INTELLIGENT BUSINESS STRATEGIES

WHITE PAPER

Maximising Business Value from Content Intelligence and AI

By Mike Ferguson
Intelligent Business Strategies
June 2025

Research sponsored by:
HYLAND™

Table of Contents

INTRODUCTION

In today's digital economy companies view data, intelligence and AI as strategic

In today's digital economy data, analytics and AI are now seen as strategic in almost every company. Every business function wants this to provide the insights, and predictions needed to make more timely and effective decisions as well as to help them automate tasks, reduce costs and improve productivity. According to one global CEO survey[1], CEOs see tangible impacts of Generative AI with 56% reporting efficiency gains, 34% reporting profitability increases and 32% reporting revenue increases. Also, about half of CEOs (49%) expect Generative AI to increase the profitability of their company over the next 12 months.

Many organisations are seeing measurable business benefits from AI

Most data and AI initiatives are focussed on structured data

To date however, almost all of the data, analytics and AI initiatives in most enterprises have been focused on the use of historical and near real-time structured and semi-structured data when producing insights. Yet, as much as 80% of the data in most enterprises is unstructured data housed in multiple content stores both on-premises and in multiple clouds. This includes unstructured data in content management systems, file systems containing office documents, email systems, data in collaboration tools, social networks and much more.

As much as 80% of enterprise data is unstructured and resides in multiple systems

Analysis of unstructured content to produce insights has not been done in many enterprises

Many companies have not yet looked at unstructured content in their enterprise at all when it comes to analysing data to produce insights which begs the question:

> "How many new insights are hidden in all this untapped unstructured content that could have a major impact on driving new business value if added to what companies are ready know?"

Many high-value insights remain hidden in unstructured content

The opportunity is, without doubt, potentially very significant. Although analysing unstructured data is not new, it has historically relied on highly skilled data scientists programming in Python and using natural language processing (NLP) and deep learning to process and analyse text, images, audio and video. However, the impact of Generative AI on unstructured data analytics is huge. It can be used to automate the processing of unstructured content, democratise query and analysis of content to produce content intelligence and the use of AI to generate explanations of that intelligence to understand the impact on your business.

Generative AI has enabled the query and analysis of unstructured data to produce content intelligence to become democratised

Now is the time to tap into a rich source of new insights to maximise business value

Now is the time to tap into a rich source of new insights to maximise business value. This paper looks at the current challenges companies face with unstructured content, what is needed to successfully implement unstructured content intelligence and AI to get maximise value from it and how one company, Hyland is stepping up to help make this possible.

---

[1] PWC 28th Annual Global CEO Survey 2025

CHALLENGES HINDERING THE USE OF UNSTRUCTURED ANALYTICS AND AI IN THE ENTERPRISE

Despite rising business demand to analyse unstructured data to produce new insights, there are a number of challenges that are hindering organisations in achieving this. These are as follows.

THE INCREASINGLY DISTRIBUTED UNSTRUCTURED DATA LANDSCAPE

Unstructured content is fractured across multiple data stores

The first of these challenges is that unstructured content is scattered across multiple content stores both on premises and in multiple clouds as shown in Figure 1.

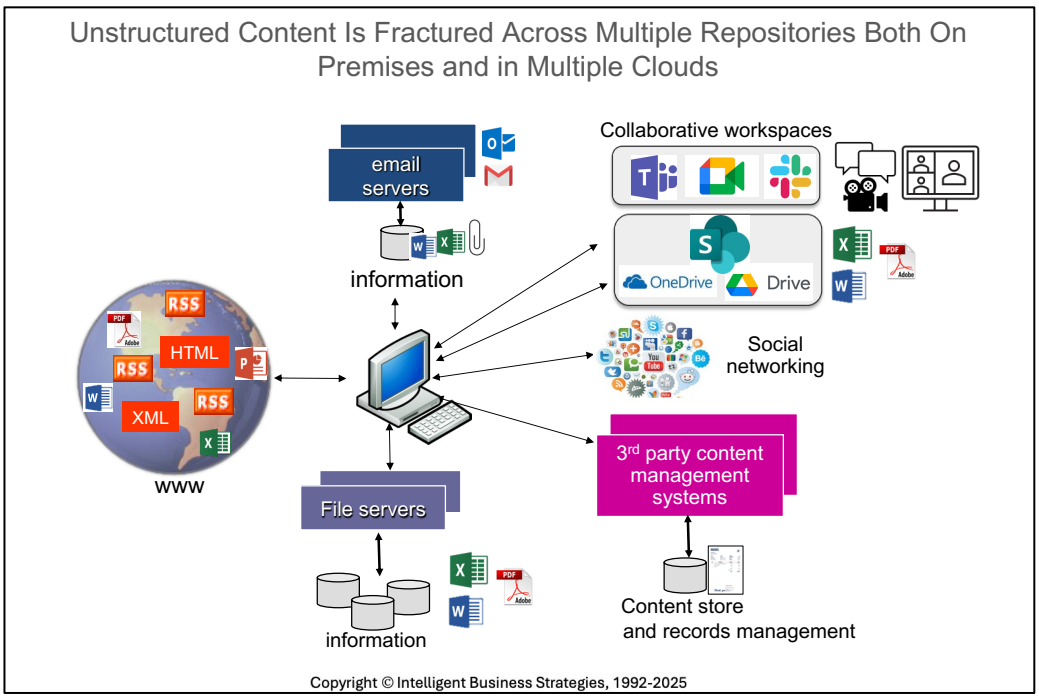


Figure 1

This makes it difficult to identify the relevant content needed for analysis

In addition, there is relevant external content scattered across the public internet that could also be valuable to specific analyses.

All of this makes it very difficult to identify the relevant content needed for specific analyses as well as making it difficult to integrate and govern.

LIMITED CLASSIFICATION OF CONTENT

Unclassified content makes it harder to find related content

In addition, many companies have done little to classify this content or create taxonomies to help organise it. This makes it difficult to find relevant, related content using mechanisms like enterprise and faceted search within the enterprise.

Furthermore, content curation and preparation for use in analytics is often very limited which means that metadata associated with content is poor.

Content curation is also limited making search far less effective than it should be

The result is that enterprise search is nowhere near as effective as it could be in terms of surfacing relevant content to meet users' needs compared to what would be the case if it could harvest rich metadata.

UNSTRUCTURED DATA IS NOT 'AI-READY'

Content is also not AI-ready so it can't be exploited by natural language queries to produce new insights

In addition, now that we are in the era of generative AI and AI-Agents, the ability to analyse unstructured content to produce content intelligence has become a lot easier but only if your unstructured content is AI-ready. However, in most enterprises, the vast majority of unstructured content is not inherently usable by AI. It lacks the structure, context, and accessibility needed to power natural language queries, drive automation, or generate reliable insights. Even with the rise of generative AI and AI-Agents, simply having unstructured content is not enough. For AI to interpret, reason, or respond with relevance, that content must first be transformed. It needs to be cleaned, parsed, and enriched. Without this foundational step, unstructured data remains disconnected from AI-driven business value.

DIFFICULT TO AUTOMATE PROCESSES RELIANT ON UNSTRUCTURED DATA

AI can facilitate automation of process tasks that are heavily reliant on unstructured data

Finally, the automation of business processes has traditionally been done with data from structured data systems where the manipulation of this data has been straightforward. However, if business process automation is heavily reliant on context within unstructured data, it has often proved much more difficult. This is where the use of AI holds a lot of promise

Given these challenges, it is clear that there are a number of requirements that need to be met if companies are going to be able to analyse and act on insights produced from content intelligence and AI. Let's take a look at what those requirements are.

REQUIREMENTS FOR SUCCESSFUL IMPLEMENTATION OF UNSTRUCTURED CONTENT INTELLIGENCE AND AI

Companies need unified unstructured content intelligence and AI capability within the enterprise to analyse content across multiple data stores

Successful implementation of unstructured content intelligence and AI is likely to follow a staged approach. The first stage is the need to build a unified unstructured content intelligence and AI capability within the enterprise to analyse content and produce content intelligence, alerts, predictions and recommendations (collectively known as insights) from the underlaying relevant content scattered across multiple underlying content stores both inside and outside the enterprise. These insights then need to be integrated into business processes and applications to support decision making in the context of tasks being performed.

Insights produced from the analysis of unstructured content need to be integrated with those from analysis of structured data

The second stage is to integrate the unified unstructured content intelligence and AI capability with existing analytical capabilities associated with analysing, alerting, recommending and acting on structured data. This ensures that decision makers have access to a complete set of insights produced from analysing all relevant structured, semi-structured and unstructured data needed to support decisions.

In order to make this possible a reference architecture is needed.

REFERENCE ARCHITECTURE

There is a need for a reference architecture for unified content intelligence and AI. An example reference architecture is shown in Figure 2

Companies need a reference architecture to help them build a unified unstructured content intelligence and AI capability that supports multiple workloads analysing content across multiple data stores

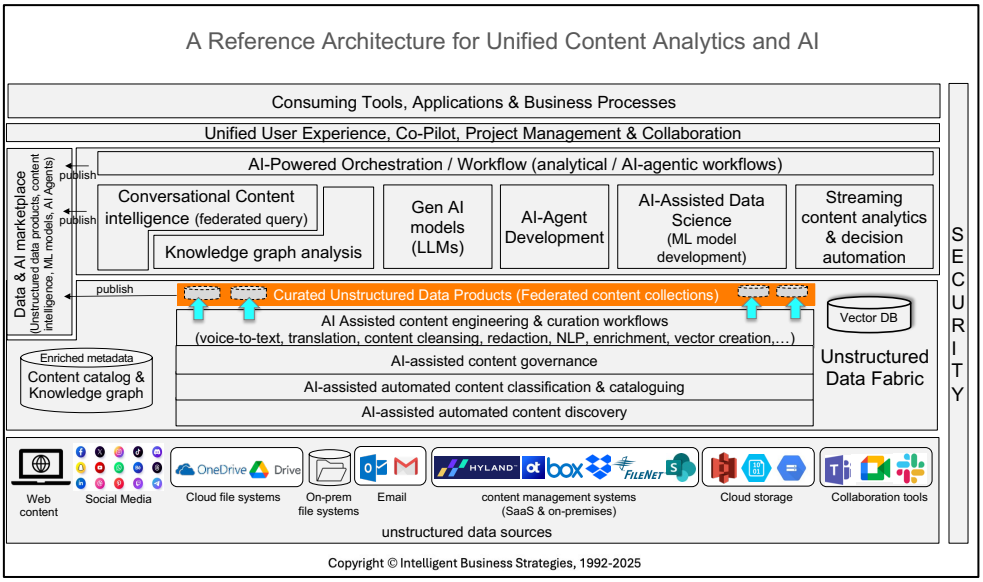


Figure 2

REQUIREMENTS SUMMARY

A summary of requirements associated with this are covered below. A more detailed set of requirements is shown in the appendices. Looking at the summary, it should be possible to support:

Connectivity across a broad range of content stores on-premises and on multiple clouds	• Connectivity and simplified access to a broad range of content sources scattered across multiple clouds and on-premises including content management systems, files shares, cloud storage, email, collaborative tools and social networks
Automatically scan, catalog and classify content	• Automated discovery and cataloguing of content to create a knowledge graph of all the unstructured data assets and relationships • Automated classification of content containing sensitive data so it can be protected before making it available for analysis and to make it easier to find and govern
Curate, clean, enrich, and transform unstructured content into AI-ready formats	• Use workflow to create pipelines that curate, clean, enrich, and transform unstructured content into AI-ready formats. This includes chunking documents, extracting metadata, generating vector embeddings, identifying entities and relationships, and organising content into structured, reusable collections.
Publish curated content that is ready for analysis	• The creation and publishing of curated, compliant, reusable, unstructured data products in a data marketplace (e.g., customer interactions - emails, chat, voice, social)
Govern content across multiple content stores	• Enable users to model their own ontology in a knowledge graph • Definition and enforcement of policies across a distributed data estate to govern content
Support conversational query, analysis, visualisation and explanation	• Conversational query and analysis of content across multiple content stores via a conversational prompt-based user interface • The creation of content intelligence including visualisations and business impact explanations on multiple types of content
Pre-built AI-Agents	• Out-of-the-box AI-Agents that recommend and take actions using intelligence from content from within interactive conversations
Create your own AI-Agents	• The creation of domain specific, user-defined analytical AI-Agents to analyse and act on content intelligence for a specific business purpose
Create powerful agentic workflows automate tasks and decisions	• Orchestration of agentic workflows to produce new insights in batch, on-demand and on an event-driven basis. This can help to automate tasks, decisions and actions
Integrate structured and unstructured analytical environments to create a holistic set of insights	• Integrate content intelligence and AI-Agents with traditional analytical environments to engineer data from structured and unstructured sources and to provide a holistic set of insights for more effective decision making
Integrate content intelligence and AI into applications and processes	• Integrate unstructured content intelligence and AI-Agents into operational business processes to leverage content intelligence on demand and to automate tasks • Integrate unstructured content intelligence and AI-Agents into business applications to provide contextual, intelligence, predictions, recommendations and actions and to automate tasks and decisions

WHAT BENEFITS CAN UNSTRUCTURED CONTENT INTELLIGENCE AND AI OFFER?

Having discussed the requirements for successful implementation of unstructured content intelligence and AI, what benefits can this offer?

Unstructured data management enables content curation and governance which in turn creates trust

The first benefit is the introduction of unstructured data management. This enables content curation, processing and governance and introduces a level of trust around the use of unstructured content. This ensures that business value from content intelligence and AI is maximised.

Then there is the value from AI-driven analytical pipelines that can invoke AI-Agents and classic ML models to convert voice to text, translate languages, classify content, extract entities, summarise text, analyse sentiment, predict taxonomies and ontologies and generate content. For example, if you combine this with topic and sentiment classification, then you can combine content from multiple sources to quickly determine customer opinions about your products and brand (See figure 3).

Powerful analytics workflows can be created to quickly process and analyse complex unstructured data to produce new high-value insights

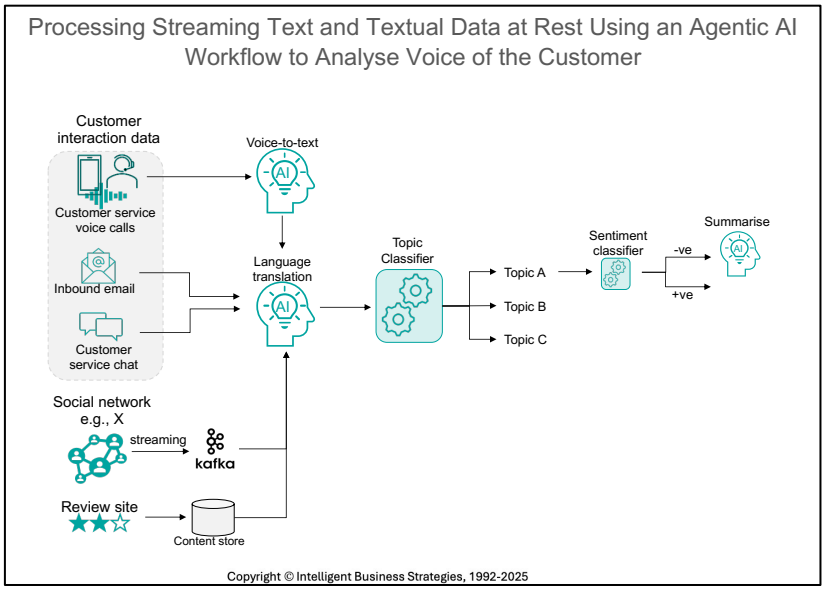


Figure 3

Insights produced from analysing content can be integrated into applications and processes so decisions can be taken at the point of need

This example shows that analysing unstructured data provides totally new high value insights to add to what is already known about your customers from analysing their transaction activity. The combination of structured and unstructured analytics reveals so much more including likes, dislikes, sentiment, and intent. Figure 3 also shows that you can combine and analyse content from multiple content stores to create high value content collections. Content does not have to be consolidated first. You can also use generative AI to query high value content, produce intelligence and act on it. This intelligence could also be embedded in applications to deliver it in context. AI can also be used to automate decisions and tasks.

More vertical industry examples are provided later in the paper.

How is Hyland Stepping Up to Meet Enterprise Content Intelligence and AI Requirements?

Hyland is an enterprise content management vendor with a global customer base

Having discussed the requirements for content intelligence and AI, this section looks at how one vendor, Hyland, is stepping up to meet these. Hyland is an enterprise content management vendor and has a worldwide customer base across many different vertical industries. It provides software that supports:

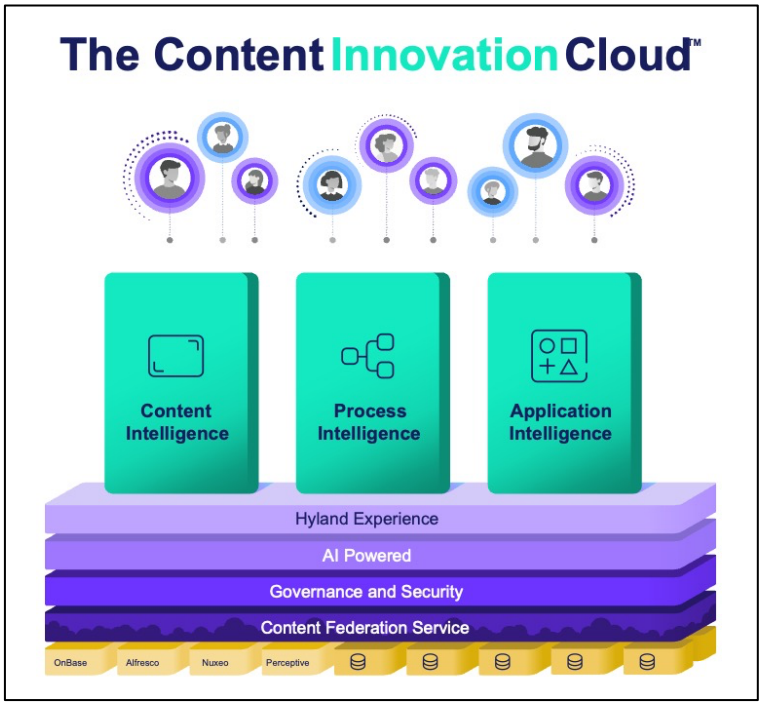
- Multiple types of unstructured content
- Multiple content stores to manage content
- Advanced search
- Content capture and processing
- Content governance including retention and records management, e-discovery, regulatory compliance, privacy and security
    - Content intelligence

Its platform can capture, process, manage, govern and analyse many types of content

Hyland's platform is the Hyland Content Innovation Cloud ™.

Hyland Content Innovation Cloud ™

This is a platform that is used to create a unified view of content across the enterprise so that content can be captured, processed, managed, analysed and integrated into processes and applications no matter where it is located. The Content Innovation Cloud is shown in Figure 4.



Hyland Content Innovation Cloud provides a unified view of content residing in multiple stores across the enterprise

Source: Hyland

Figure 4

This platform is used to capture, process and govern the flow of content in operational business processes and within applications. Users can also

produce and act on content intelligence which is the focus of this paper. Content Innovation Cloud capabilities include:

Query, analyse, predict, recommend and act on content intelligence

- Hyland Content Federation Service
- Hyland Governance and Security
- Hyland Content Intelligence

Capture, process, manage, store and govern, content

- Hyland Process Intelligence
- Hyland Application Intelligence

Hyland Content Federation Service

Hyland Content Federation Service is a shared service of the Content Innovation Cloud that simplifies access to content in multiple underlying content stores, and across networks so that people can access related content as if it was all in one central store. It provides secure connection and access (via brokers) to content stored in:

Hyland Content Federation Service simplifies access to content in multiple content stores

- Hyland OnBase, Alfresco, Nuxeo and Perceptive Content
- Third party content stores e.g. Microsoft SharePoint, Google Drive, IBM FileNet, Documentum, Box and others

Content can be processed, analysed and governed where it is - no need to migrate

There is no need to migrate existing content. You simply connect content sources to the Content Innovation Cloud and capabilities like content extraction, processing and content intelligence are all available to use across multiple connected content stores. In addition, you can evaluate content for potential content intelligence to help you see what to process to produce high-value insights.

Hyland Governance and Security

Federated content governance

Content can be discovered, classified, catalogued and governed. Policies can be enforced across multiple content stores – more on this later.

Hyland Content Intelligence

Hyland Content Intelligence enables you to curate, enrich, query and analyse content to produce content intelligence

The Hyland Content Intelligence product line consists of Knowledge Enrichment (data curation and context enrichment), Knowledge Discovery (AI-powered search), and Agent Builder (AI-Agents).

Knowledge Enrichment makes content AI ready so it can be analysed using natural language queries

Knowledge Enrichment is a data management suite that transforms raw, unstructured content into structured, high quality, contextual content for use in AI-based automation and application development. It helps organisations curate and make sense of their enterprise content by extracting contextual information, creating vector representations of that content, organising unstructured data and identifying document relationships to enrich content, bridge business intelligence gaps and support informed decision-making. Knowledge Enrichment supports the content curation over 600 file types across multiple sources.

As a result of Knowledge Enrichment, Knowledge Discovery understands content relationships ensuring that all relevant content is included when answering queries

Knowledge Discovery is an AI-powered search, information discovery and decision support application. Once Knowledge Enrichment has been done, it enables users across the enterprise to produce relevant business insights with a simple natural language query, using AI-Agents to retrieve and generate accurate information to accelerate decision-making.

Customers can also build their own AI-Agents to analyse content, automate tasks and recommend actions

Agent Builder is a no-code tool that enables organisations to build, configure, deploy, monitor and manage the lifecycle of AI-Agents to perform specific content-driven tasks or specialise in particular areas of the business, augmenting the human workforce and enabling organisations to implement AI at scale.

Hyland are also planning to ship multiple AI-Agents. whereby each AI-Agent can be configured to define the content it has access to, and also which users have access to which AI-Agents. AI-Agents use open source LLMs and non-open source LLMs to securely access content in multiple stores. This ensures that AI-Agents are governed. All LLM's are deployed on Hyland infrastructure and all models created on Hyland infrastructure are owned by the customer.

Hyland Process Intelligence

Design and orchestrate workflows to capture, process, flow, and store content at scale in Hyland and 3rd party content stores

Hyland's Process Intelligence product line includes several robust offerings for enterprise process automation and orchestration. At the core of Process Intelligence is Hyland Automate (Automate), Hyland's flagship orchestration tool. Featuring GenAI-powered prompt-based design tools and a BPMN-compliant low-code process modeller, Automate enables you to design and orchestrate operational workflows to process content residing in both on-premises and cloud content stores. It uses Hyland Content Federation Service to reach Hyland and non-Hyland managed content and supports workspaces with unified views of data and content related to specific process work items.

Hyland Process Intelligence includes Hyland Automate & Hyland IDP

Cognitive services, ML models and AI-Agents can be invoked to analyse, predict, recommend and act on content during process execution

Hyland Automate comes with a number of pre-built task automation and content processing capabilities plus the ability to invoke other Hyland and non-Hyland services. This includes Hyland IDP (Intelligent Document Processing) to execute document processing flows or any web service using a generic REST connector. This makes Automate very powerful. For example, it is possible to analyse content by invoking ML models and GenAI large language models during processing as part of an Automate workflow. AI-Agents can also be invoked to automate tasks, make faster decisions with fewer errors and generate content.

Automate flows can be scheduled, invoked on demand and on an event driven basis. Therefore, a flow could be triggered by the arrival of documents or by another event e.g., the approval of a new supplier.

Operational and analytical processes can be invoked on-demand, on an event-driven basis or scheduled

Hyland IDP is also part of Hyland Process Intelligence. It is used to process documents and includes support for optical character recognition (OCR), content classification, document separation, and extraction and validation of data.

Hyland Application Intelligence

Hyland application Intelligence enables seamless integration of Hyland operational and analytical content services into SaaS and custom-built applications

Hyland Application Intelligence is the integration of Hyland native content services with third party applications such as Microsoft Dynamics, Workday and SAP SuccessFactors. This includes the ability to capture documents, view related content, approve inbound documents (e.g., supplier invoices), access Content Intelligence, access AI-Agents and utilise IDP. With respect to analysing content, it is the embedding of Content Intelligence in these applications which is significant because it delivers content intelligence right to the point of need within the specific application in the precise context of tasks being performed. The same is true for AI, because AI-Agents can be invoked from within applications to automate and expedite tasks and to automate decisions.

HYLAND'S SOLUTION FOR UNSTRUCTURED DATA ANALYTICS AND AI

Now that we understand the Content Innovation Cloud, let's look how the platform supports unstructured data analysis and AI in the context of the requirements covered earlier. To simplify understanding of what Hyland offers, we will use the same headings as used in the requirements.

Connectivity to unstructured content and simplified access

Collections of federated content can be made available for analysis

The Content Federation Service hides the complexity of multiple content stores and networks from users by simplifying access to content in Hyland and non-Hyland content stores. It is also possible to define collections of content across multiple stores for analysis to control access.

Content discovery, classification, and cataloguing

Automatically discover, classify and catalogue content to make it easy to find and govern

To make content easier to find, Hyland offers the ability to automatically discover and classify content across multiple content stores. Hyland is also planning to enable unstructured content to be documented in third party data catalogs via APIs so that people know what content exists across a distributed data estate.

Sensitive content can be filtered out before making content AI-ready

In addition, it is possible to classify content upon capture using IDP and Automate. This includes automatic classification of sensitive data in content to filter out that content before making AI-ready.

Content curation

Content can be curated and made AI-ready using Knowledge Enrichment and orchestrated with Hyland Automate

Hyland Knowledge Enrichment is a set of REST-based APIs that are used to curate, clean, enrich, and transform unstructured content into AI-ready formats. This includes chunking documents, extracting metadata, generating vector embeddings, identifying entities and relationships, and organising content into structured, reusable collections. Curated outputs

can be used for retrieval-augmented generation (RAG), GraphRAG, or downstream analytics and automation.

Curation pipelines built with Knowledge Enrichment can include operations such as:

- Format standardisation and intelligent content segmentation
- Document classification and summarisation
- Entity extraction and metadata enrichment
- Vector creation and clustering of related content
- Knowledge graph population (planned)

These capabilities support content intelligence by enabling consistent, high-quality processing of over 600 supported file types.

Automate provides the orchestration layer that allows users to design and execute workflows incorporating Knowledge Enrichment, IDP, cognitive services, AI-Agents and other services including human-in-the-loop processing and decision approval. These workflows can be scheduled, triggered on events, or run on demand—automating document handling, content preparation tasks and decision support.

AI-Agents can be invoked within pipelines to process content for specific business needs

Both Hyland-provided and user-defined AI-Agents can be embedded in Automate workflows to analyse curated content, recommend actions, and accelerate insights. Agents can be built in Hyland Agent Builder or connected from external sources, and are governed to ensure secure, role-based access and content scope.

Curated content is published as high-value, reusable content products for downstream analysis and AI

Hyland's approach enables organisations to create curated content collections, such as customer interactions or case records, that are AI-ready and available across the enterprise for natural language querying, agentic workflows, or integration with business applications.

Content governance

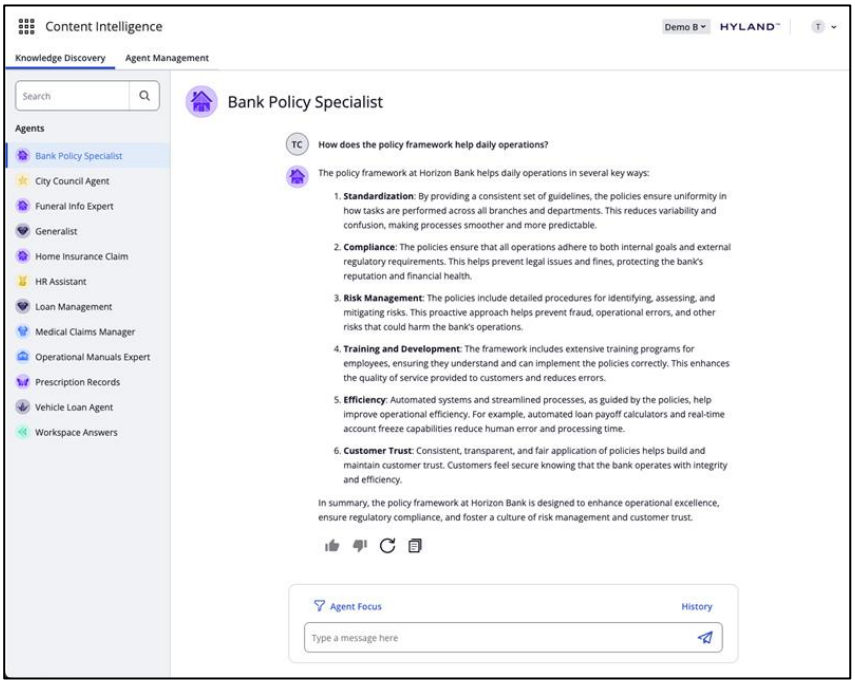
Content quality, privacy and access security can all be governed

Hyland can also improve content quality, obfuscate and redact content to protect sensitive data within it and govern access security to that content.

Content analysis

Hyland Content intelligence offers AI-assisted natural language query and analyse of content across multiple content stores

Users can use Content Intelligence to find and analyse related sets of content in multiple content stores to produce new insights. In response to natural language queries, Knowledge Discovery returns relevant files and will automatically translate from other languages. In addition, it can use tailored AI-Agents to retrieve and generate relevant insights, charts and natural language explanations of the visual insights to help you make faster decisions (See Figure 5)



Source: Hyland

Figure 5

Machine learning and NLP can be used to find patterns and provide on-demand predictions and recommendations

It is also possible to use ML and NLP, to sift through unstructured content to find patterns, make real-time predictions and recommendations.

Agent Builder lets you build your own AI-Agents to automate tasks and decisions

Furthermore, Agent Builder lets you create and deploy custom configurable AI-Agents for use in processing and analysing content. For example, you could create agents that do voice-to-text, language translation, content summarisation, and sentiment analysis. Also, departmental AI-Agents focussed on customer content for use in marketing, sales or service, or employee related content for use in HR.

Analytical agentic workflows are also possible and can be scheduled, invoked on-demand or on an event-driven basis

Beyond this, it is possible to invoke AI-Agents from Automate workflows to process and analyse unstructured content in batch and on an event driven basis to produce specific insights, drive alerts, recommendations and even automated actions. Automate agentic workflows can also be published as services for invocation on-demand.

INTEGRATING HYLAND CONTENT INTELLIGENCE WITH TRADITIONAL ANALYTICAL ENVIRONMENTS

Hyland Content Intelligence can be integrated with existing analytical environments to add new insights to what companies already know

It is possible to integrate Content Intelligence with traditional analytical environments to add new insights to what companies already know. This can be done in several ways. The first is using Automate. Automate offers powerful capabilities to build workflows that process unstructured content. It can utilise large language models (LLMs) and data extraction tools to enrich and contextualise digital assets. It can also connect with critical applications like Salesforce, Workday and SAP to access, process and extract data from invoices, purchase orders, contracts, employee documents, online forms and more. It also has a rules engine to support dynamic, complex processing of unstructured content and to support decision automation. Furthermore, it can be used to build AI-agentic workflows.

Invoke Hyland Automate content preparation and analytical flows from traditional data integration tools to produce richer data

This leads to more informed and effective decision making

When integrating Content Intelligence with traditional analytical environments, it is possible to exploit the power of Hyland Automate workflows and AI-Agents by invoking them as part of a data integration process from within 3rd party data integration workflows to extract data from unstructured content to integrate with structured data from other sources. This combination can help produce richer, more complete structured data products for analysis to yield richer insights.

In addition, content collections processed by Automate workflows can be published as unstructured data products in Data Marketplaces from where they can be shared and reused. The same is true for AI-Agents. It is also possible for traditional BI tools to route natural language queries to Content Intelligence to get additional insights.

INTEGRATING HYLAND CONTENT INTELLIGENCE INTO BUSINESS OPERATIONS

Hyland Content Intelligence, AI-Agents and agentic workflows can be integrated with SaaS and custom-built applications

Once unstructured content has been analysed to produce actionable intelligence, the next step is integrating that intelligence into day-to-day operations to support informed decision making within business applications.

Hyland Content Intelligence, AI-Agents, and Hyland Automate workflows can be embedded into both SaaS platforms and custom-built applications.

There are two primary integration paths:

- Out-of-the-box integrations: Hyland Application Intelligence provides pre-built connectors that integrate Hyland Content Innovation Cloud services, including Content Intelligence, into widely used SaaS applications like Microsoft Dynamics, Workday, and SAP SuccessFactors.
- APIs for custom integration: For other operational systems, you can invoke Content Intelligence queries, AI-Agents, and Hyland Automate workflows using Hyland's robust APIs. This allows you to embed intelligence into any business process or application environment.

Hyland Content Intelligence, predictions, and recommendations can support 'right time' decisions in context at the point of need

This level of integration ensures that insights, predictions, and recommendations are available in context right when and where they are needed. It goes beyond simply accessing insights through a Hyland interface and brings intelligence into the tools and applications people are already using, helping them make better decisions without switching context or interrupting their workflows.

VERTICAL INDUSTRY USE CASES – WHAT'S POSSIBLE?

Following on from the benefits discussed earlier, what follows are some examples of how analysis of unstructured data can help in the financial services, healthcare and insurance industries.

FINANCIAL SERVICES

In financial services, analysis of unstructured data can help to reduce risk, reduce costs, improve efficiency and personalise customer experiences.

Analysis of emails, contact centre chat and voice calls can produce content intelligence to help identify money laundering and combat fraud

For example, unstructured data such as emails, contact centre chat and voice calls can be analysed alongside structured transaction data to determine and verify customer identity, and uncover new insights and predictions that would help identify money laundering and combat fraud. Patterns found in emails, chat and voice can be combined with transaction monitoring to reduce the risk of financial crimes.

Content Intelligence can also be used in trading surveillance to help prevent insider trading

Also, for compliance, trading surveillance can be introduced by analysing emails, web chat logs and phone calls to prevent insider trading. The same is true in cybersecurity were analysis of emails, web chat logs and phone calls to help in the detection of insider threats. Also using AI to detect and classify sensitive unstructured data is needed for data privacy.

Analysis of customer interactions to identify customer sentiment and intent helps to improve personalisation

With respect to customers, it is often the case that chatbots and people involved in a customer journey may not have access to an up to date "living customer intelligence profile" for each customer. This is needed across all channels for better personalisation and customer service. Analysis of customer emails, call centre transcripts, chatbot interactions and social media interactions helps identify customer sentiment, likes, dislikes, intent etc., providing new insights to enrich these profiles which are used to help personalise customer experiences.

Content intelligence produced from analysing content during research and due diligence leads to faster and more informed investment decisions

With respect to investment banking, research and due diligence prior to an investment decision can be labour intensive and involves looking at unstructured content such as company earnings publications, analyst financial reports, research reports, and financial news in multiple content stores together with social media. Content Intelligence and Content Federation Service would make it quick and easy to produce insights from this content which could be combined with stock movement predictions to enable quicker more informed investment decisions.

AI-automation improves efficiency and lower costs

It is also possible to use AI to automate tasks to improve process efficiency e.g., capturing, processing and storing of loan applications.

HEALTHCARE

Healthcare has a lot of unstructured content and data types

In healthcare, a significant amount data is unstructured, primarily in text or image form. This includes data from handwritten clinical notes on patient symptoms and observations, and diagnostic medical images e.g.,

X-rays, MRI's, CT scans. This unstructured data is typically associated with patient narratives and recorded in multiple content stores during consultations.

Help optimise patient journeys and improve patient outcomes

Analysing unstructured data can be used to help optimise the patient's journey and treatment pathways, enhance clinical care and patient outcomes, predict a patient's future needs, and provide real world evidence of drug efficacy and safety.

Provide real world evidence for drug efficacy and safety

Help clinicians make better care decisions

For example, Automate, IDP and NLP can be used to rapidly classify medical records, create a knowledge graph of patient related content, and automatically extract medical terms and conditions from patient narratives to help clinicians make better care decisions. AI can also be used to automate image interpretation to improve the speed and accuracy of diagnosis.

Quickly identify high-value insights from patient records, spot, trends, make recommendations and generate treatment plans

Expedite pre- and post-visit processes

Content Intelligence and the Content Federation Service can be used to give clinicians quick access to years of related clinical data across multiple content stores from a conversational user interface. Also, AI-Agents can identify high value information from patient records, spot trends and recommend actions. In addition, they can support medical assistants in preparing patient charts, identifying care gaps, highlighting important screenings and in automatically generating pre- and post- visit patient correspondence.

Improve quality of patient care and treatment effectiveness

Faster and richer real-world evidence

Sentiment in patient emails, text messages and social media interactions can also be analysed to produce insights about different aspects of healthcare delivery including quality of patient care, treatment effectiveness etc. Also, NLP can be used to automatically extract first drug reactions from doctor's notes, patient interactions and social media posts around drug effects to analyse patient feedback to provide real world evidence of drug efficacy and safety.

INSURANCE

Help to reduce premium leakage and reduce fraudulent claims

In insurance, underwriting fraud occurs when people provide false information on insurance applications to obtain cover at reduced premiums. This results in 'premium leakage' which costs insurers billions. In claims, fraud from overstating losses is also an issue. Also, underwriting fraud is often a major indicator for fraudulent claims.

Faster claims processing

Improve customer and employee satisfaction

Furthermore, the claims process involves unstructured data such as claims forms, photographic images, video, claims assessor reports etc. Collection of this is often manual, data is often missing, and slow processing leads to customer and employee dissatisfaction. Intelligent document processing, ML and AI can be used to automate document handling, assist in image cleansing, data correction, and critical data extraction and fraudulent claims detection for faster claims processing.

Analysis of claims, call centre transcripts, and social media activity can also help spot inconsistencies to detect fraudulent claims.

CONCLUSION

Data and AI are now critical to everyday business decision making and task automation

Unstructured data is an untapped source of high value insights

Fractured unstructured data is hindering progress

GenAI has made it possible to democratise production of content intelligence

Companies need a unified content intelligence and AI platform that hides the complexity of fractured content, facilitates content cleansing and curation and that can democratise production of content intelligence and AI-Agents

Content intelligence and AI-Agents need to be integrated with traditional analytical environment and also integrated into applications and processes across the enterprise to maximise value

We are now in a point where the use of data and AI is considered critically important in everyday business to help to ensure highly effective, timely decisions. However, the people making those decisions are doing so using insights produced from analysing structured data. But with as much as 80% of enterprise data in unstructured form, many companies are now looking at new value that can be created from insights produced from querying and analysing unstructured content. The problem is that this content is scattered across multiple systems, many of which are poorly managed. This makes it difficult to find, query and analyse relevant content. Furthermore, until recently, analysing complex unstructured data was left to a small number of highly skilled data scientists. Today, however, since the emergence of generative AI, it is a lot easier.

For many, companies analysing unstructured content is practically a green field opportunity. What they need is a platform offering unified content intelligence and AI that hides the complexity of multiple internal and external content sources, enables them to easily find, and curate relevant related content and make it available for analysis.

In terms of analysis, it should be easy to democratise the production of content intelligence through the use of natural language queries as well as building and utilising predictive and prescriptive machine learning models and generative AI-Agents. In addition, insights produced from analysing content need to be added what companies already know from analysing structured data. This provides a holistic set of insights for more informed and effective decisions. Finally, these insights and AI-Agents need to be integrated into applications and processes across the enterprise to ensure maximum contribution of those making operational, tactical and strategic decisions towards improving business outcomes.

Looking at the Content Innovation Cloud through an analytical lens, it includes much of these needed capabilities. Content Federation Service hides the complexity of multiple content sources, and content can be easily classified and governed. Knowledge Enrichment curates, enriches, and prepares unstructured content for AI by extracting metadata, generating vectors, and organising related content into reusable collections. These curation workflows can be orchestrated using Automate, which coordinates services like IDP and AI-Agents to process and analyse content at scale. Content Intelligence offers natural language queries, AI-Insights and AI-Agents while Application Intelligence enables integration of intelligence into applications. So, it is clear that Hyland is rapidly building out a unified content intelligence and AI capability. For these reasons, it is worthy of being shortlisted as a candidate provider of unstructured content intelligence and AI in a data and AI-driven enterprise.

APPENDIX A – DETAILED REQUIREMENTS FOR SUCCESSFUL IMPLEMENTATION OF UNSTRUCTURED CONTENT INTELLIGENCE AND AI

The following is a detailed set of requirements needed for successful implementation of unstructured content intelligence and AI.

CONNECTIVITY AND SIMPLIFIED ACCESS REQUIREMENTS

As shown in Figure 1, unstructured content is often scattered across multiple clouds and on-premises content stores and file systems. Therefore, it should be possible to:

Connectivity is needed across a broad range of content stores that are both on-premises and on multiple clouds

- Connect to a broad range of content sources including:
  - Multiple content management systems
  - Records management systems
  - On-premises file shares (e.g., Windows file servers)
  - Cloud file shares (e.g., Microsoft SharePoint and OneDrive, Google Drive, Box, Dropbox etc.)
  - Cloud storage (e.g. Amazon S3, Azure Blob Storage)
  - E-mail systems that are on-premises and / or in the cloud
  - Collaborative tools (e.g., Google Meets, Microsoft Teams, Slack) that record and store web chat, video meetings etc.
  - Customer service interactions (e.g. web chat, transcripts)
  - Social network interactions (e.g. posts on X, Facebook etc.)
  - And more

This includes content management systems, file systems, email, web chat, collaborative tools, social networks and more

Users should not have to know where content is located

- Simplify access to multiple underlying content stores both on-premises and in multiple clouds by supporting federated content management

Query and search across multiple content stores

- Define and query / search user defined collections of related content that reside in one or more underlying content stores

CONTENT DISCOVERY, CATALOGUING, AND CLASSIFICATION REQUIREMENTS

To help users understand and govern what content is available for analysis, without looking in every store, it should be possible to support:

Companies need an enterprise content catalog that can automatically scan and classify content in multiple content stores so that users know what content exists and what content is sensitive

- Automated discovery, metadata inference / generation and cataloguing of content to create an enterprise content catalog and knowledge graph of all the unstructured data assets and relationships that exist within and across multiple content stores
- Automated classification of content containing sensitive data using pre-built content classifiers so that sensitive content can be quickly detected to enable it to be protected (e.g., via obfuscation) before:
  - Creating vectors for retrieval augmented generation (RAG) when using GenAI LLMs in order to avoid 'PII leakage'

    - Making that content available for analysis
- User-defined classification schemes and trainable classifiers to label content to make it easier to find and govern across multiple content stores (e.g. content considered a trade secret). This includes classification of multiple levels of confidentiality and retention

CONTENT CURATION REQUIREMENTS

With respect to content curation, it should be possible to:

Workflow for curation pipelines

- Allow users to define domain specific metadata models for content
- Support workflow to create content curation and preparation pipelines also analytical pipelines
- Use this workflow capability to create curation pipelines that:

Metadata enrichment

    - Use predictive models or fine-tuned LLMs during execution to infer metadata from content so it can be more easily found and content relationships identified
    - Chunk content to enable vector creation in support of GenAI

Make content AI-ready

    - Transform content and metadata into structured, chunked, formats and generate vector embeddings and relationships. This enables techniques like Retrieval Augmented Generation (RAG) and GraphRAG to surface contextually relevant content for accurate, responses to natural language queries

Understand related content

    - Automatically build a knowledge graph from inferred metadata to understand related content across multiple content stores

Clean and prepare content for analysis

    - Support pre-built content cleansing and preparation services that can be used in pipelines to prepare unstructured content for analysis. Examples of such services include:

Cater for voice, text in multiple languages

        - Voice-to-text
        - Multi-language translation

Group similar content

        - Content cleansing – e.g., Spell checking, noise removal (e.g. emojis, hashtags), stop word removal etc.
        - Document clustering using machine learning to create related content collections and facets prior to analysis

Summarise content and extract data from it

        - Content and content metadata enrichment
        - Content summarisation using generative AI-Agents
        - Entity extraction to extract structured data from text

Process documents at scale and automate tasks

        - Intelligent Document Processing
        - And more

    These kinds of capabilities enable the finding, preparation and analysis of related sets of relevant content across multiple content stores to be more rapid and accurate.

Publish curated content that is ready for analysis and AI

- Support the creation and publishing of curated, compliant, reusable, unstructured data products in a data marketplace that represent curated content collections (e.g., customer interactions - emails, chat, voice, social) that exist in multiple content stores

Allow people to create ontologies representing bodies of knowledge

- Enable users to model their own ontology in a knowledge graph

CONTENT GOVERNANCE REQUIREMENTS

It should be possible to define policies in a content catalog and enforce them across a distributed data estate to govern content. This includes:

Support the creation of policies to govern content across multiple content stores

- Content quality validation rules
- Content privacy policies
- Content access security policies
- Content retention policies
- Content sharing policies and contracts
- Usage monitoring
- AI governance policies

CONTENT ANALYSIS REQUIREMENTS

With respect to analysing content, it should be possible to support:

Support conversational query and analysis

- Conversational query and analysis across multiple content stores via a conversational prompt-based user interface

Visualisations and natural language explanations

- The creation of content intelligence (including visualisations and natural language explanations of business impact) using search, ML and generative AI on multiple types of content in multiple stores

Pre-built AI-Agents to score sentiment, summarise and associated with different business domains

- Out-of-the-box AI-Agents to expedite content analysis and action taking e.g.,
    - Sentiment analysis
    - Document summarisation
    - Domain specific AI-Agents (e.g., Marketing, HR, Finance) that recommend and take actions using intelligence from content from within interactive conversations

Create your own AI-Agents to analyse and act on content intelligence for a specific business purpose

- The creation of user-defined analytical AI-Agents e.g., agents for:
    - Trend analysis on document clusters
    - “Voice of the customer”
    - Competitor analysis
    - IT service optimisation

Create powerful agentic workflows automate tasks and decisions

- Orchestration of agentic workflows to produce new insights in batch, on-demand and on an event-driven basis. This can help to automate tasks, decisions and actions

REQUIREMENTS FOR INTEGRATING STRUCTURED AND UNSTRUCTURED DATA ANALYTICAL ENVIRONMENTS

Integrate structured and unstructured analytical environments to create a holistic set of intelligence, predictions, alerts and recommendations for more effective decision making

Most companies have been analysing structured data for years to produce reports, predictions, alerts and recommendations. Therefore, there is a need to be able to integrate content intelligence and AI-Agents with traditional analytical environments so that the new insights, predictions, alerts and recommendations coming from analysis of unstructured content can be added to what organisations already know. The result should be an enriched, holistic set of intelligence, predictions, alerts and recommendations to support better, more effective decision making. Requirements to make this possible include the ability to:

Engineer data from structured and unstructured sources

- Publish content curation and preparation workflows as services
- Invoke content preparation workflows from within enterprise data fabric data engineering pipelines

Publish reusable curated content, content intelligence, ML models and AI-Agents in the same place as those associated with structured data

- Publish curated federated content collections as unstructured data products in a Data Marketplace
- Publish content intelligence natural language queries and AI-Agents as services in a Data Marketplace

Enable BI tools to query content to add new insights to what is already known

- Invoke content queries as services from Data Fabric pipelines, Python, R and traditional BI tools
- Invoke BI services, ML models and content AI-Agents from within content processing workflows and from data fabric pipelines

Exploit BI, ML and AI-Agents in content processing

Integrate classic BI and content intelligence

- Integrate classic BI and content intelligence by:
  - Providing access to federated content from classic BI tools
  - Invoking content analytical AI-Agents and recommended content intelligence queries / prompts from popular BI tools
- Publish content intelligence as services in data marketplaces

Event-driven analysis of unstructured content and automated actions

- Use orchestration to create event-driven pipelines that process and analyse unstructured data and that can use prescriptive machine learning (ML) models to drive automated alerts, recommendations and actions

OPERATIONAL CONTENT INTELLIGENCE INTEGRATION REQUIREMENTS

In order to integrate unstructured content intelligence and AI-Agents into business operations it should be possible to:

Integrate content intelligence and AI into applications and processes

- Integrate unstructured content intelligence and AI-Agents into operational business processes to leverage content intelligence on demand and to automate tasks
- Integrate unstructured content intelligence and AI-Agents into business applications to:
  - Provide contextual, conversational intelligence, predictions, recommendations and actions
  - Automate tasks and decisions

About Intelligent Business Strategies

Intelligent Business Strategies is an independent research, education, and consulting company specialising in data management and analytics. Our goal is to help companies understand and exploit new developments in business intelligence, machine learning, artificial intelligence, data management, big data, and enterprise business integration. Together, these technologies help an organisation become an *intelligent business*.

Author

Mike Ferguson is CEO of Intelligent Business Strategies Limited. As an independent IT industry analyst and consultant, he specialises in BI / analytics and data management. With over 40 years of IT experience, Mike has consulted for dozens of companies on BI/ML/AI, data strategy, technology selection, data architecture, and data management. Mike is also conference chairman of Big Data LDN, the largest data and analytics conference in Europe and a member of the EDM Council CDMC Executive Advisory Board. He has spoken at events all over the world and written numerous articles. Formerly he was a principal and co-founder of Codd and Date Europe Limited – the inventors of the Relational Model, and Chief Architect at Teradata on the Teradata DBMS. He teaches popular master classes in Data & AI Strategy, Data Architecture Modernisation, AI-Assisted Active Data Governance, Practical Guidelines for Implementing a Data Products, and Embedded Analytics, Intelligent Apps, AI-Agents and AI Automation.

Telephone: (+44)1625 520700
Internet URL: www.intelligentbusiness.biz
E-Mail: info@intelligentbusiness.biz