

An aerial photograph showing the wing of a large commercial airplane on the left side of the frame, extending towards the center. The wing is white with a red-tipped wingtip. Below the wing, the ocean is filled with numerous cargo ships of various sizes, scattered across the water. The sky is a mix of blue and white clouds. The right side of the image has a white grid pattern overlaid on it. The overall color palette is dominated by blues, greens, and reds.

# A Journey in Data

How data is key across the supply chain

Manufacturing, retail and the supply chain's complex web of relationships and logistics is kept together, in no small measure, by documentation. Invoices, contracts, product specifications, quality agreements, bills of materials, bills of lading, certificates of analysis, safety data sheets, customs declarations, delivery notes... the list is long. And yet, though they are essential, documents almost invariably feel like a distraction from the main focus – the making, shipping and selling of the things that people need and want. Automating document management using an Intelligent Document Processing (IDP) solution doesn't just save time and cut costs, it frees your best people for supplier negotiations, customer service and business growth – and that's just the beginning.

# Minimise Paper, Maximise Data.

**Improve your forecasts,  
lower your risks, make better decisions**

## Controlling input costs

In a time of economic challenges, having your data at your fingertips helps you analyse your costs and processes more effectively and make better decisions.

---

## Operational Efficiency

Ready access to your data makes you more competitive and helps you keep to deadlines in a just-in-time economy.

---

## Maximising staff effectiveness

In a tight labour market automation allows time spent on repetitive, low-value tasks to be refocused on business growth and value-adding tasks including procurement, negotiation, sales and engagement with customers and suppliers.

---

## Customer Experience

Sharper command of data means faster, more useful responses and better customer service.

# Data: From Factory Gate to Shop Floor

For businesses across the supply chain, command of data is vital to control costs, organise processes, meet deadlines and build good relationships.

In many cases the quality and accessibility of the data determines just how good the decisions taken by manufacturing, retail and supply chain enterprises are.

In sectors where there are few data standards and lots of different formats, reading, interpreting and structuring data is challenging. To be useful your data must be clean and organised.

## For manufacturers:

- Ingest data from: invoices, bills of materials, certificates of analysis, mechanical artwork, safety data sheets, parts inventories, delivery notes, pro-formas, agreements, contracts and more
- Compare quotes and proposals from multiple suppliers, in a variety of formats
- Compare specifications from multiple suppliers, in a variety of formats

## For logistics enterprises:

- Process key data across a range of documents including: invoices, delivery orders, dock receipts, bills of lading, waybills (sea & air), shipping guarantees, insurance policies, packing lists, contracts, customs forms and more
- Consolidate documentation for each shipment for faster processing
- Track and compare input costs

## For retailers:

- Ensure data from a range of documents is entered into systems including: invoices, delivery notes, return notes, contracts, resumés etc
- Consolidate documentation for each shipment for faster processing
- Track and compare supplier costs

# Certificate Of Analysis (CoA) – Data Extraction

## The Customer

One of the world's largest manufacturers of fast moving consumer goods with a wide portfolio of iconic, global brands.

## The Challenge

The client handles thousands of chemical compounds from scores of suppliers. Each compound comes with a certificate of analysis that sets out the essential details of that product.

The client's data analytics team needs access to the information they contain. That means transferring the data from CoAs to the company's Laboratory Information Management System (LIMS) – to then compare products and make informed buying and manufacturing decisions on grounds of cost and quality.

Transcribing the data from these thousands of documents was done manually. Seventy-five certificates typically took 10-12 hours to process. The work was tedious, error prone and diverted staff from other essential tasks.

Processing the certificates of analysis was a prime case for automation. However, it was also beyond the capabilities of traditional OCR solutions. The certificates' formats vary widely. The quality is often poor. The data is generally laid out in tables. Last-generation solutions cannot cope with such constraints and even many current IDP systems would still need to be trained for each of hundreds of template variations.

## The Solution

For the initial phase of the project AntWorks set out to extract 33 data fields from 50,000 documents per year – around 200,000 pages. For the initial set up, the system was trained on a representative sample set to capture the majority of format variations. The data output was to be in an agreed Excel template for easy ingestion into the LIMS.

The project was scheduled to take 15 weeks from the start of the configuration process through to go-live.

## Results

By the end of the project CMR+ had achieved 82% accuracy. Further accuracy improvements will come with deep learning and further training.

There was an 80% reduction in processing time. Across 50,000 documents per year this is equivalent to a saving of around 840 days.

100%

Data Enriched

82%+

Accuracy pre human in the loop

80%

Reduction in processing time

# Solving Challenges in Manufacturing with CMR+



## Safety Data Sheets

Safety Data Sheets contain important information about the use, storage and handling of risky materials that are an essential element of many manufacturing and supply chain. CMR+ can ensure that their vital data can be extracted quickly and more accurately so it's readily available.

**Watch the demo video.**



## Certificates of Analysis

Certificates of Analysis are necessary for compliance in many areas of manufacturing and retail. CMR+ makes light work of documents like these so your colleagues can spend less time on low-value data entry and more on value-adding tasks.

**Watch the demo.**



## Bill of Materials

Depending on the sector, documents like formulas, recipes and lists of ingredients play an important role in manufacturing. But getting that data into systems where it's easy to access and analyse is slow, time-consuming and prone to errors. Automating the process with CMR+ frees your colleagues to grow your business or help your customers.

**Watch how CMR+ rips through manufacturing documents here.**

## For the retail, manufacturing and supply chain there are 5 significant benefits:

# 1

### Automate:

Inefficient, expensive processes that waste experienced colleagues' time

# 2

### Create capacity:

Perform business-critical, often highly labour-intensive processes every time

# 3

### Access Data:

A continuous push rather than an ad-hoc pull model

# 4

### Quality:

Enable compliance, auditability and better customer service

# 5

### Market advantage:

Better data gives a competitive advantage, better customer experience and opens up potential new revenue streams

## Why OCR Systems Won't Work

- They can't handle unstructured data or variations in how data and documents are set out
- OCR offers poor extraction capabilities
- Hard to read the data with the human eye
- Difficult to operate in multiple-languages
- Past deployments of neural networks needed a lot of time and money and offered no guarantee of success. These solutions tended to be very focused on individual lines of business or discrete functions and, so, couldn't be scaled across the enterprise. This pushed up the costs of purchase, maintenance and support

## Not All IDP Systems are Created Equal

- Accuracy scores are misleading – many providers have teams correcting the data without understanding the context
- Humans in the loop are great - but if they're not your humans they may misunderstand the data leading to errors
- To be really effective, IDP systems need to include the client's business rules – the client needs to be able to drive system development but most IDP providers offer a black box
- Black boxes push up the costs of purchase, maintenance and support

# How Cognitive Machine Reading Lets You Master Data Digitisation

If you want to uncover the vital insights buried in your documents and organise them so you can use them when you need them, you need to be able to:



Improve image quality



Data and document classification



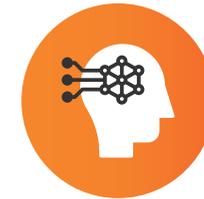
Natural Language Modelling



Creation and management of 'knowledge bases'



Ingest different formats  
e.g. pdf, Jpeg/PNG, Word, email



Learn continuously like a human



Extraction of unstructured data, paragraphs, and clauses, plus application of business rules to get the data exactly as needed



Compliant management of data



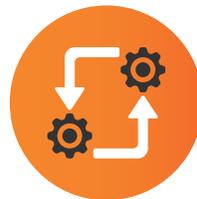
Operate in multiple-languages and interpret 'market language'



Have data export availability in multiple formats that integrate easily with downstream systems



Use supervised and unsupervised Machine Learning



Produce consistent, compliant outcomes

Using these techniques lets you extract valuable data from, for example, invoices, contracts, delivery notes, bills of materials, customs forms and more.

# CMR+ Tools

Humans create documents to make sense to other humans – not to computers. So, if we want computers to read documents we have to train them to ‘read and understand’ documents like humans.

CMR stands for ‘Cognitive Machine Reading’ because this is what we designed CMR+ to do – to read and understand documents in a way emulates the workings of the human brain.

CMR+ is powered by more than 300 proprietary algorithms that are built into our data hunting tools. Here are some of them.

## Label

CMR+ can look for a field label (or its variations) and then pair it with an adjacent value. You can set the minimum and maximum number of characters of the associated value or tell it to look for an object, such as a signature, a checkbox or other values.

Example: Tell CMR+ to look to the right of labels including ‘Invoice’ and ‘invoice number’ for a value of 5-15 characters.

## Match

The match tool excels at finding values that match specific formats for data like passport or bank numbers. Just train the system to recognise specific alphanumeric patterns and CMR+ will extract them from any document.

Example: Every country has a standard IBAN code format. CMR+ can be programmed with all the international IBAN configurations.

## Associate

Tell CMR+ to associate data in recognized formats, e.g. personal name, date, address and amount, with a nearby label of a separate value.

Example: Tell CMR+ to extract a date next to an invoice number as the invoice date even though the date isn’t labelled such.

## Derive

Derive mode takes one or more values and applies business rules to them to derive another value from them.

Example: Calculate taxes by applying a % to a sub-total or calculate a period of time by identifying the gap between two dates.

## Table

Tables are tricky for most IDP systems but CMR+'s table tool makes extracting their data easy whether or not there's a grid or if the data is single or multi-line.

## LookUp

CMR+ can be told to look for a value, like a supplier or product name, and fill in all the associated data from a linked database.

Example: CMR+ finds the company name Cranbrook Engineering on an invoice. It looks up the name in a master database of suppliers and if it find Cranbrook engineering can add a supplier code or address and contact details to the record.

## Infer

This tool learns from constant association and training and can be taught to infer meaning from context. Groupings of words can help CMR+ identify relevant data.

Example: CMR+ could search an insurance policy for a grouping of words including earthquake, tremor and cover are mentioned and then look for a check box or the word yes to confirm if the policy holder has earthquake cover.

## Zone

Zone allows you to specify that CMR+ searches a data item in a particular portion of the page, for instance left or right.

## Handwriting

CMR+'s handwriting tool can be trained to recognise both block and cursive handwriting.

# Investing in Your Future Success

Now, more than ever, data is at the heart of innovation in business; innovation that will be crucial in determining who wins out in a crowded, complex and competitive market. Cognitive technologies like CMR+ are becoming the standard tools for processing the vast flows of information vital to manufacturers, retailers and organisations in the supply chain. But it's not so much about the volume of data as the quality. It's about identifying the most relevant information and using it properly. Success or failure will increasingly rest on reading, interpreting, understanding and using the right data in decision making. The more there is at stake the more vital that will be.

It's time to automate. Many of your most experienced, knowledgeable people in, people who could be improving your customer service and helping grow your business, are wasting time on manual data entry, something cognitive machine reading does better and faster.

The race for data is on. The race for customers is on. Is it one you're in the right shape to win?

@AntworksGlobal 

AntworksGlobal 

AntWorks 