Why Relationships Matter

The Adoption of Machine Learning & Artificial Intelligence to Risk Management & Compliance

clicks®

Developed together with our friends and partners at **Deep Blue AI**

TABLE OF CONTENTS

INTRODUCING 6CLICKS

Hello – we are 6clicks!

The world is changing at an increasing pace thanks to digital transformation being spurred on by cloud computing, the internet of things (IoT), artificial intelligence and blockchain. This digital transformation is redefining every industry, from health, education and finance to agriculture, mining and transport.

The ability to witness and contribute to these changes makes for an exciting time to be alive.

We believe that for these new technologies to be successfully adopted and safely embedded into our way of life, we must address the issue of trustworthiness.

Organisations selling and operating these technologies must demonstrate that they have considered relevant threats, vulnerabilities, and risks. They must also demonstrate compliance to the applicable standards, laws and regulations. Each organisation must not only address these issues for itself, but across its entire supply chain.

This is why we founded **6clicks**.

We are on a mission to improve risk and compliance practices. By doing so, we can better connect people with technology, to ensure that this technology lives up to expectations both in terms of maximising benefits and minimising downside.

By using **6clicks**, you will shorten the time it takes to perform assessments and adjust program activities, thus enabling you to continually refine your approach, and learn what does and does not work for you. **6clicks** can help your organisation dramatically improve cyber and information security, as well as address other supply chain-related trust issues such as privacy, safety, environment and modern slavery.

Get started with a free trial account today, partner with service providers who now choose **6clicks**, or get in touch if you have any questions.

Anthony, Louis, Andrew and the **6clicks** Team

RISK AND COMPLIANCE CHALLENGES

The Overwhelming Weight of Compliance

We know that laws, regulations and standards, are being introduced and updated more frequently than ever, as regulators increase their scrutiny. If you operate in more than one jurisdiction, industry or are based in a more developed markets, it's inevitable there are a great many requirements of which your organisation must comply, and to which you must be able to demonstrate compliance.

Some of these modern world challenges include:

- Cyber security,
- Information security,
- Privacy,
- Anti-money laundering,
- Anti-corruption,
- Health and safety,
- Environmental issues (climate change), and
- Modern slavery.

The **6clicks** platform addresses all of these risk and compliance domains.

To provide context and an example, this white paper will focus on cyber and information security because of the wide-spread nature of the related risks and a general understanding of its importance across industries. That said, the research and development described in this white paper has a far broader application.

Organisations with the necessary people power to meet typical compliance obligations tend to do so through a process of mapping compliance requirements to their internal controls. This 'mapping' process may also indicate similarities across compliance requirements or the relationship between a compliance provision to something else like a question (and answers as part of assessments), or an internal control. The result is often referred to as a unified control set that maps to multiple sets of compliance requirements to individual controls.

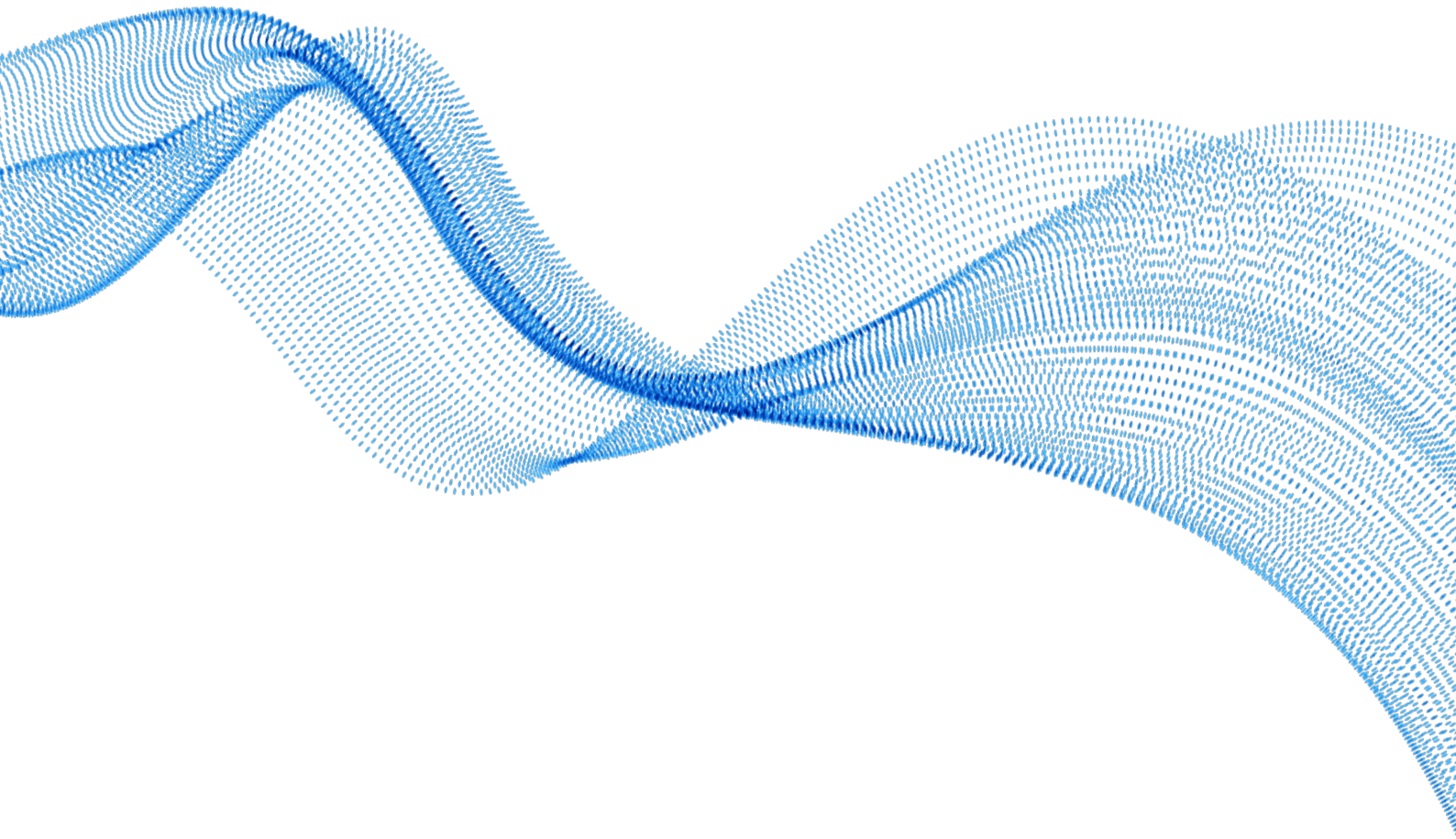
This mapping process is not only time-consuming to define but requires a high degree of familiarity and experience. It is also prone to error and almost impossible to maintain, as changes continually occur to laws, regulations and standards.

One approach you could take is to outsource this activity, but providers of these services are expensive as they are still largely dependent on humans and so still prone to error or bias. They can also be limited in coverage and tend not to be kept up to date with the pace of change.

If you have managed to get your hands on a mapping of your favourite compliance requirements, you can use it to streamline assessments and more effectively evidence compliance. Surely, asking the same question twice of the same person, over a short period of time, is not desirable. Though believe it or not, that is a very common practice. Compliance is important when tackling serious modern world challenges, but it needs to be efficient to be accepted as a routine business practice.

For example, where evidence of compliance to 'Requirement A' is needed for one standard, and 'Requirement A' is equivalent to 'Requirement B' in another standard, then you can use that same evidence to demonstrate compliance to 'Requirement B'.

Hmmm…but is all that effort worth it? It is expensive, requires an army of people and must be maintained manually today.



Risk Management Hypothesis or Hypocrisy?

Compliance requirements are becoming less prescriptive and more risk-based, which is good, but what effective risk management looks like can be unclear.

Companies typically use a combination of *top-down* (broad, general) and *bottom-up* (detail-oriented) approaches. *Bottom-up*, which is most common, is where risks are identified and managed in response to non-compliance, an incident, a threat, a vulnerability or a weakness. The *bottom-up* approach can be achieved simply and easily with the **6clicks** web application and is useful to identify and close gaps. However, when using the *bottom-up* approach it can sometimes be difficult to see the forest for the trees.

Alternatively, a *top-down* brainstorming or consultative style activity can help ensure the breadth of existing and emerging risks is addressed. The **6clicks Risk Review for Teams** mobile application can be used to support this *top-down* type of process.

Either way, following risk identification, the next step is risk assessment. The outcomes here can often seem arbitrary, i.e. subjective choices made by the person completing the assessment.

A thorough assessment should be based on likelihood and impact. We can either tell a story in the form of a qualitative assessment, or we can use numbers to quantify the risk using older approaches like *Return on Security Investment* (ROSI) or the new-age *Factor Analysis of Information Risk* (FAIR) framework.

The problem is that most organisations do not have accurate data available to use these more sophisticated models. The 6clicks Mobile App cuts through the complexity with a team-based approach to risk identification and assessment which is a pretty good and rapid way to start!

The actions taken to then address identified risks are known as 'risk treatments'. This is an area of both confusion and opportunity in risk management.

In many cases, we do not know what is going to be useful to mitigate a given risk if it is not within an acceptable threshold. We might be guided by control libraries, but there are so many control libraries to choose from, let alone controls. What if we could learn from what works and recommend the most efficient or effective controls? If only we had some form of contemporary technology that could help us with this selection process…

Misunderstanding What Is at Stake

Mitigating risk or achieving compliance is not an endgame. It is a means to an end, such as enabling objectives, maintaining privacy or being safe. If you have implemented an effective security program based on ISO/IEC 27001, for example, you will have (or at least should have) identified your information assets. Most commonly, hardware, software or files are identified on an 'asset register'.

Even though this is useful information, the purpose is to identify the information types at risk, such as personal information, financial information or some form of intellectual property. This is much harder to identify. Perhaps you use *Regular Expressions* (RegExs) to identify certain patterns like credit card numbers, other identifiers, or keywords. This is also partly useful, but you still often need a human to contextually identify, or at least validate, the full gamut of information types that have been identified.

Once you have successfully identified your information types, you need to understand where the information is stored or processed, where it goes and who has access to it, including internal groups and external stakeholders (e.g. both customers and suppliers). Firewall logs, network traffic analysers and 'WhoIs' may give you some clues. But again, you will probably need a human to correlate technical information about 'what is' and your view of what in fact 'should be' defined in your contracts. No technical solution alone can determine the value you place on certain information types. So, when it comes to classifying information types in terms of value (and not just RegEx), you will probably need that dang human again.

> *'When I think of our cloud operating system for risk and compliance, I see the remarkable advantage it gives to everyone. When I think of what happens once Hailey is unleashed, the legacy it will leave in its wake is massive'.*
>
> - Anthony Stevens, 6clicks CEO

That asset landscape is complex. There is a wide variety of components in the asset landscape that can be automatically mapped and managed today, such as hardware, software, ports and services. The relationship these components have with systems and information classification, as conceptual elements, is not easily automatically mapped.

Maintaining these associations manually is tedious. We need technology smarts that can take input from humans, in terms of value, and group the array of components into business systems with business owners. Even better, the technology smarts to augment technical information related to vulnerabilities and configuration weaknesses with this business context. No mean feat.

ENTER MACHINE LEARNING TECHNIQUES

Algorithms Begin Eating the World

The Artificial Intelligence/Machine Learning (AI/ML) revolution has dramatically changed the world we live in. From Alexa to Google Translate to Netflix, algorithms are using vast computational power to highlight insights, draw inferences and make recommendations with accuracy that would have been the preserve of science fiction only a few years ago. And not a moment too soon – the sheer volume of data available in the information age means that much of today's heavy lifting needs to be done by computers.

What Defines an AI Algorithm?

As the use of artificial intelligence proliferates, an obvious question to ask is: what defines AI? To start with, the building blocks of the AI revolution are based on good old-fashioned mathematical foundations – matrices, statistics, calculus and linear algebra – which are combined to look deeper and see further than any human could.

Briefly, artificial intelligence models based on mathematics can make predictions from existing data without being programmed with explicit rules. Simultaneously, contemporary AI models can change over time, learning from previous mistakes.

AI algorithms are trained to find patterns from existing data, without being pre-programmed with expert knowledge – in much the same way that infants learn language without reference to explicit rules (e.g. terms like 'noun' and 'verb' are learnt well after basic speech).

Why Not Rely on Human Experts?

The advantages of not relying on explicit rules derived from expert knowledge are obvious. Most humans would struggle to articulate the formal rules of grammar, despite our intuitive understanding of what is being communicated. The ability of an algorithm to adapt to its environment is also key, given that hard-coded rules will likely become outdated when the nature of business changes.

Speed is another important element of an AI model given that typical processing times are several orders of magnitude faster than those of humans. Humans are also, by nature, subject to variation, and lack the repeatability and reliability of a computer algorithm. Although domain expertise has an important role in designing and evaluating AI algorithms, in many cases it can be supplemented by relatively inexpensive, highly accurate machine learning code that adapts to new trends in the training data.

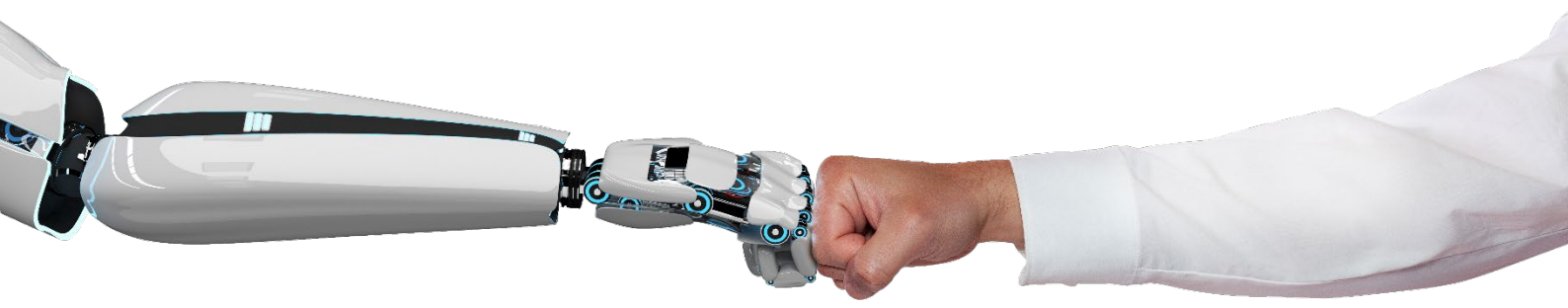
Natural Language Processing

From a technical viewpoint, Natural Language Processing (NLP) is one of the most interesting AI applications. The inherent unstructured nature of the data renders any explicit rule mapping useless, meaning that any useful AI algorithm must necessarily compute the semantic meaning of the input sentence.

The format of the input data as text rather than numbers also presents an intriguing problem. The relationship of numbers to each other is well defined and understood ('greater than', 'square root of', etc.). However, no similar objective mappings exist for words. The scale of the problem becomes apparent when one considers examples such as the grammatically correct sentence *'Buffalo buffalo Buffalo buffalo buffalo buffalo Buffalo buffalo'*, where the only word in the sentence has three different meanings.

In the world of NLP algorithms, words are converted to numerical form, called word embeddings. These representations capture, to a lesser or greater extent, the semantic meaning of words as they relate to each other. This field has been significantly advanced in recent years with organisations such as Google (Word2vec model), Facebook (fastText model) and Stanford (GloVe model) releasing word embeddings pre-trained on billions of words. Although highly specific domains still required custom word embeddings related to their specific area, researchers could, for many cases, simply apply off-the-shelf word embeddings to their specific problem.

Here, researchers now had a way to numerically represent words in a manner that captured their meaning. Yet, there remained the prickly issue of analysing entire sentences. Most traditional AI algorithms, such as 'neural networks' and 'decision trees', treat every input independently, without considering sequences. This renders them ineffective for NLP applications where the order of text matters tremendously. Take the sentences, *'John likes tea, but not coffee'*, and *'John likes coffee, but not tea'*, for example – any application that failed to consider the sequential nature of this data would not extract the correct semantic meaning.



Recurrent Neural Networks

Enter Recurrent Neural Networks (RNNs) – a class of AI algorithms that was specifically designed to handle sequential data. RNNs were initially proposed in the 1980s yet suffered from the problem of vanishing gradients. Two developments of this basic idea – *Long Short-Term Memory* (LSTM) and *Gated Recurrent Unit* (GRU) – resolve this issue by using a combination of features called 'gates' to carry information over several timestamps or phases without dilution. These architectures allow algorithms to read in long sentences, determine which words are key, and extract the required meaning. Most importantly, because this is achieved without explicit rules, these implementations can be adapted to a range of applications if applied judiciously.

Transformers

LSTMs were the avant-garde of NLP for some years but suffered from the inability to use words later in a sentence to infer the meanings of words at the beginning of the sentence. The word 'banks' – in sentences such as *'The banks lowered interest rates'* and *'The banks of the river overflowed'* – would be given the same meaning. This was somewhat mitigated using bi-directional LSTMs – one forward and one reverse – but the core problem remained: Each LSTM was unable to use future information to inform its judgement of the current word.

The next big step in NLP addressed this shortcoming by the introduction of the 'Transformer'. Transformers process the entire sentence in parallel and use attention mechanisms to weigh the importance of every input when calculating the semantic meaning of each word. Upon the introduction of the transformer architecture, several NLP benchmarks were surpassed and a new 'state of the art' was established. Models such as GPT-2, BERT and XLNet currently find use in many complex machine learning tasks such as machine translation, document summarisation and speech recognition. These clever neural network architectures have revolutionised our ability to effectively extract meaning from text and push the boundaries of NLP ever further.

THE FUTURE OF AI-GUIDED RISK & COMPLIANCE

Incidental Compliance

Our goal is for you to achieve compliance through sound decision making, and for you to demonstrate compliance as a part of normal business practice. Viewing compliance as an annual exercise in preparation for audits is not a sustainable approach for any business. It is also unlikely to meet the increasing expectations of regulators.

So far, we are thrilled to have applied and trialed research related to the concepts explained in this paper to a custom NLP model related to cybersecurity standards, laws and regulation.

We use AI algorithms to generate provision-to-provision similarities across cybersecurity standards. Operating far beyond the level of simple word matching, the model extracts the semantic meaning of domain-specific clauses.

This provides a broad view of similarities to other provisions in a fraction of the time that would have been required of a human.

The 6clicks Model

This added functionality will allow **6clicks** customers to generate a snapshot view of their compliance status against a corpus of relevant 'authority' documents, with the option to drill down into specific standards wherever required. The ability to assess opportunities to demonstrate compliance with other standards will enable companies to decrease their risk profile, while optimising their compliance outcomes.

'To improve is to change; to be perfect is to change often'

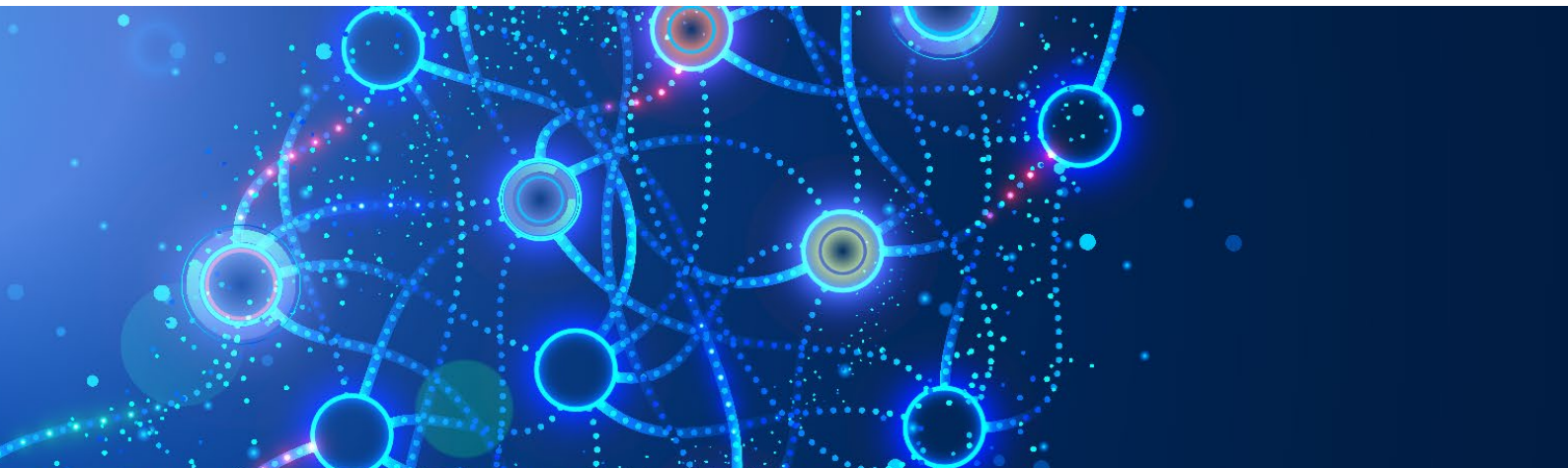
– Sir Winston Churchill

An admirable philosophy by Churchill that has been deeply embedded into the implementation framework of our model. A robust mechanism for feedback and re-training of the model allows for continuous improvement on existing data, as well as new standards that are added to the document library.

6clicks is working to achieve a general release of its provision-to-provision mapping capability. Once successfully deployed for the cybersecurity domain, **6clicks** will be able to expand the training of the NLP model to address the other domains of risk and compliance.

We are also working to streamline other pre-existing mapping functionality, including:

1. Provision-to-question mapping for traceable assessments,
2. Simplifying responses to assessment questions (with suggested responses), and
3. Provision-to-control mapping used when building a unified control framework (one framework to rule them all!).



The Quantified Risk Holy Grail

6clicks is also looking to expand its machine learning techniques to address risk identification, assessment, and treatment challenges. By leveraging the lessons learnt, we expect our machine learning will be capable of rapidly identifying critical gaps in risk identification, as well as discrepancies in risk assessment, and providing the most effective options for risk treatment.

The Asset Treasure Trove That Leads to Wisdom

Our final frontier (at least for now) is for machine learning to assist with the identification of information assets, their value, and the relationship information has with systems, hardware and software. We will help identify the criticality of various technical data, including threats, vulnerabilities and weaknesses using our wealth of contextual information.

To lead the globe in better risk management, our goal is to enable an effective translation of technical exposure to business risk that is informed by our machine learning techniques.

This will be overseen by humans yet operate at the speed of light.

Join us on our journey – follow us on **LinkedIn**, **Facebook** and **Twitter**.

About 6clicks

6clicks is a cloud-based platform built to automate inbound, outbound and internal risk assessments for cybersecurity, modern slavery and beyond.

Our solutions are available for enterprises, start-ups, government bodies and service providers looking to take a proactive and digital-first approach to risk management and compliance.

Follow 6clicks on Social

6clicks

6clicks.io

6clicks Official

6clicksTV