

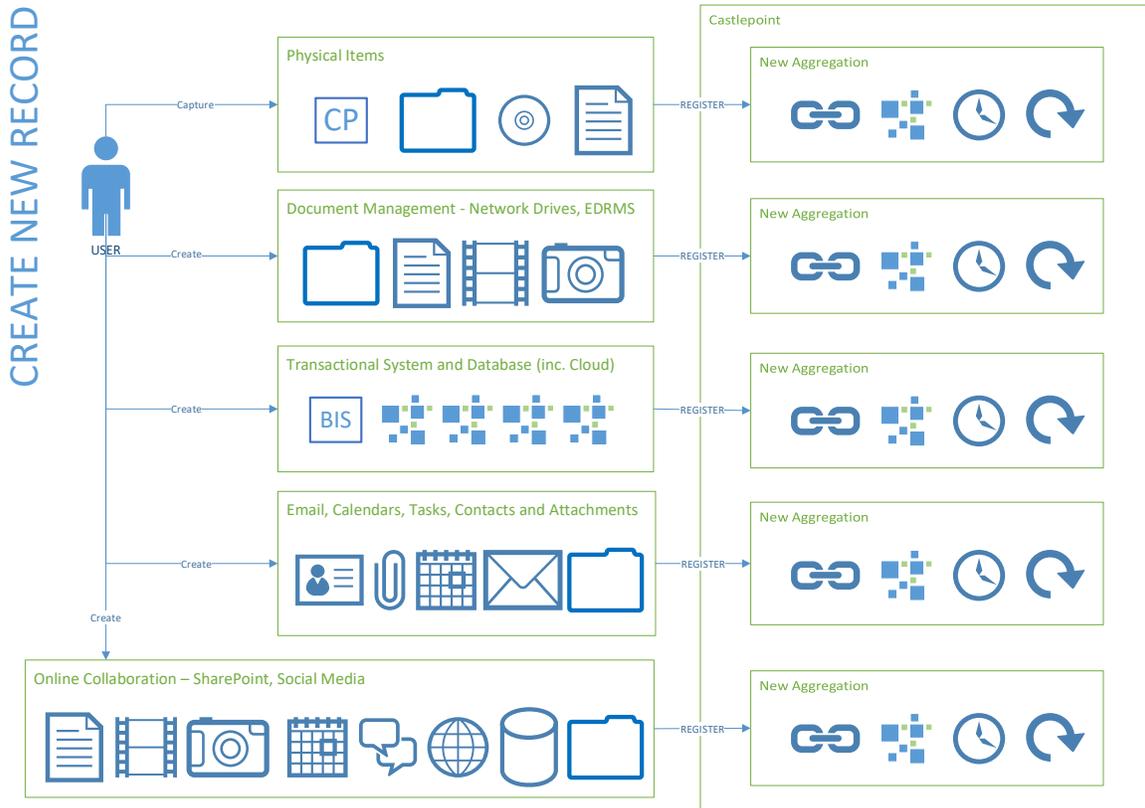
# CASTLEPOINT MANAGE INFORMATION EVERYWHERE BLUEPAPER

## SOLUTION OVERVIEW

Developed in response to the challenges seen with both traditional EDRMS and Business Information System (BIS) compliance, MIE introduces a new paradigm for digital continuity. MIE is a central system that loosely couples with BIS without the requirement to make technical changes to those systems. MIE can interact with any business information system via microservices, and manage both structured and unstructured records.

Some examples of common BIS that Castlepoint MIE can manage include

- Document management systems including network drives, Dropbox, Docusign and Google Drive
- Email systems including Exchange and Gmail
- Collaboration systems including SharePoint and Office 365
- Databases including Oracle, SQL Server, DB2, Ingres, and other ODBC-compliant databases
- Social including EventBrite, Facebook, Instagram, SurveyMonkey, Twitter, Wordpress, Yammer, YouTube
- Line of business systems including Dynamics, GitHub, Jira, PostgreSQL, and Zendesk
- Other cloud systems including SalesForce, GoTo Meeting, Webinar and Training, Google Calendar and Stripe.



MIE works by silently monitoring the BIS for new items (documents, emails, tweets, rows, web pages, images or other records). When a new item is identified, MIE uses artificial intelligence (AI) to either match it to an existing record aggregation, or create a new aggregation. No user interaction is required (including additional metadata input).

MIE uses AI to understand the context of the record (who created it, and where in the organisation it arises from) as well as content analysis of keywords present in the text. It uses this information to automatically classify the record aggregation against a Records Authority, and any relevant ontologies.

MIE applies the sentence from the Records Authority when it is registered, but also continually over time as the content of the aggregation evolves and changes, to ensure the classification remains appropriate. When the retention period is met, MIE alerts the records managers, who can review the aggregation and execute the sentence.

MIE supports compliant destruction, transfer and import and export of records, and provides a full set of reports on all records across the enterprise. MIE also adds value to users by providing an interface where they can search for other records in the organisation that relate (contextually or in their content) to the ones they are working on.

## THE TECHNOLOGY

### Business Information Systems Compliance Requirements

Agencies have a large amount of Business Information Systems. On average, an organization will have

- Over 1,000 separate business information systems containing important records
- Over 140TB of business information systems data
- Millions to billions of digital records in business information systems

Agencies surveyed by the National Archives report being prevented by cost, technical overhead and staff impacts from ensuring they meet the functional requirements from the Standards. To meet the functional requirements, each BIS must be given the ability to:

- **create records in context:** the functionality required to create and keep records that are valid, accurate and protected from both accidental and deliberate unauthorised change.
- **manage and maintain records:** the functionality required to ensure records do not lose their evidential value during management processes including classification, retention, disposal, security and access.
- **support import, export and interoperability:** the functionality required to exchange, import or export records both within and between systems, without the loss of content or metadata.
- **retain and dispose of records as required:** the functionality required in business systems to dispose of records in a managed, systematic and auditable way.

MIE has the ability to meet these requirements, and to:

- encompass millions of records in thousands of systems, both structured and unstructured
- make the management of records in those systems compliant with the standards for their entire life cycle
- avoid technical customisation or complex integration of existing BIS and EDRMS
- avoid duplication of records and expansion of records storage
- provide a simple interface for records managers to undertake their control activities
- ensure that controls for records management compliance are invisible to the user base, with no disruption
- automate the process of registration, classification, and sentencing per Department of Finance expectations

## Solution Functions

Castlepoint Manage Information Everywhere (MIE) is a solution for governing large data sets across diverse systems through a central portal. It is a single product that can be deployed on premises or in the cloud. It provides the following high level functions:

### Records Management

MIE provides fully compliant records management for all BIS in accordance with National Archives requirements.

1. **Register records:** Every new record created in the enterprise is automatically registered in MIE using microservices, at the point of capture or creation.
2. **Classify:** Each record is scanned by the Artificial Intelligence (AI) engine and given a Records Authority classification based on analysis of its context (logical structure, file type, creator organisation role, source system type) and content (body text, key words, metadata). MIE applies a sentence, setting an appropriate retention period, trigger mechanism and sentencing action rule.
3. **Index:** MIE uses an ontology to categorise the record (based on metadata, extracted keywords and text) for ease of discovery, and to help relate it to other relevant records in other source systems that are also connected to MIE.
4. **Manage:** Whenever the original record aggregation, or an item within it, is modified in the source system, MIE updates the sentence as appropriate. For example, a record with a retention of 7 years from a trigger of 'date of last action' will restart the 7-year timer whenever an item in the aggregation is updated. Additionally, as the record content and usage changes over time, MIE will review the classification and may amend the applicable Class.
5. **Disposition:** When the retention requirements of the record, and any related records, are met, Castlepoint automatically alerts the records management team using workflow, who can review the aggregated record and verify the applied sentencing action. A sentencing control record is updated with disposition details. Castlepoint can also support compliant conversion of records to long term preservation formats.
6. **Report:** MIE provides an interface for records and information managers to view all of the registered records, wherever they are stored, even where the source systems have inadequate reporting capabilities. Records managers can use the reporting dashboard and dynamic reports interface to view record management statistics, and to view, verify and manage all records lifecycle actions, including records due for disposition or review now and into the future.
7. **Control:** MIE also allows records managers to easily identify any 'lost', unsentenced or unclassified records where these actions have not been able to be automatically completed, and to help refine the machine learning capabilities of the AI system via assisted learning. Records managers can also use the interface to find and restrict or freeze related records across multiple systems, or manage records in legacy, inactive or even decommissioned systems. MIE can be used intra or interagency to understand record holdings across government and industry.

### Information Management

MIE provides high value IM capability for business to strategically manage information assets and extract best value from information holdings across the enterprise.

1. **Discovery:** Using RESTful services accessible by authenticated clients, and the OpenSearch 1.1 compliant search system and results display, Castlepoint helps normal users discover relevant, complete and useful content and context in dispersed systems or structures they would not otherwise be aware of (particularly email and shared drives). Castlepoint can expose relevant records to users without breaking security

controls, promoting better information use and re-use.

2. **Audit:** Using the discovery capabilities and inbuilt reports, find all information about a topic of concern, and drill down to see what changes were made and when, who accessed the information, and where it has been sent. Find and review all enterprise information even where it is stored in legacy or obsolete systems.
3. **eDiscovery:** Use the MIE interface to find and contain records for legal discovery and hold purposes. Manage eHolds in an MIE manifest, with details of current users and editors to contact and engage with. Flag attempted changes and deletions to records under freezes.
4. **Value chain:** Use MIE to enable a coherent view of all information assets, and relate them together for best use and reuse. Synthesise data into valuable information and enable better decision making and efficiency.
5. **Data hygiene:** Enable data cleansing, quality management, deduplication, and management of data exceptions. Identify data outliers that are not related to any other assets and are possible candidates for removal as redundant, obsolete or trivial.
6. **Data visualisation:** Display and explore entities, concepts and relationships within the full scope of your enterprise data. Present information about information assets in ways that can't be communicated textually. Identify patterns, relationships and trends that are not otherwise predictable.
7. **Data governance and policy management:** support control of data integrity, confidentiality, availability and usability of data by enforcing policy conformance. Allow data stewards and custodians to fully understand their assets across systems, and apply Master Data Management principles to standardise metadata.
8. **Data integration:** Provide a unified view of data from heterogeneous sources for research, analysis, meta-analysis, business intelligence, and management of large scale ETL for data migration.
9. **Schema management:** Implement standard schema properties and relations to existing data as well as creating your own schema to apply to your information. Include internationally recognised schema such as OWL2 as well as region and industry-specific schema like AGRIF and AGLS.

## System components and features

Castlepoint MIE leverages the following technologies.

1. **RESTful microservices:** MIE is built on a Microservices architecture hosted on the Azure API stack. Castlepoint services are integrated using loosely coupled RESTful connections between services and applications. The Castlepoint catalogue pipeline can be configured to include other available microservices as required by each organisation.
2. **Artificial Intelligence:** MIE utilises deep learning neural-net trained models to identify, categorise, and classify entities, relationships, documents and metadata across all monitored systems. The AI is continuously working in the background to ingest additional data and trained information sets, allowing Administrators to refine and improve the accuracy, scope and scale of the systems it monitors.
3. **Context and content extraction techniques:** MIE uses a number of text and natural language processing services as part of its content pipeline. The content pipeline involves a combination of data analysis, content extraction, context identification, context matching, and context cataloguing. Castlepoint uses Language API, Machine Learning, and Context Cataloguing services built on the Castlepoint microservices architecture.
4. **Cloud:** Castlepoint services are hosted on Public cloud (Unclassified: DLM Microsoft Azure) and Private Cloud (UD or Protected Vault Systems OpenStack, Microsoft Azure Stack). Castlepoint can also be provided

on premises where cloud is an unsupported model for an organisation.

5. **Open source:** Castlepoint uses a variety of open source technologies including Microsoft .NET Core, OpenStack, Docker, SipHash, Microsoft Coco Framework, Apache Tika, Newtonsoft Json.NET, REST, SVG, X3D, and Swagger and industry-standard machine learning algorithms for content and context investigation.
6. **Automated deployment framework:** MIE uses a continuous DevOps toolchain as part of its build, deploy and manage framework. User requirement designs are integrated into the Castlepoint microservice architecture; testing and quality assurance activities control the release schedule; and release automation with GitHub and Azure deployment framework includes no downtime required for updates to the components.
7. **Application integration:**
  - a. **Windows File System:** internal network drives and cloud-hosted network drives are integrated via product data feeds or via file system discovery. No standard user interaction is required with the MIE components.
  - b. **Exchange:** MIE automatically registers emails in nominated mailboxes to manage emails as they are received. MIE also includes an Outlook add-in component allowing native drag-and-drop integration within the Outlook client when manual registration is required.
  - c. **SharePoint:** SharePoint 2010, SharePoint 2013 and SharePoint 2016 information can be managed in-place using product data feeds read by the Castlepoint MIE system. No standard user interaction is required with the MIE components.
  - d. **Office 365:** integration is supported with SharePoint Online, Exchange Online, Dynamics 365, Azure Active Directory, Azure Blob Storage. No standard user interaction is required with the MIE components.
  - e. **Database:** integration is supported with Oracle, SQL Server, DB2, Ingres, and other ODBC-compliant databases. No standard user interaction is required with the MIE components.
  - f. **Other:** social media including EventBrite, Facebook, Instagram, SurveyMonkey, Twitter, Wordpress, Yammer and YouTube; line of business including Dynamics, GitHub, Jira, PostgreSQL, and Zendesk; Document management including Dropbox, Docusign and Google Drive; and other cloud systems including Salesforce, GoTo Meeting, Webinar and Training, Google Calendar, Gmail and Stripe.
8. **Keyword extraction and automated sentencing algorithms:** The MIE context pipeline includes keyword extraction and automated sentencing is built-in to the pipeline. Keyword and context information are used to determine the correct sentencing to apply using a unique aggregation automation process.
9. **Ontologies:** MIE manages entity data in highly-scalable NoSQL storage, including schema information and relationships. Ontology context and relationships are automatically extracted from objects managed by MIE without requiring users to link data or enter additional information. MIE APIs allow users to query over the relevant schemas and return relevant results semantically or graphically.
10. **Blockchain:** MIE includes integration with Microsoft Azure Blockchain-as-a-Service for cross-organisation blockchain contract ledger operations. The MIE blockchain integration allows records managed by MIE to be transacted on and transferred between organisations, while recording the contract exchanges in the blockchain ledger.
11. **W3C recommendation-compliant linked-data approaches:** MIE stores facts and values for managed data within the system. Facts and values are automatically linked by the system following W3C compliance.

Links are maintained in a highly scalable NoSQL storage system allowing schema to be managed at the entity level; this removes issues related to table join query impacts and link management. Links are automatically maintained by the system and updated as linked data changes.

12. **Scalability:** Castlepoint is hosted on the Microsoft Azure public cloud infrastructure, and is able to scale under a service consumption model. Microsoft provides the following general advice on upper limits: 500 TB per storage account (including all data stored in the Azure Storage account), 200 storage accounts per subscription (multiple are possible per customer), 4.75TB maximum file size in Blob storage, 20,000 requests per second Blob storage/retrieval, 60MB per second file throughput, 1,000 IOPS per share, 500,000 user accounts per tenant. MIE services scale with Microsoft Azure following the function consumption model. MIE is also able to scale on private cloud and on-premise deployments subject to the underlying infrastructure supporting the required scale.
13. **Business Continuity:** Castlepoint is built on the Microsoft Azure cloud service, offering 99.9% financially backed SLAs. Castlepoint can be deployed on private cloud and on-premise infrastructure as well, subject to technical fit. Castlepoint data is stored on redundant high availability Microsoft data centres.
14. **Compliance:** Castlepoint meets all relevant Information Security Manual 2016 policy and controls; NAA legislation, regulation and standards; and the Digital Service Standard.
15. **Third party options for connectors, automation and interfaces:** Castlepoint integrates directly with the Office 365 store, and can be used by other services. Subscription to the Castlepoint APIs is available in public clouds using an API consumption model.

## Meeting the requirements

To be useful and usable, an enterprise records and information management solution must have the ability to:

- encompass the millions of records in thousands of systems, both structured and unstructured
- make the management of records in those systems compliant with the standards for their entire life cycle
- avoid technical customisation or complex integration of existing BIS and EDRMS
- avoid duplication of records and expansion of records storage
- provide a simple interface for records managers to undertake their control activities
- automate the process of registration, classification, and sentencing per Department of Finance expectations

MIE meets all of these essential requirements, and adds additional value without technical overheads or high costs.

## IMPLEMENTATION

### Deployment

Castlepoint MIE services can be deployed either on-premise or on a cloud-hosted environment. Castlepoint MIE operates on the Microsoft Azure stack and requires the following environment:

1. Microsoft Azure storage services (Table, Blob, and Queue data storage)
2. Microsoft Azure application services (Microservices)
3. Microsoft Azure virtual machines (Machine learning trained models)

All specifications for the configuration including compute, storage and network requires are included in our documentation and are based on your data management requirements and projected data growth.

For cloud-hosted clients, you can deploy the Castlepoint MIE services using their own Azure subscriptions, create a new subscription, or utilise our Azure hosted environment in a fully out-sourced SaaS configuration.

For on-premise and private data centre hosting, the Azure stack provides a complete environment available as a fully compatible environment to host the Castlepoint MIE services and infrastructure.

## Integration

Castlepoint MIE services operate as both a set of background microservices as well as providing a front-facing user interface for users and administrators. Integration with Castlepoint MIE involves two primary activities:

### Log Analysis Engine

MIE automatically connects to logging services, log export locations, or application logging services to extract, analyse and record entity and metadata activities in monitored systems. Systems are registered in the Castlepoint system or via the Castlepoint RESTful interface, including information such as:

1. Type of system being monitored
2. Endpoints for the logging system(s) and/or log file locations
3. Security credentials to authenticate with the endpoints or log file locations
4. (optional) Security authorisation

Once a system has been registered administrators can test the log analysis connection and begin log processing.

### User Interface Integration

MIE provides a native interface with Microsoft SharePoint (on-premise or hosted), and all functionality for end-users and administrators is available from this interface. The user discovery interface is separate from the records and information management control centre. The Castlepoint MIE interface is implemented via a HTML5 interface and connects to the Castlepoint MIE microservices environment using RESTful service calls.

Our clients are able to enhance the user interface, build their own, or integrate with the RESTful microservices. All source code for the user interface is provided and licensed with the Castlepoint MIE system. The user interface can be re-branded, white-boxed and customised for specific user requirements.

## Licensing

Castlepoint provides value for money by

- low deployment costs – SharePoint is available at no additional license cost to agencies as part of the whole of government Microsoft agreement, and is already installed in many Departments. Castlepoint is a simple addition to the SharePoint platform which does not customise SharePoint or affect its support, sustainment, upgrade path or warranty
- low maintenance costs – our cost model includes three years' licensing and support
- low cost of integration and interoperability with dozens of systems – integration is 'loose' with no need to customise or modify bespoke or off the shelf line of business systems
- reduced user productivity cost for achieving records management compared to traditional EDRMS
- increased user efficiency for discovery – single pane of glass to discover most relevant dispersed content
- better return on investment of existing information assets due to their increased reusability, and
- minimal impact on other processes – no changes required to normal business workflows.

Castlepoint MIE is licensed under either a SaaS service consumption model or a private cloud/on-premise SharePoint farm model. In the SaaS consumption model, agencies pay for user licensing and for storage+compute usage billed monthly. In the private cloud/on-premise model, Castlepoint is licensed per SharePoint farm.

Surge events, where additional users and/or data are introduced to the system, use the SaaS consumption model for cloud hosted services whereas the private cloud/on-premise model incurs no additional costs (subject to infrastructure scale). MOG changes, where the data stays at the same location, have no additional license costs.

## OVERVIEW OF BENEFITS

### Cost Benefits

- Single customisable product that doesn't require expenditure on a BIS or EDRMS in order to connect to them
- Single license for an organisation – only one product required to manage all records in all systems
- Reduced productivity costs for records managers and staff through ease of discovery and management

### Technical Benefits

- Avoids the need for customisation or integration with bespoke or COTS products, affecting their supportability
- Avoids duplication of storage for millions of records by allowing them to remain in place in their source system
- Runs using a simple AI system that does not require a complex rules engine requiring ongoing support

### Staff Benefits

- Allows records managers to add more strategic value by freeing time from manual classification and control
- Assists users with information discovery across the boundaries of their usual systems and silos
- Allows users to work as normal in their usual BIS, keeping all records management functions invisible to them

MIE achieves the benefits of full compliance, including reduced organisational risk and increased business benefit, without the commensurate cost, technical and staff impacts that the current paradigms introduce. Using MIE, Australian government can easily achieve its DC 2020 obligations, and improve its overall information management.

## FOR MORE INFORMATION AND WHITEPAPER

<p><b>Rachael Greaves</b> <i>Chief Information Officer</i> Castlepoint Systems <a href="mailto:Rachaelg@castlepoint.systems">Rachaelg@castlepoint.systems</a> <a href="http://castlepoint.systems">http://castlepoint.systems</a> +614 88 114 767</p>	<p><b>Gavin McKay</b> <i>Chief Technology Officer</i> Castlepoint Systems <a href="mailto:Gavinm@castlepoint.systems">Gavinm@castlepoint.systems</a> <a href="http://castlepoint.systems">http://castlepoint.systems</a> +614 04 151 729</p>
---	--