# The Document Logistix guide to

# Document Management Integration

By Tim Cowell
Product Director
October 2011

## Why Integrate Document Management?

Electronic Document Management (EDM) systems are commonly deployed to enhance business processes by improving efficiency; usually based on reducing the time taken to perform a certain task.

In many cases the DM system is not the only business tool being used to achieve the task but complements existing systems by providing rapid access to supporting documents or by automation through the use of workflow.

An example of integration is 'document enabling' line-of-business applications (Accounts Payable, HR, Patient Records) to provide access to the documents relating to the transaction, capture of documents, approval of documents ( Invoices, Applications etc…) and updating of the integrated system.

In many cases the users continue to use the core business system and the DM system remains hidden from the users until it is called into play by the integration.

Typically an integrated DM system will provide far greater benefit to an organization than a standalone one, however, the benefits there can also be substantial. The result of integration is generally a more rapid Return on Investment.

## Types of Integration

Different people have different perceptions of what integration means. Simply asking 'Does your system integrate with xyz' is not sufficient to define the requirements, you should be sure about what functionality is needed from the integration and ensure that the system can deliver.

There are a number of levels of integration which can provide process improvements; however, it is important to understand the complexity and maintainability of the solution.

### *Validation of Data*

The simplest form of integration is to ensure that the information being used to index documents is valid and relates to the host system, for example the 'Supplier Name' matches one of the suppliers in the host system. This is usually achieved by defining some form of database validation between the field in the DM system and an equivalent field in the host system.

## *Automatic Indexing*

This is related to data validation and is used to reduce the amount of data entry needed, ensure consistent and accurate indexing and provide a wider range of searchable fields for retrieval at no extra cost.

Typically a key field is entered as a document index which is then validated against the host system, but in addition related fields are retrieved from the host and used to index the document. For example a 'Supplier Name' is entered and validated but additionally the supplier's code, postcode and phone number are retrieved and used to index the document.

The DM system can now locate the document by using these additional fields and no additional effort had been spent during the indexing process.

If the DM system is only going to be used to locate documents via the host system there is little benefit in doing this, however, where the DM system will be used directly by people without the host system the benefits are far greater.

## *Document Enabling*

Document Enabling means providing easy access from the host system to locate and view related documents from the DM system. This is usually done from a record view in the host system, for instance a supplier invoice, patient record, HR record, etc.

How this is achieved will depend largely on how 'open' the host system is; whether it has user configurable buttons or menus or whether it has an API and what type of application is it (Web, Rich client etc.).

There may be a one-to-one relationship between the host record and the document, for example an Invoice number will relate to an Invoice. There may be a one-to-many relationship where a host record relates to many documents, for example a list of documents relating to a job, or a patient record.

Different screens in the host system may require different levels of integration and you should consider which screens will require frequent document access.

## Assisted Document Capture

Documents sometimes require scanning and indexing in relation to a transaction in the host system, like a newly received invoice. In order to avoid manually entering the document indexes when scanning there are alternative ways to automate the process.

If the transaction has a unique reference this can be hand written onto the document and then used by the validation / auto indexing integration described earlier to semi-automate the indexing process when the document is scanned.

Better still, if a barcode label can be printed from the host transaction screen this can be affixed to the document and read automatically thereby removing the need for manual indexing. Another benefit of this method is that the barcode can also be used to separate documents with multiple pages so the scanning process is more efficient.

## Document Generation

It is surprising how many people print documents or emails and then scan them. Not only does this wastepaper and incur printing costs but is also hugely inefficient.

Ideally there should be a mechanism for saving documents generated by the host system into the DM system without the need to print and scan.

This can usually be achieved in its simplest form by using a print driver that creates an image (usually a PDF) of the document which is then saved to the DM system.

Where possible this can be automated so the users do not need to select how the printed document is indexed, but even if this is not possible and the document requires manually indexing this is still preferable to printing and scanning.

## Document Driven Actions and Updates

The DM system may also be used to automate some workflow process, for instance, an invoice approval, and during the process additional information about the document is captured.

The host system may also need to be updated during the process in order to reflect an event, like an Invoice being released.

In some cases the DM system is actually the start of a business process; for instance when a new document is created or an incoming document is scanned. This may

need to trigger an event or a transaction in the host system so it is important that the DM system has the capability to initiate this.

## *Methods of Integration*

There are different ways of achieving integration and some are easier than others. You will probably need to seek advice from the DM provider and possibly the host system provider to determine what is and is not possible.

### Database Lookup

If the host system uses an open 'industry standard' database this is usually a good indicator that validation and automatic indexing can be achieved relatively easily.

All that is usually required is some knowledge of the database schema and authentication details to the database.

For reasons of safety and integrity it is suggested the DM system be given a 'Read Only' view of the relevant tables in the host database.

Where the DM system needs to update the host system it is preferable to use the host system's API but if this is not available then it is important to ensure that the DM provider has the support and approval of the host system provider.

### Screen Scrape

Some applications that present information through their user interface in a structured field layout can have their screens 'read' by the DM system. The relevant fields are identified on the screen and these are then used to automate a search in the DM system. This technology is commonly referred to as Screen Scraping.

The benefit of screen scrape is that it provides quick and easy Document Enabling of the host system without software development. The DM System can be trained to recognize different screens in the host application, and if the screens change (due to a new version of the host software) it can be re-trained in a matter of minutes.

In order to activate screen scrape the user just needs to press a pre-configured hotkey, usually a combination of a control key and a function key, and the DM system will locate the documents relating to the information displayed, for example an order number or patient number.

## Barcode

Barcodes are an old technology widely used across many industries for rapid and reliable capture of data. Barcodes contain check digits which ensure that if a barcode is read the data is correct, they are very reliable and are ideal for indexing documents into Document Management systems.

If documents to be scanned relate to records in the host system printing of the barcode label should be automated. This may require some enhancement to the host system although some DM vendors, like Infonic, have barcode label printing software that uses the same 'screen scrape' technology discussed above to capture the required data and print the labels.

## Print Driver

One of the simplest ways of capturing documents generated by any system is to intercept the print output. Typically the DM system will install a printer driver and the user can then print their documents to the 'DM Printer'.

The DM printer will then generate a PDF of the document and file it in the DM system, either with the document indexes being entered by the user, or automatically generated depending on the level of integration achieved.

A good DM print driver will also then forward the original print onto a real printer so the original hard copy can be produced for posting if required.

## API

API is an acronym for Application Programming Interface and provides a means for a software product to expose it's functionality to other applications in a controlled and safe way.

In order to achieve a good level of integration both the DM system and the host system will need to provide an API and comprehensive technical support.

There are different types of API available and the choice will depend largely on the architecture of the product and the development language. Some

common APIs are:- .COM (also known as ActiveX), .Net and Web Services (typically XML/SOAP based).

The choice is less important as long as the API is available, documented and supported.

Ideally the DM system API should provide both a means of exposing its functionality to external applications, and be able to call external application code when events occur inside the DM system, for instance, new document creation or modification of a document index.

## Summary

The term 'integration' means different things to different people and there is a lot of room for ambiguity so it is important to be specific when asking for integration.

Compare the potential business benefits of an integrated solution against the cost of undertaking the integration, i.e. don't integrate just for the sake of it, and ensure that there is a quantifiable business benefit.

Check that your DM vendor can demonstrate experience of integration and that they have an open (and preferably free of charge) API.

All of the integration methods discussed in this document have been used by Document Logistix although it is rare to use all of them on a project. Our products include a comprehensive API that is widely used by Document Logistix, its partners and customers to provide efficient, integrated solutions.

For specific integration questions contact me tim.cowell@document-logistix.com