



Text analysis: just do it!



Instituut van
Internal Auditors

Nederland

CLIENT

IIA Nederland

AUTHORS

Van Vugt, H.C., Van Duinen, W. and Van Uden, L.



Instituut van
Internal Auditors
Nederland

© IIA Netherlands, 2023

Use of the text is permitted provided the source is acknowledged.

Contents

Introduction	4
Examples, impact and concrete analyses	4
Challenge 1: Textual data	12
I. Collecting personal and sensitive data	12
II. Processing unstructured data	13
III. Analysing with text analysis techniques	15
Challenge 2: Tools	18
Challenge 3: Organisation	21
Challenge 4: Getting started and being successful	25
Bibliography	29

Introduction

Policy documents, agreements and contracts, reports, work instructions, e-mails, minutes, complaints, description fields, newspaper articles, Twitter and websites: they all contain texts that are used daily in organisations, including the audit function. The amount of text that organisations produce and collect is substantial and increasing every day; nothing short of an explosion in data volume. Reading and using these texts in investigations is often time-consuming, while objective assessment is difficult. Analysing these texts with the help of a computer (text analysis) can offer a solution to the limitations of traditional techniques. The IIA agrees that text analysis has potential and expects that it will rapidly increase in importance in audit practice; reason for the IIA to publish this guide. The authors of this guide have experienced in their own audit practice that text analysis offers great potential for making audits and investigations more efficient and effective.

Purpose of this guide

With this guide we want to make it easier for auditors to start applying text analysis; both in the planning phase, when defining the scope, and during the audit: to collect evidence, filter relevant documents, research the entire population, etc. We start with examples and the impact our text analysis has had in our own audit context. We then discuss four challenges to make text analysis a success. Early 2022, we conducted a survey among auditors in advance of this guide. This showed that many auditors do recognise the usefulness and potential of text analysis, but also experience challenges. That is why we propose problem-solving approaches to tackle all kinds of challenges. We consider: What is necessary to ensure the success of text analysis projects?

Examples, impact and concrete analyses

E-mails that are automatically placed in your spam folder. E-mails from customers to a company that are automatically assigned to the correct department. Customer reports where unwanted, personal data is automatically masked, to prevent it from being stored in the systems of an organisation. And, one step further: the ChatGPT bot that 'helps' a teenager with homework by producing relevant texts (OpenAI, 2023). Or a model that predicts how often an article, for example on LinkedIn, is read and liked based on its content; and even adjusts the content to improve quality. These are all examples of text analysis that we encounter in practice. Text analytics can be described as a collection of techniques leading to the extraction of knowledge and actionable insights from textual data.

Audit practice includes a wide variety of examples of problems in which text analysis can be applied. Is a mandatory description or explanation field always completed? Is sustainability discussed in conversations with customers? Who is present or absent in meetings (as shown by the minutes) and who therefore are the logical key people to interview in the audit? Which transactions

contain names or country codes associated with sanctioned countries? Which documents are relevant to read for an investigation? The analyses to arrive at answers to these questions are relatively simple to perform, even by auditors with limited knowledge of text analysis.

Of course, there are also examples that require more complex analysis. Are laws properly translated into internal policy? Which complaints involve high amounts, therefore possibly representing a high risk? Is the glossary used by a masking system complete, and can we supplement it with synonyms and other related words? Has the 1st line correctly or incorrectly classified credit reports in a risk category? What is the quality of Know Your Customer (KYC) reports? Which documents from a large database have been missed, but are still crucial in large-scale investigations? See the case in Box 1 for this. For these types of analysis, the use of a data scientist is usually still unavoidable, because the analysis is more complex and knowledge of machine learning may be required. However, it is good for auditors to realise that such analyses are possible and that the impact is greater than one might think.

Reader's guide

For ease of use, we have clustered the challenges:

- **Challenge 1: Textual data:** How to deal with personal and sensitive data? What steps are needed to perform a text analysis? What are commonly used techniques?
- **Challenge 2: Tools:** What tools are available? Where can relevant knowledge be obtained to learn how to use the tools? Focus is on the tooling we already successfully apply in our work.
- **Challenge 3: Organisation:** What is needed in an audit department to make text analysis a success?
- **Challenge 4: Getting started and being successful:** What additional tips do we have to increase the chance of successfully applying text analysis in audit practice? The tips are mainly about challenges surrounding the project approach.

Data analysis is starting to find its place in accountancy audit methodology (Gold & Majoor, 2022; Liew, Boxall, & Setiawan, 2022; NBA, 2019) and also increasingly in the audit function (Johnson, Wiley, Moronay, Campbell, & Hamilton, 2021; Gartner, 2020; Wang & Cuthbertson, 2015; Eilifsen, Kinserdal, Messier, & McKee, 2020; Deloitte, 2013). With this guide, we hope to contribute to making text analysis an integral part of audit practice. Our motto is therefore: ***Text analysis: just do it!***

Impact of text Analysis

Auditors who have used some form of text analysis in their work are often enthusiastic. They have noticed that their work becomes more effective and efficient. More effective by including the entire population (viewing all documents instead of sampling); by obtaining more and different insights leading to increased quality of assurance; through insight into the operation of control measures and the selection of key people for the audit. More efficient due to savings in reading time and/or due to the decrease in documents to be analysed, which in turn can lead to increased focus and a reduced or clearer scope; by making it easier to apply the same analysis periodically. In short, the available texts are used for the benefit of the audit.

An additional advantage is that the results can often be immediately presented to the auditee after analysis. This makes fact-based discussions easier. Like other forms of data analysis, text analysis will ensure that findings are acted upon sooner (Gartner, 2017), because they are data-driven and insightful. Often analyses are also relevant for the 2nd line to carry out inspection work. Sometimes insights are so useful, relevant and/or surprising for the auditee and/or the 2nd line that internal analysts will continue to build on the text analyses performed by the audit department and even proceed to standardise these. The audit department can then build on that in the future and make a switch from data-oriented auditing with text analysis to systems-oriented auditing with text analysis.

Concrete analyses

The countless examples of text analysis applied in audit departments are often relatively simple to perform. These are easy to apply, even by auditors with limited knowledge of data analysis. Of course, there are also more complex analyses, which require more knowledge. For example, prediction models based on text characteristics. The use of a data analyst/scientist is usually still unavoidable for the more complex analyses; these are best left to data scientists for now.

The analysis examples below are (roughly) sorted from relatively simple to more complex. We always start with the question, followed by a brief discussion of the analysis. In any analysis, the auditor, wearing his 'auditor glasses', inspects, interprets and assesses the results, i.e., considers what the analysis means in the context of the audit and what follow-up questions it may raise that can once again be answered with follow-up analysis.

LENGTH OF TEXTS

- Is a (mandatory) description or explanation field always completed? If so, does it involve the use of more than three words? A volume of fewer than 4 words is often not meaningful, such as 'see file', 'not applicable', '-', or '..'

Analysis: Counting blank fields and counting the number of words in a field. Sorting by frequency.

SEARCHING FOR WORDS

- Which documents are relevant to read for an investigation?

Analysis: Making a glossary of interesting words for the investigation (e.g. 'Covid' and 'Coronavirus'). Counting these words in the documents. Sorting by frequency to subsequently select the most relevant and interesting documents.

- Is sustainability discussed in conversations with customers?

Analysis: Creating a glossary of words related to sustainability. Counting these words in meeting reports. Counting per word and per report. Sorting by frequency to quickly see which meetings did not discuss sustainability and which meetings had sustainability as a leading theme.

- Is personal information properly masked by the masking system? Is the glossary used by the masking system complete?

Analysis: Searching for synonyms of words from the existing glossary, using an existing synonym dictionary of the Dutch language. Searching for all words from the existing glossary and for the relevant synonyms in the documents after processing by the masking system. Sorting by frequency to see which words are often missed by the system.

- Which transactions contain names or country codes associated with sanctioned countries (and are possibly high risk)?

Analysis: Using country lists with country codes as a glossary. Searching in transaction fields (from, to, comments) to find transactions that are possibly linked to sanctioned countries.

SEARCHING FOR PATTERNS

- Who are the key people in a particular process or project? This helps with the correct selection of people for interviews and the reconstruction of a process.
 - Who sent and received e-mails (in a certain context)? Analysis:

Searching for patterns such as 'from:', 'cc:' 'bcc:' in e-mails (.msg files).

- Did that change over time?

Analysis: Searching for 'date:' and for patterns such as '1 or 2 digits, space, 1 word, space, 4 digits' to extract dates from texts.

- Who was present and absent from meetings in the past year?

Analysis: Searching for patterns such as 'present:' and 'absent:' in minutes and then filter and count the names entered.

- Which complaints involve high amounts, therefore possibly representing a high risk?

Analysis: searching for patterns such as '4 or more digits, space, euro'.

COMPARING TEXTS

- Are texts continuously copied and pasted into comment fields? This can be an indication of the quality of first line work.

Analysis: Grouping text fields and counting how many times each text field appears. Grouping based on 100% identical texts, or based on 'near-matches' (e.g. 90% identical texts).

- Are laws properly translated into internal policy? This is relevant, for example, to see whether new or amended legislation is properly translated into internal policy.

Analysis: Counting all words, or better: lemmas, in both laws and policy documents. Comparing the words/lemmas from each law with words/lemmas from each policy document. If words/lemmas appear frequently in a law but not in the policy document, it should be questioned.

SENTIMENT ANALYSIS

- Does the risk category into which a customer is classified match the sentiment in the meeting report?

Analysis: Searching for and counting positive words and word combinations (e.g. 'good', 'positive', 'skilful', 'improved', 'not bad', 'no points for concern') and negative words and word combinations (e.g. 'bad', 'deteriorated', 'problem', 'insufficient', 'not good', 'little improvement'). Selecting a relevant sample, for example those customers classified in the low-risk category and many negative words in the text; These may be customers who belong to a higher risk category given the number of negative words in the text.

NETWORK ANALYSIS

- Who is in contact with who in the organisation and to what extent?

Analysis: Mapping out networks through e-mail analysis: who e-mails who and how often?

CLASSIFYING TEXTS

- Which texts are correctly or incorrectly classified by the first line? Are texts with similar content classified in the same way?

Analysis: searching for topics in the texts (topic modelling). Using algorithms to classify or categorise the text or documents based on topics.

- Which documents from a large database have been missed, but are still crucial in large-scale investigations?

Analysis: Based on public sources, for example: newspaper database, generating relevant search terms through text analysis and using these to search for relevant documents in a database.

The following case (Central Government Audit Service, 2021) gives an example of applying text analysis in an investigation.

CASE

Text analysis for a parliamentary inquiry Gas extraction Groningen

In a parliamentary inquiry, a parliamentary committee orders the client to produce the relevant dossiers. The client would like to provide a complete and transparent answer and therefore enquires with the Central Government Audit Service (ADR) about the findings during the formation thereof.

An auditor can look at the control measures taken and whether they worked well during the search process. However, based on that approach the auditor cannot determine whether the result of the search process is complete. And, if the results were not complete, which documents are missing.

By applying text analysis, it can be objectively and independently determined whether the approach used is complete and which documents may have been missed. The text analysis provides answers to the following research questions:

1. Are the search queries that have been defined sufficient to find relevant documents?
2. Were any documents missed in response to the request for dossiers?

To answer these two research questions, “topic modelling” was applied as a text analysis technique. Public sources, namely parliamentary papers and newspaper articles, were used to compile the set of search questions. Subject modelling has been used to determine the main topics for all relevant parliamentary papers. These main topics were then searched in a newspaper database to realise an enriched set of search queries.

The enriched set of search queries was then used to search for relevant documents in a document management system. This resulted in roughly two million documents, which were then categorised into various topics. Ten topics were classified as most relevant. For each of these topics, documents were stratified that were found by the client’s team and documents that were not found by the team. Based on the mass ‘not found by the team’, something can be said about the completeness of the search queries (research question 1) and which documents may have been missed in the applied search method (research question 2).

Box 1. Case application of text analysis for the parliamentary inquiry Gas extraction Groningen.

```
var b, d=this, e=this
this.$el(c.router.themes
ready"), a(document.body
ready"), c.router.selected
this.undelegateEvents
this.$el("closed").toggleClass
this.$el("previewDeviceButtons
keyEvent: function
maybeRequestFilesystem
this.$el("jackbone.View.extend
listenTo(c.collection,
length), c.announceSearch
function(){c.overlay
this.$el("announce(b)))}}}, render: function
this.$el("renderThumbnails
```

Challenge 1: Textual data

We can cluster challenges related to textual data into three groups:

- I. Collecting personal and sensitive data
- II. Processing unstructured data
- III. Analysing with text analysis techniques

Each of these groups comes with its own challenges.

I. Collecting personal and sensitive data

Textual data comes either from within the organisation itself (such as e-mails, internal documents and evaluations) or from outside the organisation (such as social media, newspapers and websites). Different rules may apply for internal and external sources.

For internal sources it is important to consider how and for what purpose the data has been obtained. Textual data often contains personal data that can be traced back to an individual, such as names, passport numbers and addresses. There may also be sensitive or personal data, such as data containing information about religious or political beliefs, ethnic origin, health, genetic data and trade union memberships. Since 2018, the same privacy legislation applies throughout the EU: the General Data Protection Regulation (GDPR). The Dutch Data Protection Authority states the following: *“Every time you process personal data, it is an invasion of the privacy of the data subject. That is why you may only process personal data if there really is no other option. In other words: if you cannot achieve your goal without this information.”* (Data Protection Authority, 2022). Determine in advance whether personal and/or sensitive data is present and how to deal with it. A number of principles can be considered for processing the data (Schermer, Hagenauw, & Falot, 2018):

- **Necessity:** Limit the use of data to what is necessary to conduct an investigation (data minimisation). The processing must be linked to specific purposes.
- **Relevance:** Only process personal data that is relevant. Often an analysis can also be performed if personal and sensitive data is removed from datasets (anonymised), replaced with randomised codes (pseudonymised) or aggregated in such a way that it is meaningful for auditing. Do not keep the data longer than is necessary.
- **Safety:** Data must be properly secured. For example, do not store the data on a shared drive where unauthorised people can easily access it or do not analyse the data whilst travelling on a train.
- **Confidentiality:** Data must remain confidential. Do not share personal data with other parties without valid reason to do so.

As is the case with internal sources, purpose limitation is also important when using data from external sources. The fact that information can be found freely on the Internet does not automatically justify its use for all kinds of purposes. There are various areas of law, such as database law, copyright law, the right to respect of privacy and the right to protection of personal data, which may impose restrictions or even prohibitions on the collection of external (textual) data for personal use (Wiseman Advocaten, 2020). There may also be specific conditions for the processing of data from an external party. It is important to look into this before enthusiastically starting with, for example, scraping (retrieving information from websites) news items or reviews.

More and more organisations employ a data protection officer, also known as a privacy officer, with knowledge of internal policies to comply with (privacy) legislation. So talk to your privacy officer to find out whether and how certain data can be processed in line with your requirements.

II. Processing unstructured data

Characteristically, textual data is unstructured; the data has (virtually) no fixed format, is not labelled and sentences and words are not structured in rows and columns. It is natural language, not computer language. Unstructured data can come from all kinds of files, for example, *.doc*, *.txt*, *.pdf*, *.msg*, *.ppt* and *.xls*. In these sources, text is often already present and digitised, so that it can be read directly into the text analysis software. Sometimes we deal with spoken language, for example in the case of interviews, or with handwritten text in scanned documents. In both cases pre-processing is needed to convert the input into the correct, digital format. Spoken language can be automatically converted into digitised text by special software (Turner, 2022). Handwritten text can be converted by special software using Optical Character Recognition (OCR); Letters in images are recognised after which handwritten text is converted into digitised text.

To use text analysis, structure is applied to the unstructured data: textual data is converted into structured data in rows and columns. An example of a structured form of text is a matrix with sentences in the rows and an item such as a word in the columns; each word is noted as present or not in the sentence (one or none; see Table 1). To achieve this, one or more processing steps are completed. It is important to think carefully about the (ir)relevance of each step for the purpose of the analysis. After all, in general: the better the quality of the data (and the quality of the data preparation), the better the resulting analysis.

	A	payment	scheme	The	applicant	applies	for	credit
A payment scheme	1	1	1	0	0	0	0	0
The applicant applies for credit	0	0	0	1	1	1	1	1

Table 1. Sample matrix

Processing steps may include:

REMOVING IRRELEVANT WORDS

To achieve a simplified, shorter presentation of the text, punctuation marks such as periods, commas, and question marks are often removed from the text. Fillers are often removed as well. This yields a simplified and small matrix (see Table 2). Fillers include commonly used words with no meaning in their own right, such as articles (a or the) and adverbs (this, that, on) and words that are often used with no explicit meaning, such as 'so to say'. Filler lists are available for different languages or you can create your own filler list. Also, capital letters are often replaced by lower case, so that the computer considers 'Today' and 'today' as identical words. You should, of course, always consider if certain words can be removed and capitalised letters replaced for your specific analysis, or whether it is better to keep the words, for example, if you are interested in the exact number of words used.

	payment	scheme	applicant	applies	credit
payment scheme	1	1	0	0	0
applicant applies credit	0	0	1	1	1

Table 2. Example matrix in which irrelevant words have been removed

CUTTING THE TEXT INTO TOKENS

"Tokenization" is the splitting of text into sentences, and sentences into individual words. There are several tokenizers that each lead to a slightly different result. You indicate in the tool which tokenizer you want to use, depending on the language of the text you are going to analyse. An English word tokenizer splits the English 'she's' into the tokens 'she' and 'is'. In a matrix, the words 'she' and 'is' will now be in separate columns. A white space tokenizer will break up a text into chunks that start and end with a space: in that case, 'she's' is not split, but seen as a single token. In a matrix, 'she's' will be in one column. The Dutch phrase 'Anna's fiets' is split by the whitespace tokenizer into the two tokens 'Anna's' and 'fiets'. A Dutch tokenizer knows that 's' can be seen as a separate, third token (indicating possession). Not every tool has implemented all possible tokenizers. In that case, use an English, German or White space Tokenizer for the Dutch language. These yield pretty good results.

GROUPING TOKENS OF SIMILAR MEANING

Tokens such 'bicycle', 'cycling' and 'cycled' are related to each other. The computer will only see these tokens as related words if we make sure that the tokens are reduced to their 'stem' or 'lemma'. Simply put, the stem is the part of a word that remains when the inflection endings are removed. For example, singulars and plurals of a word (bicycle/bicycles) have the same stem (bicycle). A lemma is slightly more refined, because it takes meaning into account. For example, the word 'better' has 'good' as lemma (and 'bet' as stem). To determine this, dictionaries are needed to link a word to its lemma.

The stems or lemmas are then grouped together so that they can be analysed as a single item. As a result, all stems or lemmas will be in a single column in a matrix. In general, lemmatisation produces better groupings than stemming, but it is more difficult to implement because it requires more knowledge of the language (and not every dictionary is good enough).

III. Analysing with text analysis techniques

Many organisations therefore struggle in the initial phase with choosing the right approach to obtain the desired results from an analysis (Expert.AI, 2023). When performing a text analysis, it is important to choose the right technique to achieve the goals of an analysis. So try to have a clear understanding of the purpose of the analysis before choosing the appropriate technique. Text extraction and text classification, which are catch-all terms for a wide variety of techniques, are widely used:

TEXT EXTRACTION

Text extraction involves a certain term or piece of text being extracted from the data. These can be certain keywords, but also, for example, e-mail addresses, people or nouns. Known techniques used for this are 'bag of words', term frequency, regex, named entity recognition and part of speech tagging. A 'bag of words' contains all individual words from a text. You can then search for specific words or words from a (short) word list. Term frequency (TF) indicates the frequency of a word in a text. Regex can be used to search for a piece of text that follows a certain structure, such as an e-mail address (a text containing the '@' symbol) or an account number. Named entity recognition can be used to recognise and extract entities such as people, organisations and locations in a text. Part of speech tagging can be used to mark words as noun, verb, etc.

TEXT CLASSIFICATION

Sometimes texts in a database are classified into groups or linked to themes. For example, credit reports in a bank are classified by the 1st line in a risk category. An audit question can be: is this classification correct? By examining the texts, a relationship can be discovered between the texts and the classification and an assessment can be made as to whether the classification was correct. Sometimes texts are not yet classified, in which case having texts categorised automatically can be useful. Which group or theme an unclassified text fits in best, is determined by the similarity or disparity between this text and existing texts in the various groups. If similarity to a group is sufficient, the text can be automatically classified into that group. This way, for example, e-mails for customer service are automatically forwarded to the correct department.

Topic modelling, a machine learning technique, can be used by more experienced data analysts to find common topics in documents. This technique was used by the Central Government Audit Service in the case in Box 1. Texts can also be classified in terms of sentiment: is the emotional charge positive or negative? For example, sentiment analysis is used to analyse reviews and can be used by audit to investigate complaints and complaints management.



Challenge 2: Tools

Our survey showed that, despite the desire to apply text analysis in audit work, the knowledge to actually do so is often lacking. Some organisations employ data analysts or data scientists who provide the necessary knowledge of text analysis and software. However, more and more auditors want to work data-driven and perform text analysis themselves. Many text analysis tools and recent overviews of free and commercial solutions can be found on the Internet (Pat research, 2022; Datamation, 2021; Mousa, 2020). For some of the tools mentioned here we noticed that these users do not control the functionality themselves. In this guide we have chosen to focus on the tools that we ourselves use, or have used, for text analysis and therefore have the most knowledge of.

I. Basic functionalities in Microsoft Excel

Microsoft Excel has a number of functions for performing basic text analyses, such as the 'Find' function to check whether a word occurs in the text of a column. The functions that Excel offers are divided into categories and the 'Text' and 'Lookup & Reference' categories contain potentially interesting functions. Of course, Excel functions are very limited compared to the features that the tools discussed below have. We are not familiar enough with other tooling that Microsoft has developed. For example, Azure that offers possibilities for text analysis ([Text Analytics | Microsoft Azure](#)). Azure probably offers more possibilities than Excel.

II. No code environments: KNIME and Alteryx

No programming knowledge is required for KNIME and Alteryx, as they are 'no-code' environments. They have a user-friendly visual interface, allowing users to quickly get to grips with the possibilities and applications of text analysis. Both tools allow the user to perform all kinds of basic and more complex text analyses, such as the examples mentioned in the first part of this guide.

KNIME can be downloaded for free, including an extension that allows text analysis. There is a paid version that offers additional functionality (see <https://www.knime.com/knime-software/knime-hub-pricing>). We do not use this ourselves. Alteryx involves more costs. A free trial version can be downloaded from [Alteryx Designer Free Trial | Alteryx](#).

KNIME offers a learning platform ([KNIME Learning | KNIME](#)) where you can find support material to get started with KNIME: books, courses (online, onsite and at your own pace), technical documentation, certification and more. From words to wisdom (<https://www.knime.com/knimepress/from-words-to-wisdom>) is a relevant book focused on text analysis.

Furthermore, there are some online videos on YouTube (e.g. [Text Mining Techniques - YouTube](#)), as well as sample analyses that can be downloaded on the KNIME hub (<https://hub.knime.com/>). Alteryx also offers all kinds of possibilities (videos, training courses and examples) on [Alteryx Academy - Alteryx Community](#) to learn how to use the tool.

III. Programming code: Python or R

Python and R are programming languages. You need programming experience to use these for text analysis. Data scientists often work well with Python or R, but the threshold for auditors to learn these is too high. Python and R can be used in all kinds of applications, such as Anaconda, which uses so-called Jupyter notebooks ([Using Jupyter Notebook – Anaconda documentation](#)). For text analysis, you can use the NLTK (Natural Language ToolKit; Natural Language ToolKit; <https://www.nltk.org/>) and spaCy (<https://space.io>) packages. There are all kinds of courses to learn how to use these packages (for example via <https://www.datacamp.com/>).

L I F T I N G
O T H E R S
W E
R I S E

Challenge 3: Organisation

Both the structural and incidental application of text and other data analyses in audits probably means a considerable change in the way of working. If audit departments want to sustain their analytical efforts and see these develop in practice, they must embrace change (Deloitte, 2013). And that in a profession where great value is attached to existing methodologies, and in which a conservative attitude and scepticism towards data analysis is common (Li, 2022). Countless books, courses and websites have been written about making changes like this a success. We believe that the eight change steps¹ of Kotter (Kotter, 1996) provide a good overview of what is needed. We will not describe these steps in detail here. We will, however, largely in line with the Kotter model, discuss a number of tips that were mentioned by participants of the IIA Professional Practice day in 2022 and which we ourselves experienced as crucial:

I. Explicitly ask for support from management

The belief that data analysis in audits will (continue to) play an important role in the future is the core to make it a success. It is important that management embraces the change to more data-driven auditing, so as to actually get auditors to work with it and to master a new, data-driven approach. After all, how else do auditors feel supported to invest time and effort in it? Most auditors do not have a data background and need time to build up knowledge of (text) analyses, apply this knowledge and learn from it. If courses have to be followed or tools need to be purchased, it can also mean a monetary investment. This is difficult without explicit support from management. Furthermore, positive appreciation from management for more innovative, data-driven auditors can help as an incentive to apply a data-driven approach more and more.

II. Discuss challenges

Every organisation faces challenges to make text analysis a success. Making goals explicit and discussing them, as well as the challenges to achieve these goals, help to move forward (Van Vugt, 2022). Important questions are: 1) What level of maturity does an organisation or organisational unit currently have? Is text analysis absent, present to some extent, is it

1. Tangible, visualising the need ('sense of urgency')
2. setting up a strong steering group with sufficient resources to manage the necessary change
3. Developing a guiding vision annex strategy to realise that vision
4. Communicating the new vision
5. Encouraging and enabling to act in accordance with the new vision
6. Ensuring visible short-term successes
7. Consolidating improvements and continuing to implement changes
8. Embedding changes in the corporate culture

emerging, or is text analysis already being used at a structural level? 2) What level of maturity is the organisation aiming for? What can and does the organization want to achieve by using text analysis? 3) What obstacles are experienced to achieve the desired level of ambition? How can obstacles be removed?

Challenges, or obstacles, can be various (Shahim, van Praat, Harmzen, & Matthijssse, 2018; Stuurman, 2022): a lack of knowledge, limited data quality and usability, concerns about data security and privacy, the lack of a clear perspective on the application of data analysis, a high expectation of effort (for example, to obtain data, to understand the technique and the correct interpretation of the results), limited facilitating circumstances (lack of time, knowledge sharing and tools) and the complexity of information systems. With the right attitude, challenges will appear to be less complex than expected and can make way for opportunities.

An aid in discussing these challenges is the Optimal Digital Game, which has been adapted to the audit context (Bruinenberg, 2022). It can be played in the start-up phase of audits, during different types of challenges. Discussing challenges in a structured and fun way will greatly help the audit team. Those involved reach consensus and find practical solutions together, which in turn increases productivity in the audit.

III. Work together with 1st and 2nd line

Data that is relevant in an audit context is often produced in the 1st line. Often the 1st line already creates dashboards for this data, for example in a BI department. 2nd and 3rd line are interested in the same data, but often need to view this data from a different perspective. They need their own insights to be able to carry out their auditing activities. Collaboration with 2nd line is important to build on (text) analyses performed by them and to learn from each other's approach and focus. Collaboration with the 1st line is important to understand what data is available, what the data means, how the data is produced, how it is used in processes and dashboards and what analyses are on the wish list, but which are yet to be performed. A positive relationship with business owners is also important, because this makes getting permission to use data in analyses easier and faster, within the framework of the policy of course.

IV. Win support from the IT department

Being able to use analysis software often depends on the policy of the IT department in the organisation. Programming environments are often supported, but auditors are often not programmers. So-called 'no-code' environments make it easier for auditors to get started themselves, but these are often not supported or even recognised by the IT department.

When auditors come across an IT department that does not support certain analysis software or whose policies do not even allow them to use it, it is hard. Creating your own lab environment and integrating this into the corporate architecture at a more advanced stage could be an option to avoid being dependent on IT. If internal knowledge is lacking, external help can be called in to set this up.

Audit departments from different organisations can also help each other by sharing software assessments (analysis that software is safe to use). In our experience, it helps to enter into dialogue with the IT department to discuss needs and solutions within the policy framework.



Challenge 4: Getting started and being successful

We have experienced that even if organisations embrace data analysis and the right tooling and knowledge are available, projects are not automatically successful. Projects can become too large-scale, take too long, data is not available, focus disappears or the impact is marginal. To increase the chances of success, we have a number of relevant practical tips, which we have found to work:

I. Generate ideas and prioritise on impact

Text analysis has more potential to be more efficient or better in some investigations than in others. Apply text analysis in an investigation where relatively large gains can be made. For example, because many documents are involved, or because important audit questions or sub-questions can be answered by means of text analysis. It is of course important to know which questions are most important for the sector, strategy and priorities of your own organisation (Deloitte, 2013).

Workshops and brainstorming sessions are effective for coming up with ideas in groups, for example the entire audit team, on how to apply text analysis to available texts in the context of an audit. Relevant questions are: Which (periodic) audits involve a lot of text analysis? Which audit questions are essential to answer? Which documents or texts are important in this? What does an auditor spend a lot of time on and could text analysis help? The 1st and 2nd lines can be involved in this as well, so that an analysis can lead to important insights for them too.

Once ideas have been generated (for example, on paper or digital sticky notes) they can be plotted on an impact-effort matrix for prioritisation purposes. Ultimately, we are looking for an idea that can have a lot of impact with relatively little effort. As knowledge of text analysis increases over time, the effort to perform the more difficult text analyses will become less and less.

II. Start on time

It sounds obvious, but it holds true for almost any kind of data analysis. And yet this tip is relevant: start on time. For example, if text analysis is used for sampling, it is necessary to do the analysis right at the beginning or just before the official start of the audit or investigation. In general, before carrying out an investigation, it is important to start with:

1. collecting the data. Experience shows that this often takes longer than expected. Firstly, permission is required to analyse data for the intended purpose. It is useful to include the data protection officer in the plans to obtain data from the outset, so that they can join in the deliberation process. Also consider overarching agreements to prevent getting stuck again in red tape for every

small investigation. Secondly, the appropriate supplying party must be found to retrieve the data from the (source) systems and to anonymise the data where necessary. This party too often works with backlogs, so you are not immediately ranked at the top. Both the process and the data collection itself can be time consuming.

2. an analytical approach. Consider how to perform the analysis once the data is available. Hypotheses can be formulated for testing in the analysis phase. For example, in the hypothesis “*Description field x is always filled with more than three words*”, the counting of words in the description field is obvious. A hypothesis-oriented approach makes it easier to interpret the results by focusing on what is important.

3. reading in and exploring the data. Before the actual analysis, it is useful to get an idea of the quality, format and size of the texts. This provides insight into possible difficult points about the data (for example, what do certain codes mean?) for which the business may need to be consulted.

III. Use the available data and check data quality^a

In an ideal world, we perform analyses on a recent, complete set of good quality data, for which the metadata is available as well, so that we know exactly what all the columns mean. However, the reality is rarely ideal. We are dealing with document collections of such size that it is impossible to (quickly) analyse all relevant documents, with fields that have been completed incorrectly, with data from different systems that are difficult to match, with outdated data, etc. We have to make do (analysing) with what we have (data). Do not be discouraged. Prior to the analysis, test whether the data quality is suitable for your purpose (fit-for-purpose). Poor data quality does not automatically mean that you cannot come to relevant insights. Focus on which problems and sub-problems can be answered with the data and which insights can be obtained. In addition to clear conclusions, it is always good to indicate the limitations of the investigation, so that the auditee also understands the situation.

The state of the data can also say something about the organisation itself: poor data quality can be an indication of poorly performing processes or outdated systems. This information can in turn be discussed with the auditee: to what extent are the processes and systems in working order? Using (text) analysis in audits can help the organisation take steps to improve data quality.

IV. Start with basic analyses

The knowledge in text analysis courses can be quite profound and the number of applications and possibilities can be overwhelming. The trick is to start with simple analyses, based on the relevance to a problem, for example by counting or searching for words. If an initial analysis does not have the desired effect, the trick is not to give up, but to consider what the first analysis brought (for example, knowledge about the tool) and to start with a second simple analysis that has the potential to generate impact. This brings us to our last tip.

V. Start small and evaluate

Here the 'think big, start small' principle applies: start with a proof of concept (PoC) and only expand once its value has been demonstrated or others are convinced. Making an initial presentation about the goal and approach early in the process, possibly combined with some initial results, helps to discuss the usefulness of the analysis, to enthuse other auditors or the 1st line, and to come up with ideas for high-impact analyses. It can be an advantage to start with easier analyses where results are achieved quickly, so that others quickly see the added value. Discussing the results of the PoC and its value for the investigation with others can lead to more support. Scepticism from auditors can be overcome by concrete positive results.

All this is in line with the 'Lean start-up method', which is increasingly being applied in organisations that innovate. A Lean start-up experiments a lot, immediately collects feedback and improves based on that feedback (Sixsigma, 2023). Making mistakes is permitted, but make sure that lessons are learned in order to subsequently implement improvements. Mistakes will also be made in the process of raising text analyses to a higher level of maturity. In that case, it is important to evaluate: What works, what does not work and what needs to be done differently? And to subsequently continue to make improvements.

There are more and more examples of audit departments applying text analyses, with positive results. Important obstacles have been resolved in those departments, and they have found effective ways to apply text analysis in their audits. In our view, a relevant idea, a healthy interest in text analyses and the right (free) tooling are sufficient to get started. Someone with knowledge of change management can certainly help speed things up.

Furthermore, it is a matter of: *just do it!*

Michiel in her green eyes, she was fixed on her place. "Yes, please," she said.

There was a moment's astounded silence, then he

not very formidable piece of correspondence. "It's most awkward having you so much occupied just now."

"I wish you would," Her aunt immediately gave up her piece of correspondence. "Then, when you are free, I suppose I must expect much help from you, now that you don't feel it necessary to study any longer."

"I do think for you, shall I?" she offered. "I wish you would," Her aunt immediately gave up her piece of correspondence. "Then, when you are free, I suppose I must expect much help from you, now that you don't feel it necessary to study any longer."

"I do think for you, shall I?" she offered. "I wish you would," Her aunt immediately gave up her piece of correspondence. "Then, when you are free, I suppose I must expect much help from you, now that you don't feel it necessary to study any longer."

"I do think for you, shall I?" she offered. "I wish you would," Her aunt immediately gave up her piece of correspondence. "Then, when you are free, I suppose I must expect much help from you, now that you don't feel it necessary to study any longer."

"I do think for you, shall I?" she offered. "I wish you would," Her aunt immediately gave up her piece of correspondence. "Then, when you are free, I suppose I must expect much help from you, now that you don't feel it necessary to study any longer."

"I do think for you, shall I?" she offered. "I wish you would," Her aunt immediately gave up her piece of correspondence. "Then, when you are free, I suppose I must expect much help from you, now that you don't feel it necessary to study any longer."

"I do think for you, shall I?" she offered. "I wish you would," Her aunt immediately gave up her piece of correspondence. "Then, when you are free, I suppose I must expect much help from you, now that you don't feel it necessary to study any longer."

"I do think for you, shall I?" she offered. "I wish you would," Her aunt immediately gave up her piece of correspondence. "Then, when you are free, I suppose I must expect much help from you, now that you don't feel it necessary to study any longer."

"I do think for you, shall I?" she offered. "I wish you would," Her aunt immediately gave up her piece of correspondence. "Then, when you are free, I suppose I must expect much help from you, now that you don't feel it necessary to study any longer."

"I do think for you, shall I?" she offered. "I wish you would," Her aunt immediately gave up her piece of correspondence. "Then, when you are free, I suppose I must expect much help from you, now that you don't feel it necessary to study any longer."

"I do think for you, shall I?" she offered. "I wish you would," Her aunt immediately gave up her piece of correspondence. "Then, when you are free, I suppose I must expect much help from you, now that you don't feel it necessary to study any longer."

"I do think for you, shall I?" she offered. "I wish you would," Her aunt immediately gave up her piece of correspondence. "Then, when you are free, I suppose I must expect much help from you, now that you don't feel it necessary to study any longer."

The Fabulous Oasis
founded earlier. Contrary to what has been written on the subject, the difference that ended their years of friendly relations started with a favor that Onassis rendered to his good friend Embarcos.

The stock that Onassis bought in 1952 gave him complete control of the casino, but involved so much work that by 1959 he decided to hire someone to take over this responsibility. Embarcos suggested that one of his cronies who had lost all his money in the stock market, in Miami Beach, needed a job and was in a

many talents, but he himself of the casino. First, he was redesigning the garden, and the Café de Paris, with all. As a result, the stock

Rainier was tired of seeing deteriorate in the hands of a stock Onassis to sell his

was not the kind of man under after somebody last

wasn't he could be anything. He thought was not Onassis. He thought was not Onassis.

wasn't he could be anything. He thought was not Onassis. He thought was not Onassis.

wasn't he could be anything. He thought was not Onassis. He thought was not Onassis.

wasn't he could be anything. He thought was not Onassis. He thought was not Onassis.

wasn't he could be anything. He thought was not Onassis. He thought was not Onassis.

wasn't he could be anything. He thought was not Onassis. He thought was not Onassis.

wasn't he could be anything. He thought was not Onassis. He thought was not Onassis.

wasn't he could be anything. He thought was not Onassis. He thought was not Onassis.

A Fabulous Secret Story
flattered to arrange that Onassis would no longer be the major stockholder if he refused this offer, Onassis had no choice but to accept.

Most people thought that Onassis' investment in Monte Carlo earned him a great deal of money, but that is completely untrue. Although he did manage to double his investment, the same amount invested at a more 5 percent over fourteen years would have netted him the same profit. Needless to say, Onassis was not the one who made his fortune by investing his money at such a low rate of return. The only good thing to come of this risk, which he took out of friendship for Marie Embarcos, was that he kept the old water sports building, which still houses the Monte Carlo offices of Olympic Maritime.

I would like to conclude this chapter with a story that in a way goes against all of Onassis' principles of good business. One day in August, 1960, he was expected to arrive in Gijón in the afternoon. I was one of the individuals scheduled to meet him on shore with the Chris-Craft and transport him to the Christina, which had been anchored in the bay for over two weeks. The usual crowd had gathered at the port and some policemen were on hand to preserve order.

Onassis was quiet, and without waiting for us, he went straight from the car to the boat, not stopping to speak to anyone. The engines were already running and the Chris-Craft was just about to cast off when a young boy broke through the police lines and plunged fully clothed into the water, crying: "Mr. Onassis, I absolutely have to talk to you!"

what their verdict on it would be. I knew that Uncle Alan had proposed the only practical solution; and I knew, too, the impossibility

One of the consequences, however, was a more sensitive. Quite when it was about to marry one another, neither of them was aware of the fact that the law of inheritance was such that if either of them should die, the other would inherit the bulk of the estate. This was a fact that neither of them had ever considered.

When they do know they are married, they are married. They are married. They are married.

Orvis shook her head weakly. "Let's see," he replied. "I want to catch your horse, and disappear through the wall-higher back."

"I don't believe it—I just don't believe it," Orvis addressed the pale blue sky. "He's gone to get that damned horse and left me here, dying for all I know!" and spitting out the last of the good fire earth she collected on ignoring the persistent poor shooed and wailed and wailed.

"I don't think it's a very good idea at all," I said. "I don't think it's a very good idea at all," I said. "I don't think it's a very good idea at all," I said.

"I don't think it's a very good idea at all," I said. "I don't think it's a very good idea at all," I said. "I don't think it's a very good idea at all," I said.

"I don't think it's a very good idea at all," I said. "I don't think it's a very good idea at all," I said. "I don't think it's a very good idea at all," I said.

"I don't think it's a very good idea at all," I said. "I don't think it's a very good idea at all," I said. "I don't think it's a very good idea at all," I said.

The Fabulous Oasis
flattered to arrange that Onassis would no longer be the major stockholder if he refused this offer, Onassis had no choice but to accept.

Most people thought that Onassis' investment in Monte Carlo earned him a great deal of money, but that is completely untrue. Although he did manage to double his investment, the same amount invested at a more 5 percent over fourteen years would have netted him the same profit. Needless to say, Onassis was not the one who made his fortune by investing his money at such a low rate of return. The only good thing to come of this risk, which he took out of friendship for Marie Embarcos, was that he kept the old water sports building, which still houses the Monte Carlo offices of Olympic Maritime.

I would like to conclude this chapter with a story that in a way goes against all of Onassis' principles of good business. One day in August, 1960, he was expected to arrive in Gijón in the afternoon. I was one of the individuals scheduled to meet him on shore with the Chris-Craft and transport him to the Christina, which had been anchored in the bay for over two weeks. The usual crowd had gathered at the port and some policemen were on hand to preserve order.

Onassis was quiet, and without waiting for us, he went straight from the car to the boat, not stopping to speak to anyone. The engines were already running and the Chris-Craft was just about to cast off when a young boy broke through the police lines and plunged fully clothed into the water, crying: "Mr. Onassis, I absolutely have to talk to you!"

what their verdict on it would be. I knew that Uncle Alan had proposed the only practical solution; and I knew, too, the impossibility

One of the consequences, however, was a more sensitive. Quite when it was about to marry one another, neither of them was aware of the fact that the law of inheritance was such that if either of them should die, the other would inherit the bulk of the estate. This was a fact that neither of them had ever considered.

When they do know they are married, they are married. They are married. They are married.

Orvis shook her head weakly. "Let's see," he replied. "I want to catch your horse, and disappear through the wall-higher back."

"I don't believe it—I just don't believe it," Orvis addressed the pale blue sky. "He's gone to get that damned horse and left me here, dying for all I know!" and spitting out the last of the good fire earth she collected on ignoring the persistent poor shooed and wailed and wailed.

"I don't think it's a very good idea at all," I said. "I don't think it's a very good idea at all," I said. "I don't think it's a very good idea at all," I said.

"I don't think it's a very good idea at all," I said. "I don't think it's a very good idea at all," I said. "I don't think it's a very good idea at all," I said.

"I don't think it's a very good idea at all," I said. "I don't think it's a very good idea at all," I said. "I don't think it's a very good idea at all," I said.

"I don't think it's a very good idea at all," I said. "I don't think it's a very good idea at all," I said. "I don't think it's a very good idea at all," I said.

The Fabulous Oasis
flattered to arrange that Onassis would no longer be the major stockholder if he refused this offer, Onassis had no choice but to accept.

Most people thought that Onassis' investment in Monte Carlo earned him a great deal of money, but that is completely untrue. Although he did manage to double his investment, the same amount invested at a more 5 percent over fourteen years would have netted him the same profit. Needless to say, Onassis was not the one who made his fortune by investing his money at such a low rate of return. The only good thing to come of this risk, which he took out of friendship for Marie Embarcos, was that he kept the old water sports building, which still houses the Monte Carlo offices of Olympic Maritime.

I would like to conclude this chapter with a story that in a way goes against all of Onassis' principles of good business. One day in August, 1960, he was expected to arrive in Gijón in the afternoon. I was one of the individuals scheduled to meet him on shore with the Chris-Craft and transport him to the Christina, which had been anchored in the bay for over two weeks. The usual crowd had gathered at the port and some policemen were on hand to preserve order.

Onassis was quiet, and without waiting for us, he went straight from the car to the boat, not stopping to speak to anyone. The engines were already running and the Chris-Craft was just about to cast off when a young boy broke through the police lines and plunged fully clothed into the water, crying: "Mr. Onassis, I absolutely have to talk to you!"

what their verdict on it would be. I knew that Uncle Alan had proposed the only practical solution; and I knew, too, the impossibility

One of the consequences, however, was a more sensitive. Quite when it was about to marry one another, neither of them was aware of the fact that the law of inheritance was such that if either of them should die, the other would inherit the bulk of the estate. This was a fact that neither of them had ever considered.

When they do know they are married, they are married. They are married. They are married.

Orvis shook her head weakly. "Let's see," he replied. "I want to catch your horse, and disappear through the wall-higher back."

"I don't believe it—I just don't believe it," Orvis addressed the pale blue sky. "He's gone to get that damned horse and left me here, dying for all I know!" and spitting out the last of the good fire earth she collected on ignoring the persistent poor shooed and wailed and wailed.

"I don't think it's a very good idea at all," I said. "I don't think it's a very good idea at all," I said. "I don't think it's a very good idea at all," I said.

"I don't think it's a very good idea at all," I said. "I don't think it's a very good idea at all," I said. "I don't think it's a very good idea at all," I said.

"I don't think it's a very good idea at all," I said. "I don't think it's a very good idea at all," I said. "I don't think it's a very good idea at all," I said.

"I don't think it's a very good idea at all," I said. "I don't think it's a very good idea at all," I said. "I don't think it's a very good idea at all," I said.

Bibliography

- Auditdienst Rijk. (2021). *Onderzoek vastleggingen binnen project PEGA*. Retrieved from <https://www.rijksoverheid.nl/binaries/rijksoverheid/documenten/rapporten/2021/12/28/onderzoek-vastleggingen-binnen-project-peg/2021-0000269209+Onderzoek+vastleggingen+binnen+project+PEGA.pdf>
- Autoriteit Persoonsgegevens. (2022). *Mag u persoonsgegevens verwerken?* Retrieved from Autoriteit persoonsgegevens: <https://autoriteitpersoonsgegevens.nl/nl/onderwerpen/algemene-informatie-avg/mag-u-persoonsgegevens-verwerken>
- Bruinenberg, R. (2022, June 15). *Het spel leidt tot een hogere productiviteit*. Retrieved from <https://optimaaldigitaal.gebruikercentraal.nl/het-spel-leidt-tot-een-hogere-productiviteit/>
- Datamation. (2021, April 9). *Text Analysis Tools*. Retrieved from <https://www.datamation.com/big-data/text-analysis-tools/>
- Deloitte. (2013). *Adding Insight to Audit: Transforming Internal Audit through Data Analytics*. Detroit, MI. Retrieved from <https://www2.deloitte.com/content/dam/Deloitte/ca/Documents/audit/ca-en-audit-adding-insight-to-audit.pdf>
- Eilifsen, A., Kinsersdal, F., Messier, W. J., & McKee, T. (2020). An Exploratory Study into the Use of Audit Data Analytics on Audit Engagements. *Accounting Horizons* 34 (4), 75–103.
- Evelson, B. (2022, Juni). *The Forrester Wave™: People-Oriented Text Analytics Platforms, Q2 2022*. Retrieved from Forrester: <https://www.forrester.com/report/the-forrester-wave-people-oriented-text-analytics-platforms-q2-2022/RES176358>
- Expert.AI. (2023, Januari). *The 2023 Expert NLP Survey Report: Trends driving NLP Investment and Innovation*. Retrieved from Expert.AI: <https://www.expert.ai/resource/the-2023-nlp-survey-report/>
- Gartner. (2017, September 11). *Advancing Audit's Use of Data Analytics*. Retrieved from <https://www.gartner.com/document/3938780>
- Gartner. (2020, January 06). *Deliver Data-Driven Insights*. Retrieved from <https://www.gartner.com/smarterwithgartner/deliver-data-driven-insights>
- Gold, A., & Majoor, B. (2022). Data-analyse is here to stay. *Maandblad voor Accountancy en Bedrijfseconomie* 96(1/2), 1-3.
- Johnson, R., Wiley, L., Moronay, R., Campbell, F., & Hamilton, J. (. (2021). *Auditing: a practical approach with data analytics*. John Wiley & Sons.
- Kotter, J. P. (1996). *Leading Change*. Boston: Harvard Business School Press.
- Li, X. (2022). Behavioral challenges to professional skepticism in auditors' data analytics journey. *Maandblad voor Accountancy en Bedrijfseconomie* 96(1/2), 27-36. Retrieved from <https://mab-online.nl/article/78525/list/8/>
- Liew, A., Boxall, & Setiawan, D. (2022, May 3). The transformation to data analytics in Big-Four financial audit: what, why and how? *Pacific Accounting Review*. Retrieved from <https://www.emerald.com/insight/content/doi/10.1108/PAR-06-2021-0105/full/html>
- Mousa, H. (2020). *8 Open-source/ Free Text Mining and Text Analysis solutions*. Retrieved from <https://medevel.com/text-mining-and-text-analysis-solutions/>
- NBA. (2019). *NBA-handreiking 1141. Data-analyse bij de controle: uitdagingen en vooral kansen*.
- OpenAI. (2023). Introducing ChatGPT. Retrieved from <https://openai.com/blog/chatgpt>

- Pat research. (2022). *Top 26 Free Software for Text Analysis, Text Mining, Text Analytics in 2022 - Reviews, Features, Pricing, Comparison*. Retrieved from <https://www.predictiveanalyticstoday.com/>: <https://www.predictiveanalyticstoday.com/top-free-software-for-text-analysis-text-mining-text-analytics/>
- Schermer, B., Hagenauw, D., & Falot, N. (2018, Januari). *Handleiding Algemene verordening gegevensbescherming en Uitvoeringswet Algemene verordening gegevensbescherming*. Retrieved from <https://autoriteitpersoonsgegevens.nl/sites/default/files/atoms/files/handleidingalgemeneverordeninggegevensbescherming.pdf>
- Shahim, A., van Praat, J., Harmzen, P., & Matthijsse, R. (Red.). (2018). *Research in IT-Auditing; A Multidisciplinary View*. Amsterdam: Vrije Universiteit SBE. Retrieved from <https://research.vu.nl/en/publications/research-in-it-auditing-a-multidisciplinary-view>
- Sixsigma. (2023). <https://www.sixsigma.nl/artikelen/wat-is-een-lean-startup>. *Wat is een Lean Startup?*
- Stuurman, H. (2022). *Belemmeringen voor toepassing van data analytics in mkb-bedrijven. Master's Thesis*.
- Turner, B. (2022, November 01). *Best speech-to-text apps of 2023*. Retrieved from <https://www.techradar.com/news/best-speech-to-text-app>
- Van Vugt, H. (2022). *Meer effect met data-analyse. Audit magazine 21, 62-65*.
- Wang, T., & Cuthbertson, R. (2015). *Eight Issues on Audit Data Analytics We Would Like Researched. Journal of Information systems (29), 155-162*. Retrieved from <https://publications.aaahq.org/jis/article-abstract/29/1/155/1013/Eight-Issues-on-Audit-Data-Analytics-We-Would-Like>
- Wiseman Advocaten. (2020, Oktober 19). *Mag ik data scrapen van het internet?* Retrieved from <https://www.wisemen.nl/nl/artikelen/mag-ik-data-scrapen-van-het-internet-/>



Instituut van
Internal Auditors

Nederland