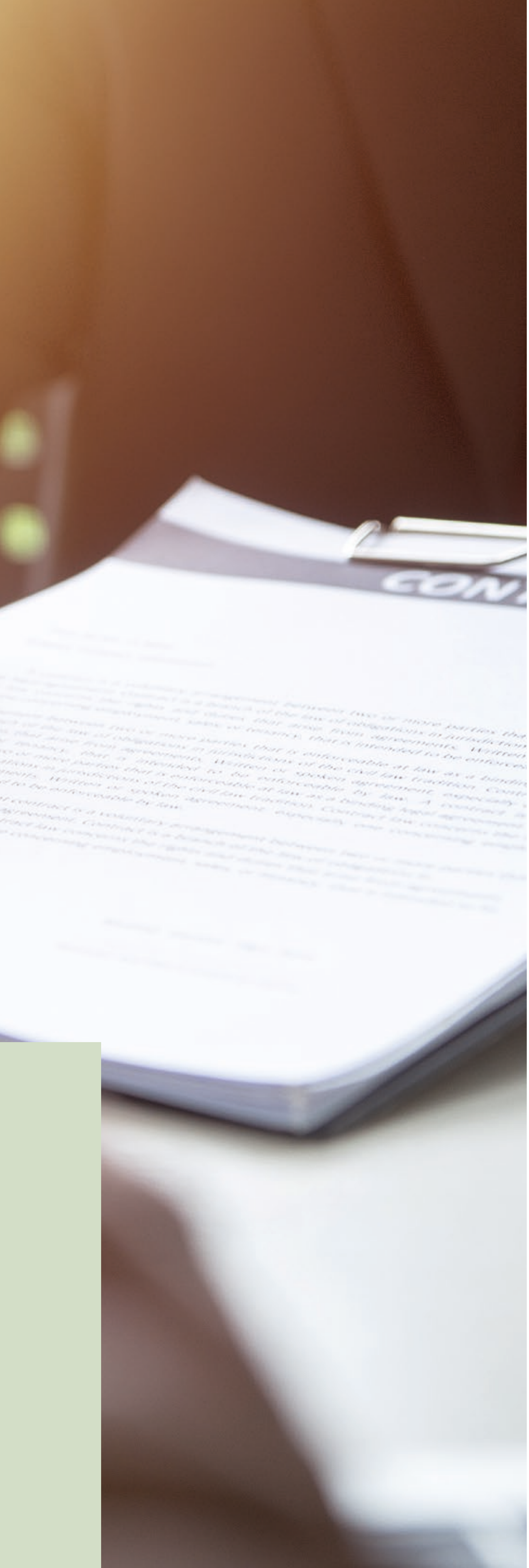
Tackling the “trust barrier”

The challenge of developing Contract Extraction tools

Obtaining more and better insights, gaining competitive advantage and improving business processes: these are some of the reasons why organizations want to make data-driven decisions based on the use of innovative tools. But how do we know if the insights obtained from these tools can be trusted? Accompanied by an example of a Contract Extraction tool we discuss the approach to Trusted Analytics and how trustworthy tools can be realized.

Hanna Nap MSc
is Consultant Forensic Technology
at KPMG.

nap.hanna@kpmg.nl



INTRODUCTION

The Industry 4.0, Internet of Things, Artificial Intelligence, robotics, innovation, and machine-to-machine are all buzz words in the world of innovation. The digital world is growing and companies are aware that they need to keep developing and make better use of their abundance of data. By centralizing data, analyzing it by using innovative techniques and turning this into actionable insights, business processes can be improved and even new business models can be realized. Data-driven actions and decisions can help companies to better understand their customers, to improve their supply chains, to become more productive and realize more profit, and to gain a competitive advantage.

Not only start-ups but also multinationals are developing innovative tools for their own organization or clients to add value, make better decisions and improve their business, because a sense of urgency is present. According to the KPMG Global CEO Outlook of 2017 ([KPMG17-2]), which is an annual research of the issues and priorities CEOs are focusing on, 72% of the companies expects to have a high investment in data analytics tools in the coming three years. Furthermore, 67% expects a high investment in cognitive technologies (among which Machine Learning and Artificial Intelligence).

However, although the sense of urgency is present, which is essential to adopt innovations, there is a crucial barrier concerning the usage of innovative tools, i.e., “trust”. Can tools, which perform actions and make decisions based on the data without intervention of a human being, be trusted? How can the developers of tools identify and tackle this “trust barrier”? And did we tackle it when developing the Contract Extraction Suite tool?

> It is all about the risk that people are willing to take

Below we will try to answer these questions by shortly discussing the KPMG model of Trusted Analytics, as already discussed in a previous issue of *Compact* ([Pato17]), and by discussing whether and how KPMG Forensic Technology covered the four anchors of the trust model when developing the Contract Extraction Suite (hereafter: CES).

TRUSTED ANALYTICS

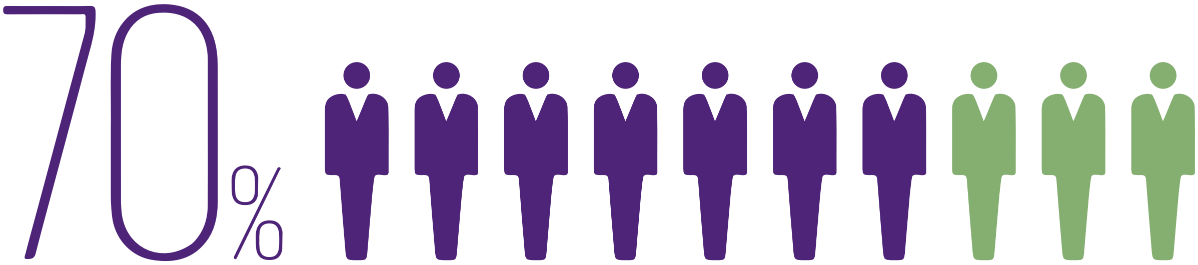
Measuring trust

As knowledge of technology increases, so does the number of innovative tools which use techniques such as text analytics, artificial intelligence and neural networks. Where in the past only IT specialists were able to analyze data and make data-driven decisions, nowadays each employee within an organization is able to use tools, obtain insights and translate them into decisions and actions. However, although users do not need to understand the underlying approach, such as the chosen algorithms, the approach needs to be trustworthy. This trustworthiness is among other things based on the data quality, the effectiveness of the chosen algorithms and the decisions made by the machine. In other words, it is not only about the data that we use (which is also still important), but how we use it, and how it results in the decisions that are made. It is clear that only when the tool is trustworthy and the profit outweighs the costs, people are willing to use it. But how can we measure trust in algorithms, decisions and data quality?

It is all about the risk that people are willing to take, better known as "risk appetite". Instead of asking "Do you trust the underlying approach of the tool?", you need to ask yourself the following questions: "What is the risk when the tool uses the wrong technologies or makes the wrong decisions and am I willing to take this risk?" In some situations a small error can already have a big impact. Think for example of self-driving cars. Self-driving cars are cars that drive without the help of a human being by the use of sensors and algorithms. A small error in the system can result in a car accident, which is a risk that, according to a research by Multiscope, most people are not willing to take ([MULT15]). Although 62% of Dutch consumers is positive about self-driving cars, 80% still wants to be able to take control of the car.

We know from experience that the risk appetite concerning the usage of innovative tools is also very low, since these tools are the foundation for the decisions made within the organization. Obtaining incorrect insights and making wrong decisions may not only affect the supply chains and profit, but in the worst case may also affect the brand reputation and therefore the trust by customers and stakeholders in the product or service. According to the KPMG Global CEO Outlook of 2017 ([KPMG17-1]), 61% of the questioned CEOs mentioned that building greater trust among customers and external stakeholders is a top three priority, due to the awareness of the potential impact on business by negative public opinions and the growing importance of the reputation and brand concerning business success. Eventually this results in organizations decreasing their risk appetite.

Figure 1. The reputational risk when using data and analytics ([KPMG16]).



of organizations agree that by using **data and analytics**, they expose themselves to **reputational risk** (e.g. data breaches, mis-selling of products and services).

The Four Anchors of Trusted Analytics

As discussed in the *Compact* 2017/2 edition ([Pato17]), Trusted Analytics is a term being used for the implementation of analytics that can be trusted. People want to know if the output of the implemented analytics is correct by using the correct data, implementing the right technologies and making the right decisions. In 2016, Forrester Consulting, in commission of KPMG International, examined the power of trust in D&A by exploring organizations' capabilities across four anchors of trust ([KPMG16]). These anchors are: quality, effectiveness, integrity and resilience. By focusing on and strengthening these anchors, developers will be able to tackle the obstacle of trust, since it makes tools more trustworthy. The potential users of tools can also use this model to determine the trustworthiness of the tool and to identify risk areas. The four anchors of trust are discussed below.

Quality

Quality is a broad term and one of the most mentioned anchors concerning trust. Organizations are aware of the importance of data quality during the whole process of data analytics (from importing data to obtaining results), but it is also a challenge, since the storage of data and its regulations are growing.

To determine the quality of the tool, different aspects need to be investigated, since the quality depends on multiple factors. These aspects include the following:

- appropriateness of the data sources;
- quality of the data;
- rigor behind the analytics methodologies;
- methods used to combine data sources;
- knowledge and implementation of best practices;
- expertise of data analysts and scientists.

From these key gaps in quality, organizations consider good quality of the data as the most challenging one.

Effectiveness

Effectiveness is about the performance of the tool. Do the tool and its output work as intended and does it deliver value to the organization? The effectiveness of a tool can be measured by determining the confidence in the:

- effectiveness of the tool in supporting business decisions;
- way the tool and its output are used across the organization;
- accuracy of its model in the prediction of results;
- appropriate use of the tool by employees to make decisions and complete tasks.

However, according to the survey of Forrester Consulting ([KPMG16]), many executives find it difficult to measure the ROI and its value to the organization. Only 47% of the executives declare that they check and monitor the effectiveness of data models in supporting decision-making. Furthermore, 42% says they track and monitor the impact of incorrect insights and actions by misusing/incorrect analytics.

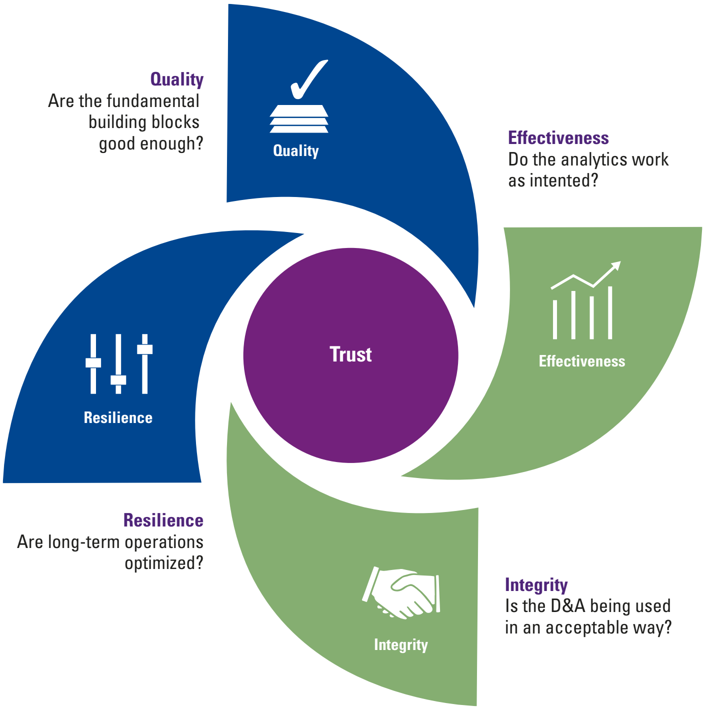
Integrity

The anchor integrity is especially related to the "correct" use of the data, from being compliant with rules and regulations to the ethical use, such as profiling. The questions that need to be asked are: How does the tool use the data and is it in compliance with laws and regulations, for example concerning data privacy? To answer these questions and determine the integrity of a tool, the confidence of the organization in the following aspects need to be checked:

- alignment with relevant rules and regulation;
- transparency (with customers and for regulatory purposes) of the way the data is collected, stored and used;
- evaluation of how customers think of the use of their data;
- alignment to ethical responsibilities and policies.

Although it might sound to some people like a new topic in the field of analytics, it is a very important anchor, due

Figure 2. The four anchors of trust ([KPMG16]).


Quality
Are the fundamental building blocks good enough?
Quality
Effectiveness
Do the analytics work as intented?
Effectiveness
Trust
Resilience
Resilience
Are long-term operations optimized?
Integrity
Integrity
Is the D&A being used in an acceptable way?

to the rapidly changing regulation (one may think of the GDPR) and the impact when actions are unethical or not compliant. It may not only have an impact on the internal trust, but also on the public trust, and may therefore cause brand damage.

Resilience
The last anchor concerns the resilience of tools when challenges and changes occur. Is it secure against cyber attacks for example? And if the organization needs to extend or change functionalities of the tool due to new data sources, is that possible?

Resilience of a tool can be measured by investigating the organization's confidence in:

- the ability to adjust governance policies to data use scenarios;
- how the authorizations to access and use data are controlled;
- how data changes are tracked and reviewed;
- how cyber assurance is managed.

TACKLING THE OBSTACLE

By measuring against the four trust anchors, identifying the gaps and closing these gaps, we are able to tackle the obstacle of not being able to trust the usage of innovative techniques and tools. As mentioned earlier it not only helps (potential) users to determine the trustworthiness, but also the developers of tools. That is, during the process of development they need to keep measuring against quality, effectiveness, integrity and resilience in order to identify gaps and improve their tool by closing these gaps. In other words, to create a trustworthy tool.

To examine how the trust model can be applied to the development of tools in practice, we zoom in on the KPMG solution CES.

Contract Extraction Suite

The CES is a tool that extracts information, in form of pre-defined data points, such as the start date, end date and price, from unstructured contracts. By extracting this information through using innovative techniques and creating a relational database, an organization can easily obtain an overview of all current and former contracts.

The sense of urgency
The reason for the development of this tool, and also the cause of the sense of urgency, is the new lease standard IFRS 16, which will be active from 2019. This lease standard requires leaseholders to add lease constructions to their balance sheet, which will make the assets and liabilities visible on the annual financial statement. This change in regulations has an impact on many companies when they do not have an overview of their current contracts, for example because of the high number of contracts or the existence of hardcopy contracts. Furthermore, when the number of lease contracts is high, going through these files and extracting information from the contracts manually is time-consuming. Therefore, a tool that scans these documents, extracts lease terms from them by using text-mining techniques and centralizes these data, can help organizations efficiently turn data into insights.

The trust anchors
The reason to use this tool differs per organization. Each organization wants to extract information and obtain insights from contracts, but the actions and decisions that result from these insights differ. This means that also the level of risk when making the wrong decisions or taking the wrong actions differ per organization, which affects the risk appetite. That is, when the tool is used for IFRS 16, obtaining incorrect insights may result, in the worst-case scenario, in a substantive error which makes the risk tolerance low.

In order to be able to trust the CES and to trust the insights, decisions and actions that result from using it, a trustworthy tool needed to be created.

Quality
To guarantee sufficient quality, different quality checks were performed during the development of the tool, but are also done when using the tool. The first check is after scanning the hardcopy contracts and performing Optical Character Recognition (OCR) on these documents, which extracts text from an image. To determine the quality of the contract, the tool determines the percentage of words occurring in a created dictionary. When this percentage is low, this may indicate that the OCR is not performed correctly due to contract quality. Therefore, the scanned document is not selected for automated term extraction. An example of this is when someone wrote with a pen on the contract or the text is faded. During the development of the tool, an additional manual check was performed to determine whether the assumption concerning the performance of the OCR was correct and whether the contract indeed should not be selected.

After preparing the data, automatic language detection and template detection are performed, which also have the purpose to select contracts that work well with the chosen algorithms concerning term extraction. That is, the developed search engine supports different languages (e.g., Dutch, Spanish and French), and next to lease contracts, it supports procurement contracts and subsidy contracts. When the language, contract type or template cannot be identified or is not supported, the contract is not selected for automated term extraction. And besides that; template detection has another purpose. By clustering the contracts based on their characteristics, knowledge of the variety of contracts is obtained, which is used by the developers to improve the tool.

Based on the identified language and type of contract, a developed algorithm is chosen to extract the information from the contract, by using text-mining techniques. To develop effective algorithms, the developers train the algorithms on the different contract types and languages, with help from subject matter experts and by manually comparing the outcome with selected contracts, performing root cause analyses when mismatches occur, and improving the algorithms, which is a continuous process.

It was a conscious choice to not create a self-learning model using Artificial Intelligence. That is, a self-learning model works well when it is trained on a large number of contracts, with different contract types, different setups, and different languages. However, when the model is only trained on contracts of the same contract type and same setup, and a contract with another setup is loaded into the tool, it may not identify the data points well.

The steps concerning data preparation and processing are illustrated in Figure 3. It shows that the CES actually consists of two different tools: the extraction tool and the validation manager. The extraction tool contains OCR, language detection, template detection and term extraction. When the selected data is processed, it is imported in the validation manager to validate the results, which is an important element concerning the trust anchor effectiveness.

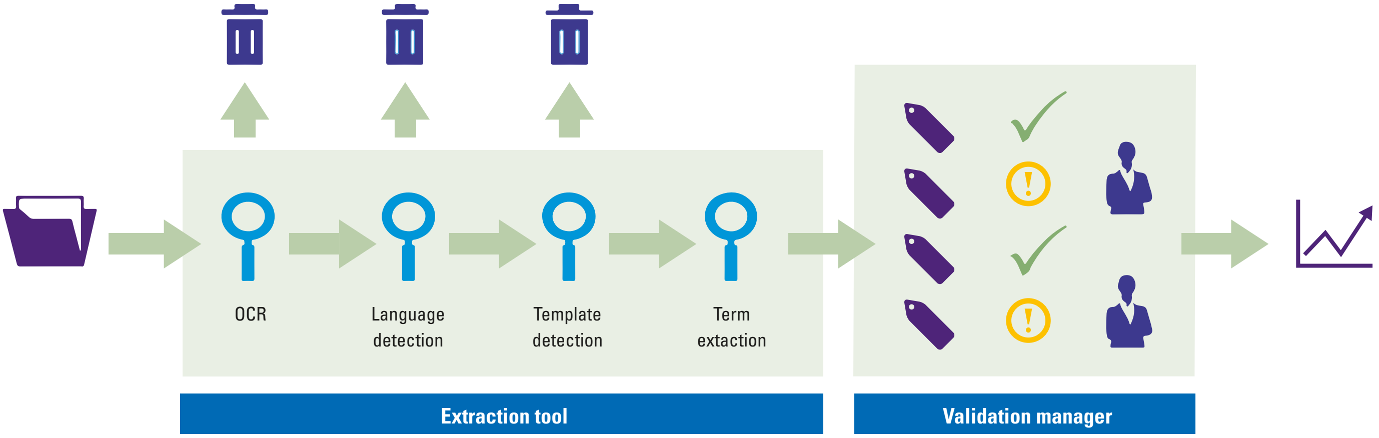
Effectiveness

To guarantee the effectiveness of the tool, the CES incorporated a data validation workflow (the "validation manager" tool), as shown in Figure 3. This workflow focuses on automated validation procedures and minimizes the number of manual checks and corrections that have to be made. That is, when performing term extraction, each detected term obtains a numeric value, which indicates the likelihood that the selected text is indeed correct. When the likelihood is less than a user-defined value (depending on the risk appetite), the user of the tool needs to validate the identified term and change it if needed. In this way, users are able to improve the accuracy, without being time consuming. This functionality in the user interface creates trust, since users feel that they are in control. As an additional validation, also the terms that are considered to be extracted correctly by the tool can be sample-checked via the same workflow. In this way, users that need to rely on the tool can validate the results.

Integrity

The integrity of the CES is a fundamental ingredient due to the impact when the integrity is violated. First of all, most contracts contain limited personal information since in most cases it is business-to-business. Furthermore, each change in the workflow is tracked by logging, which makes it transparent and easy to determine which actions are taken and by whom. And lastly, the tool

Figure 3. The Contract Extraction Suite.



contains user access management and different user roles so that access and rights can be controlled. When a user is added to the tool, he/she has no access to the contracts by default until contracts are assigned.

Resilience

The authorizations of users are managed, but there are more factors that influence the resilience of the tool. Looking at the flexibility to change or extend functionalities, this tool can be improved. It takes time to search for data points in new contract types and identify new data points, since the developers need to create new algorithms or adjust the algorithms to stay effective.

However, when a data point is not identified correctly it does not or barely affect the identification of other data points. The same holds for the different contracts. When one contract contains an issue and therefore has poor results, it does not affect the performance on other contracts.

Lastly, the term extraction can be performed separately from the validation, which makes it optional to perform the extraction in a separated network. This will decrease the risk of cyber threats.

Is the CES a trustworthy tool?

Whether the CES is trustworthy still depends on the risk appetite of the potential users. It is clear however that the developers of the tool focused well on the trust anchors (mainly on quality and effectiveness). That is, only contracts with a certain quality and with certain characteristics are selected so that the developed algorithms are effective and correct insights are obtained. Furthermore, together with subject matter experts the developers were able to create and train the algorithms. Lastly, the users are able to validate and adjust the results by using the validation manager tool. This last element does not only improve the effectiveness of the tool, without being time-consuming, but it also creates control. The users of the CES feel they are in control without being forced to know the back-end of the tool. Although the CES can be improved to increase the trust (mainly concerning the trust anchor resilience), the quality and effectiveness of the tool are high.

CONCLUSION

More and more innovative tools are developed by not only start-ups, but also multinationals. The sense of urgency is there, but the crucial barrier for potential users is trust. Are tools trustworthy? Yes, they can be. Users are able to trust tools – without understanding the underlying techniques and methods that are used – when the developers focused on the four anchors of trust when developing the tool. That is, during this process they need to keep measuring against quality, effectiveness, integrity and resilience in order to identify gaps and improve their tool by closing these gaps. In other words; to tackle the “trust barrier”.

How these gaps can be closed differ per solution. Concerning the CES, different elements have improved the trust, but one of the most powerful elements is the validation manager tool, through which the users feel that they are in control without having to understand the back-end of the tool and without needing to undertake many steps manually.

Jori van Schijndel, Patrick Özer and Bas Overtoom contributed to this article.

References

[KPMG16] KPMG International, *Building trust in analytics – Breaking the cycle of mistrust in D&A*, https://assets.kpmg.com/content/dam/kpmg/xx/pdf/2016/10/building-trust-in-analytics.pdf, 2016.

[KPMG17-1] KPMG, *Disrupt and grow – 2017 Global CEO Outlook*, https://home.kpmg.com/content/dam/kpmg/nl/pdf/2017/advisory/2017-global-ceo-outlook.pdf, 2017.

[KPMG17-2] KPMG, *Economic outlook and business confidence*, https://home.kpmg.com/content/dam/kpmg/nl/pdf/2017/advisory/2017-global-ceo-outlook-dutch-results.pdf, 2017.

[MULT15] Multiscope, *Laat de zelfrijdende auto maar komen!*, http://www.multiscope.nl/persberichten/laat-de-zelfrijdende-auto-maar-komen.html, 2015.

[Pato17] J. Paton MSc and M.A.P. op het Veld MSc RE, *Trusted Analytics – Mind the gap*, Compact 2017/2, https://www.compact.nl/articles/trusted-analytics/.

About the author

J.H. Nap MSc is Consultant Forensic Technology at KPMG. She is a data analytics expert who is interested in smart technologies and innovative tools. Within the organization she focuses on fraud detection by identifying fraud risks, translating these risks into analyses and implementing them in order to help companies be in control.