

Ashley Keil, VP sales, EMEA/APAC at IBML explores how systems integrators can improve data accuracy as part of complex outsourcing projects

Improving data accuracy



Ashley Keil

Systems integrators working on client projects involving complex data and information management operate in a world with challenging demands; meeting service level agreements, controlling costs, managing data which arrives from multiple sources and formats, of varied quality and in greater quantities than ever before.

The implications are significant. Inaccurate data is detrimental, bad data erodes operational efficiency, makes decision-making less accurate, adds commercial risk, delivers a poor customer experience and can be expensive too given GDPR rules and the associated penalties.

Data practices are good but there are gaps

Many systems integrators have made considerable headway automating data capture processes, investing significantly in best-of-breed intelligent capture technology which integrates easily into line of business systems because of the use of open APIs. This helps process the tsunami of information coming in whether it's extracted from postal mail, email or other sources.

Artificial Intelligence and machine learning platforms today perform complex data capture with minimal operator invention delivering accuracy rates of anywhere between 80 and 95 per cent. The variation comes when dealing with crumpled or torn paper, text where a highlighter pen has been used, or illegible handwriting on a form, for example. It's just more challenging for the recognition engines to process so that data can be ingested into downstream business systems.

To boost accuracy, barcode technology has been used with some success. But it is not a panacea and only works with a small percentage of data capture situations. Some organisations have employed staff to manually rekey information or to review capture results for each document to ensure accuracy. These approaches to eliminating data errors are costly, time consuming and far from foolproof.

So, what are the options for systems integrators if accuracy rates of 80 to 95 per cent aren't good enough? How can data capture be improved to get to 100 per cent without the expense of adding more headcount?

Achieving data perfection is possible

The answer lies in using a multifaceted approach. First, utilise a scanner which comes with real-time intelligence that understand documents, extracts data early in the process so as to minimise errors downstream.

Second, set business rules to capture and validate field-level metadata. So, for example, the scanner can review whether an application form has a signature and whether the right numbers of pages are scanned. Remedial action can be taken if they don't.

Third, AI-driven matching solutions are available – integrating with the scanner or independent of it – to enable the cross referencing and matching of multiple incomplete or incorrect data fields against master database sources so that errors can be flagged and dealt with immediately.

This means that a number of partial metadata captures, which are inaccurate

in their own right, can be pieced together and combined to correct and validate the information being processed before it is accepted into a business system.

An example would be scanning a postal form where a system integrator is the 'prime' on an election project. Bits of the person's name, address, postcode or all three might be obscured. By assessing all the fields and the text and then cross referencing this extraction in a master database – which might hold millions of customer records – the AI solution can bring these partial 'reads' together to get a qualified and accurate result. Complex algorithms are used to do this, with it all taking just milliseconds.

Triple data entry

The fourth way to achieve clean data is to use a scalable automated crowdsourcing approach to do what's called triple data entry. This pretty much guarantees data accuracy.

Crowdsourcing pushes snippets of the same information to two online data entry clerks who are connected to a management platform via the internet. They check the same snippets of unmatched or poor quality data from an image before entering it into a system. If there's a mismatch between what the two individuals input, it goes to a third person for exception handling which solves the issue of manual errors creeping in. This is how 100 per cent accuracy rates are achieved.

Crowdsourcing data checking is ideal where intelligent word or character recognition technologies have struggled to recognise handwriting in a field and more validation is required. Working with a specialist crowd partner is a fraction of the cost of employing staff.

Data's exponential growth has created opportunities to leverage it in new ways for better business outcomes. Accuracy is therefore key for systems integrators tasked with delivering complicated projects. Triple data entry is a relatively new concept but the approach is cost-effective, fast and secure.

www.ibml.com