



EXCLUSIVE
CHECKLIST
REPORT



Getting Smart About Compliance Automation

The Gift of Compliance Automation

By Marydee Ojala, Conference Program Director, Information Today, Inc.

Documents, documents, documents. The business world is full of them—and there’s no indication there will be fewer in the future. In fact, the trend line is going in exactly the opposite direction. Not that long ago, when people talked about “documents,” it was clear that they meant written materials, usually on paper. In today’s world, the definition has expanded considerably. A “document” could resemble its paper antecedents, as a memo, a contract, or a report, but in electronic form. Storing documents electronically is vastly more efficient than sticking paper into file cabinets.

However, the nature of documents takes on additional meaning when you consider the types of communication we now enjoy. Documents could be emails, transcriptions, drawings, photographs, scribbled notes, audio files, and more. Each has the capability of being digitized and stored electronically, and each presents its own set of challenges. Plus, multiple formats can exist in one document. A report, for example, could include an embedded video and hyperlinks to external sites.

The explosion in documentation has several repercussions. One is the sheer volume of available information, both within a company and external to it. Growth has been driven not only by the expanded definition of what a document is but also by advances in computing technology. Storage capabilities are verging on the unlimited. Now that most employees work from their devices, the opportunities for storing information in multiple places and the ensuing silos of data have exploded. Add regulatory oversight to the complexity of document automation and you have the potential for disaster if documents, in all their many forms, are not handled well.

HANDWRITING RECOGNITION

One of the more exciting trends in document automation is handwriting recognition. Although word processing may have supplanted handwritten documents for many purposes, handwritten notes and correspondence have not disappeared. Neither have handwritten signatures. Some school districts are phasing out instruction in cursive writing, but there remains a large body of documentation that shows up in handwritten form.

Early attempts at simple handwriting recognition, such as that used by the Post Office to understand an address written on an envelope, revolved around optical character recognition (OCR). Although that worked reasonably well for some purposes, it had a number of shortcomings when applied to longer documents, even when they were printed and not handwritten. You need only look at early digitization efforts of historical newspapers to see errors introduced by the OCR process.

Advances in machine learning, coupled with deep learning algorithms, have contributed greatly to the ability of computers to understand not only printed but also cursive handwriting that

appears in a document or is entered into a legal or medical form. A significant amount of research into character and feature extraction involving neural networks and deep learning has resulted in broad applications of handwriting recognition with astonishingly improved accuracy, particularly in light of the early OCR-based attempts. For businesses, particularly in industries that rely heavily on multi-line forms, being able to decipher handwriting streamlines many processes.

PRESENTING DOCUMENT AUTOMATION

The changing nature of document automation can be seen as a gift. We have more types of information available and accessible. We can view it as a glorified child’s birthday party, with a myriad of brightly wrapped packages surrounding the happy child. But the other side of the story is disappointment. The box marked with the branding of a coveted toy has been repurposed by the giver to contain a sweater. Our expectations are high, but disappointment can be managed.

Dealing with disappointment is a fact of adult life. The promises of technology must be tempered with the knowledge that just because something is technologically feasible doesn’t mean it should be done. The reality of balancing document automation with regulatory compliance pervades much of business life.

Let’s just suppose you are assigned to the compliance area of your company. Let’s further assume you didn’t ask for the job and you don’t think it will further your career. You’re disappointed. Let me introduce you to Greg Council, VP of marketing and product management at Parascript, who will prove you wrong for taking a negative view of compliance.

BENEFITS OF COMPLIANCE

Council is a realist. He’s happy to admit that all those pesky rules and regulations related to compliance are frequently regarded a “necessary evil,” even an impediment to running a profitable business. I’m pretty sure you make yourself unpopular when you point out that a new marketing initiative, public relations idea, or IT project will run afoul of current laws. You’re greeted with “But we’ve done this before and no one objected” or “I don’t see why we can’t store this particular piece of data since we need it to understand the customer better” or “Those regulations only apply if we’re in [insert geographical entity here] and we’re not” or “Complying with that rule just isn’t technically feasible for us.”

Council suggests that, instead of viewing compliance as a burden, we consider the possibility that maximizing our ability to access and use information becomes an asset to the company. It’s not a thankless task; it’s a benefit. The notion of transforming compliance from a negative to a positive rests on automation. He asks you to imagine the automation of compliance as something

“The commonalities of the mortgage industry and healthcare, when it comes to document automation, are striking. ... The documents originate in paper, sometimes mailed or even faxed. Handling the documents is labor-intensive and, thus, a costly endeavor.”

that will get you recognized as a positive influence on the bottom line—which reminds me of the old adage about making lemonade when life gives you lemons.

It's typical that compliance, despite its complexity, is often given few resources to accomplish its mission. Yet the process of document automation can uncover important data currently lost in manual processes and, probably, in multiple silos. Think of it as that present you bought 6 months before a child's birthday that you carefully wrapped and put in the back of a closet—where you promptly forgot about it until well after the child's actual birthday. So you spring it on them as an “unbirthday” present. The valuable information unearthed during an automation project provides the same unexpected excitement. Hidden information can provide insights to guide the company towards a better future.

FINANCIAL SERVICES AND HEALTHCARE INDUSTRIES

Two of the most regulated industries are financial services and healthcare. Council uses these as specific examples of the value of automating documentation. As I'm sure many of us know from experience, obtaining a mortgage loan requires so many signatures on so many individual pieces of documentation that your hand could cramp up. That might send you into the healthcare system to have the hand pain alleviated, where you are confronted with another myriad of forms. All that data, often collected on paper, needs to be transferred to electronic formats.

Although a borrower might think signing multitudinous documents is the finale of the transaction, from the mortgage lender's perspective, the paper trail doesn't end with signatures at closing. Instead, the paperwork becomes a critical component of servicing loans. Underneath the sale of service rights is “the need to ensure that all key data from a loan that impacts the servicing process is collected and verified.”

Since document requirements are not standardized, automation streamlines the transaction. It starts with the intake of the loan file, where it is imported and analyzed. Next comes verification of loan data. Natural language processing (NLP) plays a big role in locating and extracting key information.

In healthcare, Council points to HEDIS (Healthcare Effectiveness Data and Information Set) Audits as a prime area where document automation of patient medical charts not only ensures compliance, but lowers its cost. If audits are done using manual means, it takes up the time of healthcare staff, primarily nurses, as a significant cost to the organization. Automation reduces the time spent by individuals, concomitantly reducing the amount of money spent and releasing medical staff back to caring for patients. Council was quick to tell me that Parascript does not see individual patient charts. That data stays in the custody of the client.

AUTOMATION PROCESSES

The commonalities of the mortgage industry and healthcare, when it comes to document automation, are striking. Both need data captured, evaluated, and verified. The documents originate in paper, sometimes mailed or even faxed. Handling the documents is labor-intensive and, thus, a costly endeavor. Although the content of the documents differs, the process of handling them is surprisingly similar, as is the need for security and compliance.

Parascript's approach begins with sorting the deluge of records, parsing it at the document level. Each industry has key phrases

relevant to the process. For mortgages, it might be “income”; for healthcare, “symptoms.” NLP, along with deep learning and neural networks, simplifies identifying relevant information within documents and relating it to other pages and records. This is vastly more efficient than relying on people to spot the patterns. By automating tedious work, Parascript cuts the time required to package mortgage loan documents for the secondary market and to review patient information for audit purposes.

And just in case you're thinking, “Well, that's fine for those industries, but it won't work in mine,” Council would beg to disagree. It is true that compliance is more front and center for regulated industries, yet every business today must come to grips with the increasing number of privacy laws being enacted in many jurisdictions, from GDPR (General Data Protection Regulation) to CCPA (California Consumer Privacy Act). Employing compliance automation to your customer records and making access to information more efficient can not only avoid compliance disasters, but also win acclaim from the higher-ups in your organization.

Compliance is an area where, all too frequently, recognition only comes when you've done it badly. Turning this around so that your compliance efforts gain praise is the goal. Similarly, convincing others in your organization that compliance is worthy of their attention is more effective if you talk about monetizing it. Gaining buy-in to compliance happens when you stress the positive implications rather than threats about the negative results of non-compliance. It's simply more productive to win people over by convincing them that doing more with less is actually desirable and can improve the revenue picture. The process of automating data verification and cross-validating it with data found in other documents or records ensures speedy access to information that aligns with overall business goals.

Keeping your company out of hot water doesn't seem to have the same panache as generating hot profits, but it does have its moments. Your compliance job? It's a gift, a well-wrapped present. Trends in compliance automation revolve around innovations in machine learning, NLP, deep learning, and associated AI technologies. These open up enormous opportunities to understand documents that already exist and those being created. They will grasp the meaning that might currently be hidden within non-textual formats and unstructured data.

Technology, however, is only one trend to watch. Increasing interest by governments at all levels about privacy will affect how technologies are applied and how automating document-based tasks are accomplished. On the bright side, these trends can lead to improvements within companies when they discover hidden and unappreciated information that can provide important insights into their customers and products. ■



Marydee Ojala is conference program director for Information Today, Inc. She works on conferences such as Enterprise Search & Discovery, which is co-located with KMWorld, and Data Summit University, among others. She is a frequent speaker at U.S. and international information professional events. In addition, she moderates the popular KMWorld webinar series. Ojala is based in Indianapolis, Indiana and can be reached at marydee@infotoday.com.

Getting Smart About Compliance Automation

By Greg Council, Vice President of Marketing and Product Management, Parascript



When it comes to automation, most organizations are rightfully focused on top-of-line processes that aid with improving revenue through customer acquisition and retention or through the ability to enter new markets. Also, redesigning processes are usually on the “automation docket” so that they are less costly to support—in essence, to do more with less.

STREAMLINING COMPLIANCE-ORIENTED PROCESSES

Increasingly regulated organizations such as those involved with lending or in healthcare are turning to automation to streamline compliance-oriented processes. Traditionally, organizations viewed compulsory requirements associated with regulations or internal policies as a “necessary evil” to be dealt with in the least costly and disruptive manner possible; a case of “good enough” is fine. Often “good enough” means a lot of manual effort. This is largely due to the complexity of the processes and the information contained within them.

Buried within many compliance-oriented requirements lies the opportunity to improve business outcomes. It is more than just complying with a regulation. The underlying enabling benefit common to these processes lies in the ability to access and use key information, which without automation would remain largely untapped. This is where document automation comes into play.

TURNING LENDING COMPLIANCE INTO A REVENUE ACCELERATOR

Take for instance lending within the mortgage industry. While most efforts involving automation revolve around attempts to streamline and improve the customer experience with loan origination, there are a number of post-funding processes that can be improved, not only to better support compliance, but to improve overall business.

For example, service rights acquisition involves the transition of the business of servicing a loan over its lifetime from the originator/lender to a third party. The process of the sale of servicing rights requires both the buyer and the seller to work together to quickly share required information and the procedures involved must adhere to both internal compliance and regulatory requirements. The quicker the process, the faster a seller can recognize revenue and the faster the new servicer can benefit from service fees. Underlying all of this is the need to ensure that all key data from a loan that impacts the servicing process is collected and verified.

Each services acquirer has a different set of documentation requirements—including the order in which the documents are to be arranged, which is called the “stacking order.” The seller must quickly and efficiently comply with each seller, which on the

surface may appear simple, but in reality is often cumbersome and error-prone. The buyer, on receiving the documentation, verifies the receipt of all required information and then verifies the loan summary data against the documentation. Even though documents are provided in a specific order to support quickly locating each one, the process performed manually is still slow. The process of verifying specific data within each document is slower yet.

This is where automation comes into play. Even in an ideal situation where the seller has provided a remarks page along with the loan file, automation can make the process quick, efficient and controllable.

“Traditionally, organizations viewed compulsory requirements associated with regulations or internal policies as a ‘necessary evil’ to be dealt with in the least costly and disruptive manner possible; a case of ‘good enough’ is fine. Often ‘good enough’ means a lot of manual effort. This is largely due to the complexity of the processes and the information contained within them.”

The first step is the intake of the loan file. Typically, documentation is provided as a comprehensive sorted PDF file. This file is imported and analyzed to determine that the file is complete and adheres to the sequence defined in the stacking order. To automatically identify the presence of the required documentation, the software is trained on examples of these documents. The software “learns” what attributes of each document provide a reliable way to identify and classify them.

Whether attributes consist of the presence of certain text, the presence of titles, visual information such as logos, or a combination (usually the case), the software identifies patterns for each document very quickly and often identifies attributes that are hard to identify with a visual inspection. A similar learning process identifies how to determine the beginning and end of each document, which is helpful to support locating specific documents. An added plus is the ability to produce a derivative file that includes bookmarks for each document to allow a reviewer to quickly move from one document to the next.

The second step involves verification of data within the file. Since documents have already been identified and separated, the process of applying rules to locate specific information is made easier. Depending upon the nature of the document, different

techniques for data location are applied. Keywords or labels can be used to locate specific data while grammatical parsing using natural language processing (NLP) techniques can locate needed data in documents such as a remarks page.

In the ideal setting that a remarks page is provided, NLP parsing can locate specific information within paragraphs of text and import it into processes designed to review the data based upon data located and extracted from the accompanying documentation. For instance, employment information can be located and extracted from the remarks page and cross-validated with the corresponding income documentation, allowing for an automated calculation to verify income. The same cross-verification can occur for other key loan data.

While any such automation applied to document-based information will not be 100% accurate, use of automation as an “assistive technology” can significantly improve the quality and timeliness of reviews in support of quickly onboarding a new loan. In this case, compliance automation also accelerates revenue.

USING HEDIS AUDIT AUTOMATION TO IMPROVE OUTCOMES IN HEALTHCARE

Within healthcare, a certain process called the Healthcare Effectiveness Data and Information Set Audit, or HEDIS audit, is used by regulators, providers, insurers and other stakeholders as a benchmark for healthcare quality and effectiveness. Most healthcare plans use HEDIS to consistently evaluate the performance of their network of providers, but the process is complex, involving a lot of service-related data in a variety of formats.

Ideally, all required data would be collected and transmitted as structured data within an electronic health record and the industry is making strides to do just that. Still, the lack of a standard way to record and transmit all the necessary data leads to a costly and slow process to collect, sort and review data to support an audit. In all, the cost to report this data has been estimated to be over \$15B annually. The costs associated with review are just as large or larger.

Based on a review of specific patient cases, insurers make information requests from each of their network providers. While more measurements rely on structured claims data, a good amount of information is supplied according to the “hybrid method” that consists of patient medical charts.

Charts are submitted according to a health plan request and the review begins. Review involves skilled staff, often nurses, trained to identify key patient encounter data to determine appropriateness of the services rendered. The process can be so taxing that an industry has been built around servicing the requirements for both the providers and the health plans.

While review of data is still best accomplished by subject matter experts, there is a lot that can be done to automate the supporting processes, which can reduce both the time required and the associated complexity.

WHERE AUTOMATION BEGINS

Automation starts with records sorting. The submissions of medical information from providers often come through the mail or fax machine as a large multi-page file of medical records.

Each record represents a specific encounter where medical services were provided. Ideally, each individual record is identified and separated to evaluate independently, otherwise expensive professionals must spend extra time performing this work. Traditionally, this process has been labor-intensive as there has been no reliable way to distinguish one record from another. Each document is different and does not always contain the same data. Additionally, two service records can be on the same page or begin and end on different pages. Therefore, attempts at constructing a rules-based approach to automate this process often fails. Another technique needs to be used.

PROCESSING TECHNIQUES

Using natural language techniques along with deep learning neural networks, key information such as dates can be analyzed along with other text to understand which dates are likely to be the service date. This data needs to be parsed at the document level to get an understanding of how one page relates to the next and if a single page is part of two different records. Other textual information is further analyzed to locate the optimal “split point” for each logical record.

Once individual records are created, further automation is used to analyze each record, looking for key phrases that indicate symptoms, medical conditions and treatment recommendations including referrals to specialists. This data can be located and highlighted for reviewers such that the process can remove most of the tedious work associated with stepping through individual records and scanning text looking for specific information.

OVERALL WORKFLOW

The overall workflow resembles one where the reviewer is presented and shepherded through the information where they can easily skip through each record and note specific, pre-identified data in order to cut significant time from a process that can take hours for each patient audit. The additional benefit is that processing and reviewing actual patient medical charts offers a much more comprehensive and “richer” data set than can be achieved from a structured claims database. This review can also identify inconsistencies with the structured claims data leading to improved patient data. Both of which support improvement in overall patient care.

COMPLIANCE FOR BUSINESS IMPROVEMENT

When applying automation, processes used to support compliance aren't always a single-minded effort; these same processes can also improve business functions in a variety of ways. Utilizing document-based automation for the purposes of quickly identifying documents and making typically unstructured data very accessible for current business systems often means there are many additional benefits other than just meeting compliance.

Generally, through more efficient access to information, processes can implement automated verification of data, cross-validate data in one document with that in another document or system of record, and quickly identify related data. Immediate access to high-quality data not only ensures that compliance-related processes work more efficiently but also provides broader access so that key decisions are made with the most accurate, richest level of data. ■