until too late). As shown in Figure 5, an EngCM provides a common hub for the interchanges and transactions (which have a significant temporal misalignment in engineering disciplines) and manages the distribution of data as necessary. (See also concurrent engineering.)

Figure 5: *An EngCM provides a hub for interchanges and transactions without the fragility of point-2-point integrations.*

3.1.14 Design Basis Re-use for Collaborative Engineering

With Engineering Application Integration comes the ability to share data between tools whose application data models are dissimilar (e.g P&ID, Instrumentation, Electrical, Structural, Process etc.). This provides a platform for design basis re-use where the data within the EngCM is available to be re-used between many different applications (write once read many). This is completely different from the traditional Check-In/Out provided by an ECM where the same tool is used for changes to the same document. Data sharing by integration has been shown to dramatically improve the integrity of the engineered asset, reduce the potential for error and improve overall productivity and time to market.

3.1.15 Change Impact Analysis (CIA) and Management of Change (MoC)

The combination of all the EngCM functionality indicated above provides a framework for analyzing the impact of a change. With engineering data being so highly inter-related and inter-dependant a seemingly minor change to a value on a datasheet can have a ripple effect throughout the entire design basis – changing data and documents in many applications. Once the impact of this change has been assessed within the integrated data of the EngCM a structured workflow process that overarches the individual tools can be orchestrated to ensure that the change is executed and approved throughout the design basis.

3.1.16 Engineering Application of Documents (such as Mechanical Completions)

Most folks don't go to work to manage documents; it just so happens that the job invariably involves documentation. This may be in the form of instructions, documents to read, documents to prepare, or records/forms of the process collected. Invariably, though, the document is the universal carrier of the data. Leading EngCM systems therefore embody the document in the work process;

use it to prepare the work, use it as part of the work process, capture the result of the process, and pass it on to the next process. An example of this in SmartPlant Foundation is the Mechanical Completions process. Here the engineering record is used as part of the pre-commissioning of a plant. Completeness is assessed and certified, and handed over to the operations commissioning team for startup preparation.

4. Capability Matrix for ECM and EngCM

The table below illustrates the matrix of different capabilities provided by both ECM and EngCM systems. There are some capabilities that are provided by both types of systems to a greater or lesser extent. No one system on the market today offers all the requirements for Enterprise ***and*** Engineering content management. Therefore, it would seem that integrating those different but aligned technologies would bring significant benefit to an organization embarking on a broad-based information management strategy. But what would those integrations look like? What would some of the key integrated functions achieve? How would some of the organization's work processes evolve as they take advantage of such a strategy?

Capability	**Provided by ECM**	**Provided by EngCM**
Business Process Management / Workflow	Y	Y
Content And Documents	Y	Y
Scanning	Y	
Document Imaging	Y	
Forms Processing	Y	
Recognition	Y	
Categorization, Taxonomy	Y	Y
Indexing		
Document Management	Y	Y
Records Management	Y	Y
Email Management	Y	
Web Content Management	Y	
Digital Asset Management	Y	
Repositories	Y	Y
Storage	Y	Y
Content Integration	Y	Y
Migration	Y	Y
Backup/Recovery	Y	
Search/Retrieval	Y	Y
Syndication	Y	Y
Localization	Y	
Personalization	Y	Y
Publish	Y	
Paper Electronic	Y	Y
Security	Y	
Collaboration	Y	Y
Long-Term Archival	Y	Y
Unstructured vs Structured Information	Y	Y
Hierarchical vs Molecular		Y
Non-document objects		Y
Structure Management (WBS, PBS etc.)		Y
Configuration Management		Y
Termination		Y
Concurrent Engineering		Y
Compound and Composite Documents		Y
Search / Find / Query – find what you need by what you know		Y
Document Control		Y

Capability	Provided by ECM	Provided by EngCM
Comparison		Y
Engineering Application Integration		Y
Design Basis Re-use for Collaborative Engineering		Y
Change Impact Analysis and Management of Change		Y
Engineering application of documents – e.g. mechnical completions		Y

5. Integration Between ECM and EngCM

The sum of the parts of the integration between two such systems should provide a level of capabilities that is greater than the whole. It should ***not*** become an exercise to rationalize and use a function from one or the other, when clearly there is benefit to both. In fact, the provision of one service or another should be completely seamless to the end user; they should not have to know which system the data is in, how to find it, how to access it, how to use it within their business process, how to move it around the organization, and so on. The following provides some examples of potential valuable integration points. This is by no means an exhaustive list and is presented here as an illustration of the potential.

5.1 Federated Search and Retrieval

One of the most basic and fundamental requirements of any such system is to easily and intuitively find the information in the organization, no matter where it resides and what form it is in. Providing an integration between ECM and EngCM systems that simply presents the user with two or more different search screens (for each system) – on the assumption that the user knows in which system the data resides ***and*** what it was called in each – is not a federated search capability. A federated search is provided by a common or uniform search paradigm that is in turn executed against the one or more repositories supplying the user. Now clearly it is not as simple as providing a single search screen; if only that were the case! But there are a number of different approaches.

5.1.1 Query-By-Example – A Common Taxonomy

Providing a common "query-by-example" search tool is a challenge where naming conventions and terminology are not common between systems. This is compounded by differences between local/site terminology, company nomenclature, and national language; even acquisitions or partnering adds new language variations. It is not uncommon for many companies to employ content management experts whose role it is to provide a common taxonomy[2] and ontology[3] for their organizations. (Convincing other collaborating companies in their entire supply/value chain to adopt that same vocabulary is another matter). Therefore the query-by-example (and of course the corresponding data loading mechanisms) may need to support a meta-data mapping[4] capability able to translate from a common definition to multiple definitions[5] required for each search engine.

---

[2] **Taxonomy** is the practice and science of classification - http://en.wikipedia.org/wiki/Taxonomy

[3] **Ontology** is a study of conceptions of reality and the nature of being. It seeks to describe or posit the basic categories and relationships of being or existence to define entities and types of entities within its framework. - http://en.wikipedia.org/wiki/Ontology

[4] **SmartPlant Foundation** provides such a meta-data mapping and meta-data adaptor for integrating with systems whose inherent data models are different

[5] Sharing common tools such as popups, pulldowns and other menu'ing concepts would require careful creation and consolidation

5.1.2 Inter-relationship Navigation

Most modern ECM and EngCM systems exploit the inherent interrelationships between object instances, whether it be the ECM cabinet>folder>file paradigm or the richer plant structures of an EngCM such as plant>area>unit>system>tag. Also, the majority of modern ECM/EngCM systems are Web-centric, or provide an interface through a common portal (e.g. Microsoft SharePoint). An advantageous integration point is to provide a common drill-down or tree structure whose node/branches are links (URLs), or queries formulated for their respective systems. User perceives one common integrated tree structure, but the data is sourced form multiple locations.

5.1.3 Search Proxy – 'Google Everything'

There are many cross platform engines to index and categorize text and provide linkages back to the source, such as Search engines, Web crawlers, and full text retrieval. We have become "Googlized," as the term has become an everyday verb for us. "Did you Google it?" now means to search using some free-format, unstructured query. This offers another potential integration point for our federated search. However, we must take care in the construction of such a system marrying up the unstructured ECM and the structured EngCM. For instance, how is relevance ranking determined? How is access and security honored?

5.2 Federated Workflow Execution (including task management and calendaring)

Workflow – the automated and electronic routing of tasks, distributing information to those that need it, when they need it, and the capturing of the results for subsequent auditing – is a technology that has become embodied in many ECM and EngCM systems today. No department or discipline should be an island! So, federated workflow execution may support one of a number of integration points.

5.2.1 Overarching Cross-functional Workflow Processes

The ECM and EngCM sub-systems may provide different workflow technology[6] and processing engines, each supporting their own sub-process execution. An overarching workflow process execution may be able to provide a task to a sub-process. The sub-process may able to receive the task, execute another workflow process, and provide a result back to the overarching workflow for continued execution.

5.2.2 Shared Folder Routing (You use mine, I'll use yours.)

Another mechanism to support federated workflow is to concentrate on the routing of the content, e.g. put an ECM object through a EngCM workflow and vice versa. This is useful for execution

[6] **Workflow Management Coalition** (WfMC) http://www.wfmc.org/standards/framework.htm and the **Business Process Management Institute** (BPMI) http://www.bpminstitute.org/index.php working towards standards for interoperability of workflow systems

processes where the content is shared, but the process is not. For example, consider an engineering change workflow executing in the EngCM. An EngCM folder may contain not only engineering objects from the EngCM but also documents from the ECM – e.g. a folder containing engineering drawings, standards, specifications, purchase requirements, and authorization procedures. A pre-requisite for such may be the federated search indicated above.

5.3 Collaboration "Team Spaces"

Collaboration has come to mean different parties working together on common tasks and projects, and sharing of files with a common technology that presents/binds it all together – without the need for each party to understand the originating core applications that created the source content. The Microsoft SharePoint family of products, most notably Microsoft Office SharePoint Server (MOSS), is one of a number of technologies that embody this working paradigm – served up via the Web, through a browser, simple point-and-click intuitive interfaces, and content delivered through user-role or application-specific windows[7]. Some of these systems are hosted on the Internet as paid services. (Some within the company intranet are secure portals for partners only.) The real power in these teamspaces comes when they are integrated as role-specific applications, such as project review, where the Web parts from the different systems communicate with each other over secure connections.

5.3.1 Project Dashboards and Rollups

Dashboards and rollups provide excellent mechanisms for management or project staff that need quick access to and rapid visualization of status/rate-of-change of projects. Providing integrated reporting capabilities (a report that can execute on either repository and return the result), progress rollups, status change notifications, completion status (traffic lights) – all these key indicators are delivered using reporting built on federated search.

5.3.2 Project and Supplier Portals

In any engineering project, huge volumes of data are exchanged within the value-chain. This content may be in the form of formal multi-page documents, drawings and models, email messages, databases, and multimedia – all in a variety of file/media types. Much of this content traffic goes through the traditional document control or records management function of an organization, to capture, classify, and record the existence of the exchange and transmission. however, much of the electronic media bypasses the recognized channels. For an owner operator, much of the discussion is not adequately captured at all. A project and supplier portal, backended with ECM and EngCM systems to formally capture all of the communication (formal exchanges, review sessions, decision support, threaded discussions, and contractual commitments) would channel the communication to ensure that there was a full and proper record of all transactions[8]. Figure 6 illustrates the project and supplier portal structures.

---

[7] SharePoint refers to these as 'web parts', Netweaver as 'iViews', 'portlets' is another common term to describe the window on a web page that points back to an application

[8] Note this portal may be created and managed by the Plant Owner, the Operator, the EPC(s) or indeed may be an independently provided service, the key is all content traffic is directed through it

Figure 6: *Project and supplier portals enable better communications.*

5.4 Common or Shared Services

There are other integration points between an ECM and EngCM that are worth indicating, and would provide back-end services and not integration that the user would interact with directly. These could be considered as integrations that use APIs or even Web services (within a service-oriented architecture).

5.5 User Profile Management (including access, security, and IPR)

Of course, if the user is to access content from multiple systems, these systems need to honor each others' mechanisms for accessing and using the content stored within. In some cases, this content – particularly in the case of an EngCM system, will be stored in a very granular data form. Clearly, document-based systems are not designed to manage such a level of security requirements. Therefore, the systems should provide positive deniability in the case of doubt. In any case, the content should not be allowed to flow beyond the domain of the user, group, discipline, department, or company without the necessary controls in place.

5.6 Presentation: Rendering, Streaming, Viewing, and Markup Services

As indicated earlier, content comes in many forms, from many sources. It is not appropriate, cost-effective, nor feasible for a single user to understand and learn every application to support every format that will ever be provided to them during the course of a project. Even if the user was skilled in the tool and had access to it at the desktop, it still may not be appropriate to deliver the content to that user in that form. For instance, if a document that has been approved and issued, it may be more appropriate to deliver it to the user in a non-editable form, such as PDF, streaming 3D model, or other similarly rendered forms, and within a simpler viewing tool. This concept is not new to either ECM or EngCM systems, but what is new is the concept of one of these systems providing rendering, streaming, viewing, and markup services to the other.

About Us

About Intergraph

Intergraph Corporation is the leading global provider of spatial information management (SIM) software. Security organizations, businesses and governments in more than 60 countries rely on the company's spatial technology and services to make better and faster operational decisions. Intergraph's customers organize vast amounts of complex data into understandable visual representations, creating intelligent maps, managing assets, building and operating better plants and ships, and protecting critical infrastructure and millions of people around the world.

Headquarters
Intergraph Corporation
170 Graphics Drive
Madison, AL 35758

For more information about Intergraph, visit our Web site at www.intergraph.com.

PPM104A0 – revision 1