Cognitive computing for Customs agencies: improving compliance and facilitation by enabling Customs officers to make better decisions

By Stewart Jeacocke,
GLOBAL CUSTOMS EXPERT, IBM,
and Norbert Kouwenhoven,
EU CUSTOMS LEADER, IBM EUROPEAN UNION TEAM

DATA ANALYTICS – for example, automated selectivity rules – has become an increasingly important tool for Customs agencies. However, conventional data analytics has some critical limitations. It can only look for pre-defined patterns and rules, and cannot make use of unstructured data, which comes in the form of emails, social media, blogs, documents, images and videos. Cognitive computing allows Customs agencies to extract insights from both structured and unstructured data, discover new patterns and rules, capture the experience of top performers, and improve the quality and consistency of decision-making.

The dawn of the 'cognitive' era

On 11 May 1997, the Deep Blue computer system beat Grandmaster and world chess champion, Garri Kasparov, after a six-game chess match. This represented a major milestone in the evolution of computer systems. The Deep Blue project inspired a more recent grand challenge: building a computer that could beat the champions at a more complicated game – the American game show Jeopardy!

Over three nights in February 2011, this computer system – named Watson – took on two of the all-time most successful human players of the game, and beat them in front of millions of television viewers. The technology in Watson was a substantial step forward from Deep Blue. By understanding and processing a massive supply of unstructured data, Watson demonstrated that a whole new generation of human-machine interaction is possible. Watson had ushered in a new era of computing: the 'cognitive' era. Since then many large technology firms, such as IBM, Google, Facebook and Apple, have all developed cognitive computing technologies.

Cognitive systems have three characteristics that distinguish them from programmable era systems. They:

- understand unstructured data, through sensing and interaction;
- reason by generating hypotheses, considering arguments, and making recommendations;
- learn from training by experts, from every interaction, and from continually ingesting data.

It is estimated that 80% of all data generated today is unstructured. The dawn of the cognitive era means that machines can, for the first time, create insights from this enormous volume of data. Unstructured data is data which is not organized in a pre-defined manner; it is typically text heavy and often consists of documents, reports and articles, but can also include images, social media and videos. This previously unintelligible pile of documents is now a rich source of data for cognitive systems which can process it faster and more accurately than humans ever could. Processing the entire content of Wikipedia is a piece of cake for a computer!

Early successes in oncology

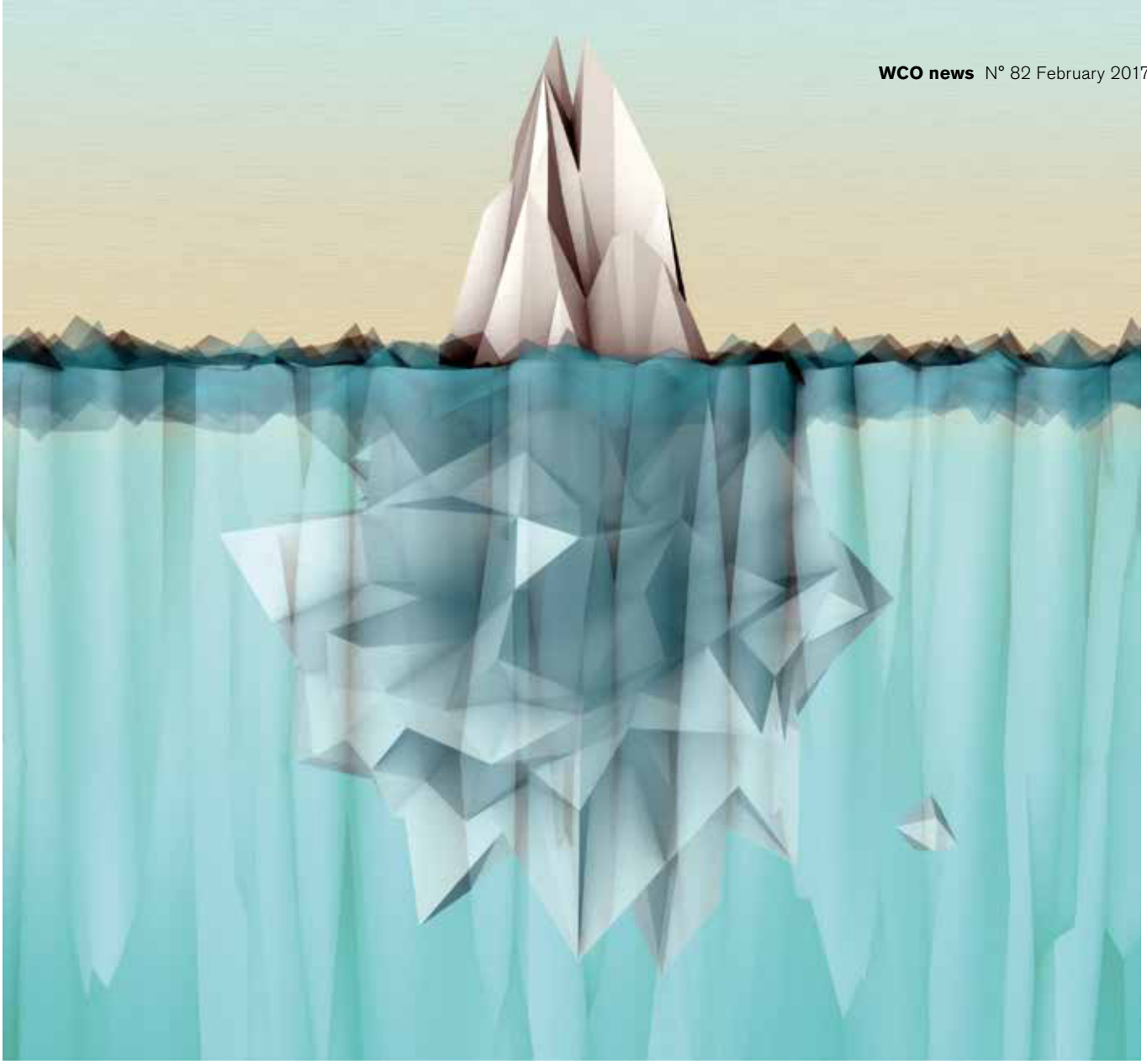
The first successful commercial use of these cognitive software capabilities can be found in the medical world, in the field of oncology. Watson is being used to process unstructured medical and research notes, as well as test results. These notes include descriptions of symptoms, patient characteristics, chosen therapy, and the corresponding therapy results. Watson is able to process this vast stream of information, generating correlations, patterns and hypotheses; this along with training by expert clinicians is used to create a corpus of knowledge.

Using this corpus of knowledge, Watson helps physicians identify personalized therapy options for patients. Given a patient's symptoms and history, Watson will suggest various therapy options, and a confidence score for each option. It allows the physician to drill down into the documents used to generate each hypothetical therapy option so that they can understand the evidence on which Watson is basing its recommendation. The more relevant information Watson finds to support a particular therapy, the higher the confidence score. Where additional test results would assist in improving Watson's recommendations for the patient, it will suggest these to the physician.

The Watson solution allows local physicians and their patients to benefit from the hundreds of thousands of pages of medical research and results – a knowledge base that is too large and evolving too fast for any individual to be able to read everything in order to stay current. Watson provides local physicians with advice on treatment options, but the treating doctor is still in control of making the final treatment decision along with the patient.

Cognitive systems for Customs

Informed compliance, risk management and audits continue to be some of the most powerful techniques used by Customs agencies to deliver their mission. These techniques allow Customs agencies to assist the vast majority of economic operators who wish to be compliant, and enables them to concentrate their resources on controlling the smaller number who represent the greatest risk to a country. Data analytics – for example, automated selectivity rules – has become increasingly important in enabling these techniques. However, conventional data analytics has some critical limitations. It can only look for pre-defined patterns and rules, and cannot make use of unstructured data, causing it to ignore 80% of all available data.



© Philippe Put

The amount of data available to Customs agencies is exploding, and much of it is unstructured. Initiatives aimed at creating a global platform of information, such as the WCO's Customs Enforcement Network (CEN), are increasing the amount of information shared between agencies. The ease of publishing information on the Internet has dramatically increased the number of relevant open source intelligence sources. Supply chains are becoming more instrumented and interconnected, and economic operators are sharing more of this information with Customs.

Cognitive systems allow Customs agencies to extract insights from this mass of big data, discover new patterns and rules, capture the experience of top performers, and improve the quality and consistency of decision-making across their organizations. In doing so, they improve an agency's ability to engage, decide and discover:

- Engage – Informed compliance can be improved with a virtual assistant, which engages in a conversation with travellers or infrequent importers; guiding them through the import process, or highlighting relevant preferential schemes. Recently, Dubai launched Saad, a virtual assistant which guides individuals through the process of opening a business. During inspections, Customs officers could use an iPad to take a photo of a foreign language document, and see it instantaneously translated into their native language on screen. Or, they could take a photo of an unknown item and automatically search for similar images in their agency's knowledge base, assisting them to determine the correct course of action. Targeting officers could query a single system to access all intelligence and information relating to a consignment, allowing them to determine quickly whether the consignment should be held for inspection;

- Discover – Cognitive systems can be used to 'mine' Customs data, enabling previously unknown patterns, relationships and indicators of fraud or other risks to be discovered. This data is no longer restricted to just that which has been captured electronically in information technology (IT) systems as part of the Customs declaration process. Commercial documents along with open-source intelligence sources, such as news sites, organizations' websites, social media, company filings, container tracking and vessels' automatic identification system (AIS) broadcasts, can all be processed, merged with classified government records, and used to derive insight;

- Decide – Conventional selection/ targeting systems can be enhanced with cognitive technology so that they also process the free-text information in economic operators' filings, and use it to improve targeting accuracy. Goods descriptions can be automatically compared to the declared tariff classification, in order to identify cases of misclassification; a technique that has already been adopted by the Government of the Bahamas and other Customs agencies. In a world that is not too far away, the images from scans of trucks and containers could be automatically compared to economic operators' filings as a means of highlighting those shipments where the scan does not appear to match the goods description.

Cognitive targeting

Some countries have already implemented cognitive solutions to improve their targeting/selectivity processes. Canada Border Services Agency (CBSA) enhanced its targeting rules so that it could apply cognitive processing to free-text goods description fields. By doing this, they were able to improve the accuracy of the targeting rules compared to their previous keyword search approach.

A simple example of the improvements is that the cognitive solution understands that a goods description including the words "no wood crates" should not match against a profile for controlling wood packaging. Previous keyword matching techniques would have matched against the wood packaging profile because the text included the word "wood." The cognitive solution understands that the "no" indicates a negation. This reduction in false positive matches translates into reduced inspection efforts, benefitting both the CBSA and economic operators.

Cognitive advisor for intelligence and investigations

Several countries are experimenting with using cognitive solutions to assist with intelligence and investigations. A common solution is an intelligence advisor which allows investigators to analyse structured and unstructured data sources by posing queries in natural language. The investigator logs into the system, and types in a query in natural language: for example, “What connections are there between football gangs and drug trafficking?” or “Which members of the Ultras football gang are connected to companies undertaking activities at the Port of Sevenberg?” The cognitive intelligence advisor responds with a series of hypotheses.

For the first question the responses of the intelligence advisor might include “Ultras”, “Baltic Nations” and other hypotheses. The investigator can drill into each hypothesis to understand the evidence behind it. This evidence might be document excerpts from classified intelligence databases, or open sources such as press articles or grey literature. With a small team and three weeks, one Customs agency was able to create a pilot system that processed over 20,000 documents from six different open source intelligence sources, such as Reuters, national news agencies, and company websites and press releases, allowing users to interact with it in the manner described above.

The simplicity of the interaction hides a significant amount of system intelligence. Cognitive systems work in a similar manner to the way humans undertake a task; they learn by being taught by another person. Older programmable systems, by comparison, must be programmed with specific rules. The system must gather data in a huge range of formats from a wide range of sources. It understands the language contained in the documents and users’ queries; for example, that “Ultras” is the name of an organization, and that the user is looking for relationships between that organization and the Port of Sevenberg.

A small group of intelligence officers train the system by querying it in the manner described above, and then indicating which of the responses brought back by the system are most relevant to the question being asked. This process trains the system by creating a corpus of knowledge. The training teaches the system which parts of the data that is has access to are most relevant to answering particular types of questions.

After sufficient training the system will start to answer queries in the way that an intelligence officer would, and will highlight connections that could take an officer years to discover manually. Furthermore, the system scales the knowledge of a small group of expert intelligence officers to all Customs officers who have access to the system, lifting the performance of all staff up to that of leading experts.

Customs agencies need a cognitive roadmap

Cognitive computing has arrived and is already being exploited by forward thinking Customs agencies. Organizations which successfully apply cognitive computing follow a structured, but agile roadmap which nurtures optimism whilst breaking down any complacency.

Cognitive computing helps Customs officers make better decisions in a wide range of different scenarios by understanding the vast amount of unstructured data that is inaccessible to conventional IT systems and reasoning, generating hypotheses and learning from expert officers. It allows all officers to benefit from the experience of an agency’s best experts, whether they are investigating complex cases of fraud, or determining whether to inspect a shipping container.

The cognitive era has arrived; every Customs agency should have a roadmap for adopting cognitive computing, and use it to usher in a new era of Customs facilitation and compliance.

More information

stewart.jeacocke@uk.ibm.com
Norbert.Kouwenhoven@nl.ibm.com