July 2023

EUROSAI
European Organisation of Supreme Audit Institutions

N° 7

Data science for investigations and audits

*By **Pierre Moscovici***
First president of the Freench court of Accounts

Editorial

EUROSAI Innovations is a newsletter dedicated to innovative audit approaches, new ways of engaging with stakeholders or new management concepts. It is published by the Project Group "Innovations in EUROSAI" and is part of EUROSAI's strategic goal 1 (SG1).

This July 2023 issue - the first edited by the French *Cour des comptes* since France took over this project - is focused on **data science for investigations and audits**.

Aside from the frenzy over ChatGPT, algorithms are more and more often used by SAIs to help them to analyse huge amounts of data collected. This can guide investigations, help test hypotheses and gain a better understanding of how the various stakeholders affected by a public policy may react to that policy or to policy reforms.

For example, thanks to data science, the French *Cour des comptes* has been able to **better understand how disabled people navigate through administrative procedures and the residency path**, and adapted an algorithm to confirm and objectify the audit team's intuitions.

In another example, the Danish *Rigsrevisionen* used **logistic regression to calculate and quantify the errors in the invoicing of general practitioners** and to demonstrate the efficiency of the new scanning system.

And in the United Kingdom, the National audit Office used **innovative approaches to audit the models** used by the audited entities and built **a Framework to Review Models** aimed at people commissioning, carrying out or assuring analysis.

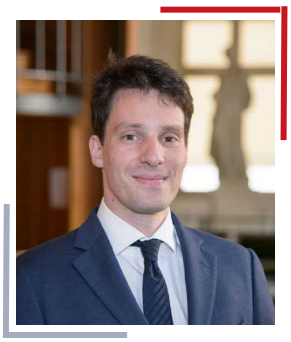
Although data science is increasingly commonly used, **40% of SAIs do not yet have dedicated staff to conduct data analysis, while 35% lack a department or staff responsible for data management**, according to the survey led in 2022 by the EUROSAI Project Group "*Auditing the New Normal*". The Turkish Court of Accounts, who leads this Project Group, provides an article on this Survey and how SAIs use data science in their audits.

Collecting, assessing and analysing the data is crucial, but reporting and data visualization is essential for proper appropriation of the analysis. The European Court of Auditors created an **interactive dashboard for animal transport that gives our readers a very good example of data visualization**.

Visualization is also a key role within the Austrian SAI and its **"R-Community" build to spread and foster knowledge on data science** and to provide tools and visualization solutions for the auditors.

Finally, you will discover how the German SAI, the *Bundesrechnungshof*, has organized its data team and read with interest the very informative **whitepaper on "*auditing algorithms*"**, drafted along with 4 other SAIs. This whitepaper comes with a helper tool to aid auditors with data science.

Whether you are an auditor, a data scientist, or a senior manager of your SAI, I hope you will enjoy the reading of the new EUROSAI Innovations newsletter and find it helpful and practical!



*By **Robin Kreling***
Data Scientist in the Data Analysis and Science Department

Understand how disabled people navigate through the residency path with optimal matching

The number of people over 50 receiving disability allowances increased by 36% in France between 2011 and 2019. The Court of Accounts wanted wanted to check whether the needs of the populations concerned were being adequately met.

To this end, the optimal matching technique has been used, consisting of showing similarities in the succession of events and thus possible causality between them, by adapting an algorithm derived from genetics, to verify possible breaks in the residence and administrative path, depending on the different situations encountered (whether or not people have access to specialized care or recognition of disability, and whether or not they receive care at home, etc.).

The data scientists' team was particularly inspired by the approach used by the Toulouse University Hospital for a previous survey. This highlights **the value of having a memory of past work and continuity of activity in the team, to capitalise on methods, practices and innovations**. The effort to adjust the algorithm to the data used by the Court in this investigation consisted of choices of calculation method intended to limit the calculation time.

The data was extracted from the digital service platform ViaTrajectoire, which connects people with institutions and helps them manage queues. The available information indicated whether people had open rights for dedicated medical or social care or not. The data was pseudonymised, with different encryption keys for each department.

It should be noted that, as a result of this initial work, the **Court is considering to match this data, with other administrative data for other surveys**, still anonymously, thus capitalising on the knowledge gained from these databases.

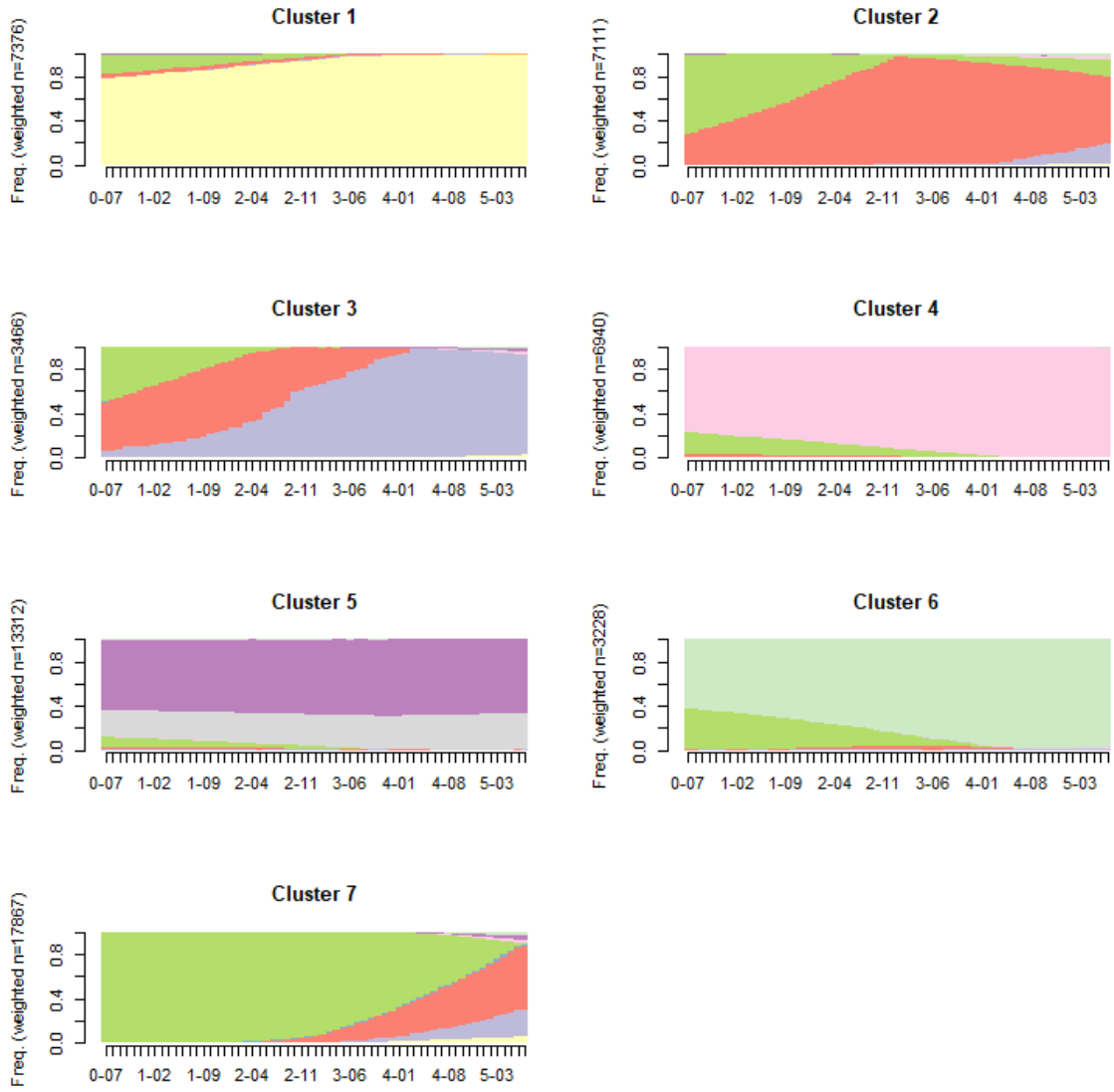
An algorithm to confirm and objectify the audit team's intuitions

The optimal matching technique applied to these data consisted of defining a similarity metric between the sequences, i.e. calculating a number giving an indication of the distance between two data sequences: while many changes are needed to transform one given sequence into another, they are considered very dissimilar and remote. If they require little or no changes, they are very close. This metric is then used to group the sequences into proximity clusters.

The grouping in clusters of typology of administrative and residential paths confirmed the intuitions of investigators and refined their understanding of the administrative, care and individual residential paths. For example, 12% of individuals in the 45- to 50-year-old sample are grouped in cluster 2. They are persons who have an administrative recognition of a disability but who do not submit any known application to an institution for a long period of time: these may include persons who refuse the recommended guidance or conduct the recognition procedure as a precautionary measure for future needs for formal assistance. Knowing and quantifying the existence of such precautionary approaches is useful for calculating indicators of the tension of accommodation solutions.

Collaboration between the audit team and data scientists is essential for the success of the audit

Exchanges between the control team and the data scientists of the Data Science and Analysis Department of the Court took place as soon as the feasibility note was drawn up, prior to the initiation of the investigation. This eased collaboration and facilitated the operation of the bases during the investigation. In particular, the audit team had identified very well the databases useful for the investigation. Its exchanges with data scientists were weekly and made it possible to produce indicators requested by the auditors, to point out unanticipated situations (like preponderance of some less visible handicaps) and highlight shortcomings in administrative databases.



Reading graphs: in X-axis, the months and years (sequences of 5 years complete, from 0-01 to 5-12); in Y-axis: the proportion of cluster observations, which ranges from 0 to 1. Each cluster has a different size (the size is indicated by "Freq. (weighted n=[number]". For example, cluster 6, we see that nearly 60% of the 3,228 people in this cluster do not experience any changes in their situation over a period of 5 years.

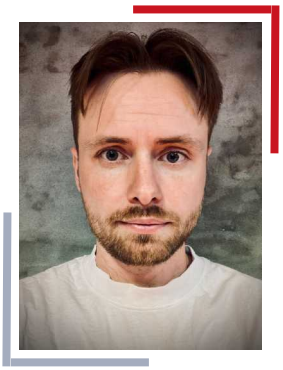
In this respect, it is important to stress that **the audit of the data itself contributes to the audit of the steering of the public policy** in question: these shortcomings, identified by the data scientists, gave rise to explicit recommendations in the report, which should be published by the end of July 2023.

The resulting database is indeed very recent, but will be useful for future investigations into disability and dependency. Over time, its historical depth will increase and represent longer and more representative sequences of a life course. It will also allow causal analysis of the effects of future reforms of public policies on autonomy and inclusion of people with disabilities.

To go further, you can contact the Court Data Science and Analysis Department.

Robin Kreling, Data Scientist in the Data Analysis and Science Department

Cour des comptes
FRANCE



*By **Christoffer Mejer Nielsen***
Auditor at the Rigsrevisionen (National Audit Office of Denmark)

Tracing the pattern of GP invoices

Every year, Danish general practitioners (GPs) send 60 million invoices, worth EUR 1,2 billion, to the five danish regions that are responsible for a wide range of healthcare services in Denmark, including services provided by GPs. Although the vast majority of the invoices are error-free, intentional or unintentional mistakes sometimes creep in, and all invoices are, therefore, checked before payment is transferred to the GP.

All invoices are electronically screened for errors before payment is made. If, for example, an incorrect price is charged for a service, or the patient's social security number is not found, the invoice will be rejected by the system. However, not all errors are caught in the electronic screening, and the case managers, therefore, also carry out manual checks, which involve **assessing the potential risk and significance of errors and prioritizing the checks accordingly**.

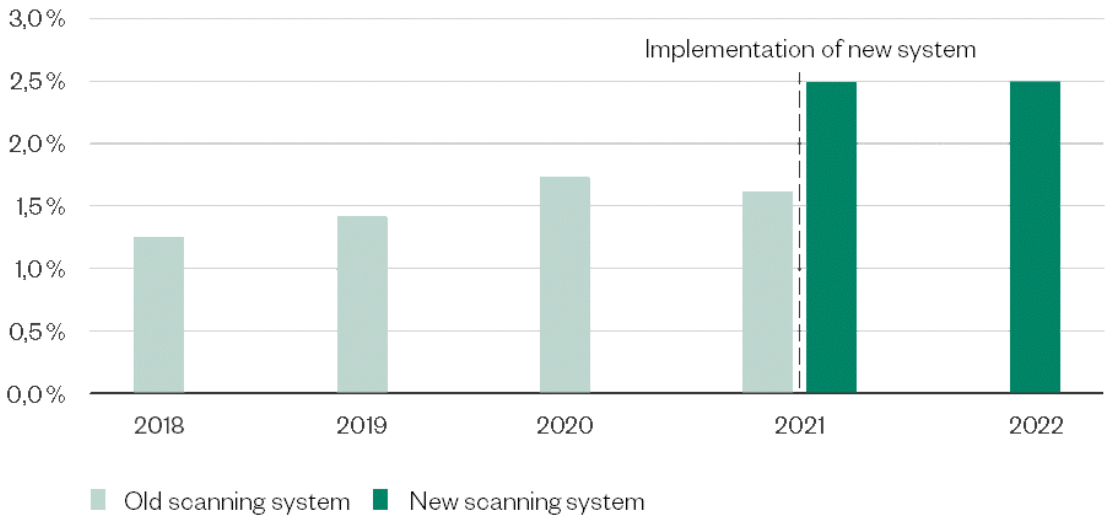
Our audit of the control of the fee calculation revealed a number of sources of error that remain undetected through the scanning system; some because the scanning system lacks the required functionality, and others because the control required a manual estimate.

Here **we used the patterns in the millions of bill data to uncover the weaknesses in both electronical and manual control**. In this article, we provide insight into how we used total data from payment systems to review the effectiveness of the regions' controls.

What we did

Firstly, we wanted to investigate whether the invoices were affected by errors that the scanning system failed to detect. To determine this, we extracted all GP invoices submitted in the period 2018-2022, which came to a total of 270 million invoices.

The quantity of rejected invoices across scanning systems (2018-2022)



Source: Rigsrevisionen using invoice data provided by the regions

All the invoices included data on the GP's service number, anonymized social security numbers, services invoiced, when the services were provided, whether the invoice had been settled or rejected and much more.

We then reviewed the contracts between the GPs and the regions to determine the agreed requirements for remuneration of health services. For example, a GP is not allowed to charge for two lung examinations of the same patient, or for more than one follow-up home visit per patient per year.

In total, we identified 11 essential requirements. To test whether the GPs were paid for services provided in conflict with the requirements, we coded more than 200 validation rules and applied them to the dataset. For instance, we recorded each time a GP had received payment for more than one follow-up home visit to the same patient within a calendar year.

The analysis showed that during the investigation period, payments were made in conflict with all but one of the mandatory requirements. In this analysis, **the advantage of using total data was obvious, as it allowed us to say with certainty that a payment was in breach of the contract's requirements and calculate the exact number of affected payments**.

Secondly, we wanted to look into the effectiveness of a new scanning system which the regions had implemented during the period under investigation, and we wanted to determine whether this new system was more sophisticated and effective than the old one. Specifically, we used total data from the old and the new scanning system to test whether the transition had significantly affected the number of rejected invoices.

We used logistic regression based on 95% confidence intervals ($p<0.05$) and analyzed the numbers for all the regions combined and for each region separately. **This was the first time we used logistic regression analysis on a data set of such size**. Coding was done in 'R'.

The analysis showed that the new scanning system detected significantly more errors and rejected 2,5% of invoices, worth approx. EUR 19,9 million annually, against 1,5%, worth approx. EUR 13,2 million with the old system. In other words, the new system now rejects invoices that previously probably would have been settled, worth approx. EUR 6,7 million annually.

Using total data also for this analysis turned out to be beneficial. Contrary to a sample-based approach, it **allowed us to demonstrate, with a high degree of certainty, a significant effect and to calculate the effect** based on the number of rejected invoices.

The study also pointed to interesting variations across the regions; some of them had discov-ered systematic sources of error in the scanning system early on and had initiated compensa-tory checks. Another region carried out compensatory checks of certain types of consulta-tions and vaccination fees that the GPs are not allowed to charge in combination with certain other services.

Our analysis of the payment pattern for these particular combinations of services showed that the controls had a significant effect and led to the rejection of 83% of the invoices in question. In another region, that did not carry out compensatory checks, the scanning system rejected 0.6% of the invoices that covered the same services.

The data analysis allowed us to verify whether compensatory checks, which should theoretically lead to more rejected invoices, also did so in practice.

Based on our findings, we listed a number of recommendations on how the regions could use the data available to them to target their control of GP invoices further. (Read an English version of an extract of the report)

If you want more information on the report and the methods applied please contact me.

***Christoffer Mejer Nielsen**, auditor at the Rigsrevisionen (National Audit Office of Denmark)*

FOLKETINGET
RIGSREVISIONEN

*By **Ruth Kelly***
Chief Analyst at the United Kingdom National Audit Office

What makes a super model? Innovative approaches to audit models

The United Kingdom government relies on hundreds of models for activities like estimating costs, distributing funding within organisations, and testing policy options – and they underpin decisions that affect people's lives.

In recent years the UK government has used models to, for example, plan the COVID-19 pandemic test and trace service, identify how many new teachers it needs to train, and estimate the cost of the financial settlement when leaving the EU. So, it's really important that people who depend on outputs from models can feel confident in the quality and robustness of these models.

How the UK National Audit Office (NAO) uses models

At the UK NAO, part of our financial audit work involves **scrutinising the models that underpin significant estimates in departments' accounts**. My expert Modelling team looked for innovative ways to do this, and to support government departments in improving the way they produce and use models.

The complexity of micro simulation models makes traditional approaches to auditing models challenging and amplifies potential errors that can be easy to miss.

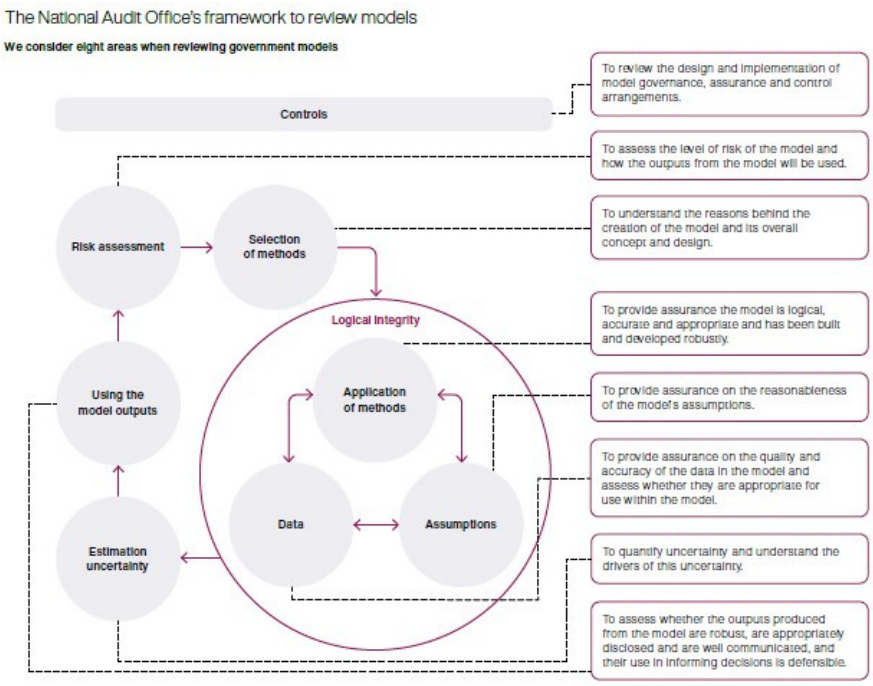
To help us audit the estimate, we built an independent reproduction of the model in the R software language. Running the reproduction separately allows us to **produce an independent estimate and helps us to identify and investigate any discrepancies to the original model**. This has enhanced confidence in the outputs for key stakeholders.

How we managed uncertainty in auditing models

Modelled outputs are inherently uncertain. As well as checking that the central estimate is reasonable, we also wanted to understand the full range of plausible outcomes. We built in fully automated uncertainty analysis to our reproduction, which lets us stress test the estimate under extreme scenarios.

It also lets us test what happens to the estimate when several inputs change at the same time, by running thousands of simulations to generate a likely range of outcomes. This is something not carried out in many of the models we audit and is an area where our independent model assurance can provide additional value. It gives us confidence that the estimate will not be materially wrong, even when economic shocks are considered.

This fully working model reproduction has transformed the way we audit the estimate and is a great example of what is possible in terms of model quality assurance. It has enhanced the quality of our work: quality assurance checks are automated, including more advanced sensitivity analysis. And **it's helped us to be more efficient**: the quality assurance checks in the reproduction are quicker to produce, freeing up our analysts to focus on creating greater insights.



The NAO's framework to review models considers 8 areas.

What next?

We think there are opportunities to replicate this approach across the portfolio of models that we audit and help enhance our quality assurance work. We want our audit work to help build confidence in the quality of government's models and support government in making plans that don't place value for money at risk.

Our report on Financial Modelling in government looks at how government produces and uses models and identifies the systemic issues that can lead to value for money risks.

To find out more about the way we audit models, see our Framework to Review Models. This framework is aimed at people commissioning, carrying out or assuring analysis. It provides a structured approach to review models, which organisations can use to determine whether the modelling outputs they produce are reasonable and robust.

We have also developed a tailored version for audit and risk committees, to help audit committees know what to ask and what to look for when questioning management on the quality assurance of models and the handling of modelling risks.

Discover the Framework to review models

For any further information, please feel free to contact me.

Ruth Kelly, Chief Analyst at the UK National Audit Office

National Audit Office

By **Abdulkadir Karadağ**
Auditor at the Turkish Court of Accounts

Survey: how european SAIs use data science

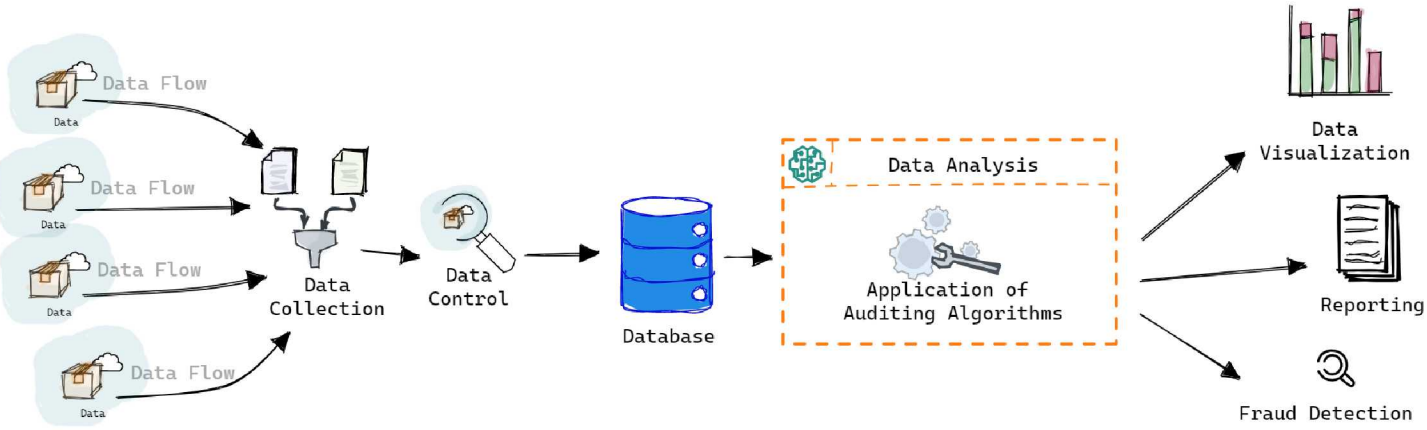
Rapid advancements in data science offer the chance to change and improve the methods used to audit an organization's operations, processes and accounts, and to detect errors, corruption, or irregularities.

This article aims to address and assess the results of the Survey Assessment Report conducted in 2022 by the EUROSAI Project Group "***Auditing in the New Normal: Connecting Technology to Audit Processes***" with the participation of 20 Supreme Audit Institutions (SAIs) with respect to the level and form of use of data science by the participating SAIs.

Use of Data Science Techniques in Audit and Investigation Processes

The use of data science for the purposes of audit and investigation involves a process that follows certain steps, as outlined in the following chart.

1. The Data Collection and Preparation phase consists of collecting the data obtained automatically or manually from various sources and making them ready for use. According to the survey results, almost all SAIs use e-mail for data collection.



60% of SAIs participating in the survey have access to the information systems of the auditees. Web service integration, direct database access and FTP (File Transfer Protocol) connection, on the other hand, are less common methods of data collection.

2. During the Data Control and Storage phase, the collected data is consolidated and stored after being examined in terms of quality and deficiencies. As per the survey results, **85% of SAIs review the data collected for audit purposes in terms of quality and completeness before using them**. All SAIs rate their overall data quality as either good or fair, while none rates it poor or excellent.

3. The Data Analysis and Modelling phase consists in detecting patterns, anomalies or relationships in the data by using machine learning algorithms, such as statistical analysis, cluster analysis, classification or regression models and data visualization techniques on the data set. While **85% of the SAIs** responding to the survey **rate their current level of use of data analysis tools and techniques as moderate or good**, compared to 10% who rate it as poor.

40% of SAIs do not have dedicated staff to conduct data analysis, while 35% lack a department or staff responsible for data management. 75% of SAIs use a standard tool to conduct analytical analysis on data, while 45% employ an advanced analysis tool.

4. During the Detection of Irregularities and Risk Assessment phase, potential irregularities or risky situations are identified, based on the findings of the data analysis, by using the previously determined irregularity detection patterns or risk assessment models. The results of the survey indicate that **SAIs carry out preliminary checks to ensure and maintain data quality and validity from the very beginning of the audit process**, confirm the accuracy and reliability of the data collected for this purpose, and conduct risk assessments.

5. Data Visualization and Reporting is the phase where the results obtained are visually presented and reported through data visualization tools. In this step, the data is made easier to understand through the use of charts, graphs or interactive visuals. Reports may include audit or investigation results and recommended actions. According to the results of the survey, **statistical analysis and visualization stand out as the two most frequently used techniques to process existing data and produce usable outputs**.

6. Finally, the Evaluation of Results and Improvement phase takes place, where audit or investigation teams carefully examine and confirm the results and reports. In this step, actions such as verifying the detected findings and irregularities, taking corrective measures or improving processes can be performed. In addition, the effectiveness and accuracy of the results obtained by the use of data science techniques are constantly evaluated with the aim of improvement. According to the survey results, **only 30% of SAIs evaluate the impact of the use of digital technologies on audit outputs**.

Survey results demonstrate that **all SAIs think that digital transformation directly affects the effectiveness and efficiency of audits and considers the use of digital technologies as a key strategy for enhancing SAI performance.** Also, they rate the importance of using new and/or existing digital technologies in audits as 8.55 out of 10. This positive approach demonstrates SAIs' readiness to integrate technology into their audit procedures.

In conclusion, data science is a valuable tool for using analytical power to improve information and detect errors in audits and investigations. According to the survey, **SAI staff's adoption of data science skills is rated between 5 and 7, with an average value of 5.5, on a scale from 1 to 10. The results show that the human resources capacity of the institutions is sufficient, yet there is still room for improvement**.

In this context, for a successful transformation process, the competencies of auditors should be developed, and a recruitment policy focusing on digital skills and abilities should be developed for personnel at all levels.

If you want more information on the report and the methods applied please contact me.

Abdulkadir Karadağ, auditor at the Turkish Court of Accounts

By **Lukasz Kolodziej**
Data and Technology for Audit – DATA Team

Harnessing data science for investigative audits: an interactive dashboard for animal transport

Data science has become an increasingly powerful tool for conducting audits and investigations. It enables auditors to achieve greater efficiencies, to better target audit work, and to identify patterns, trends, and anomalies. In this article, we will present the creation and benefits of an interactive online dashboard that was published with the European Court of Auditors' (ECA) 'Review 03/2023: Transport of live animals in the EU: challenges and opportunities'.

The **DATA team was created by the ECA in 2021**. It brings together data experts, IT auditors and project managers **to work as a single-entry point for audit teams when it comes to leveraging data as audit evidence and technology as audit object**. The review on animal transport was adopted by Chamber I "Sustainable use of natural resources", headed by ECA Member Joëlle Elvinger. The task was led by ECA Member Eva Lindström. The review describes the main factors surrounding the transport of live animals and illustrates the trends in animal transport.

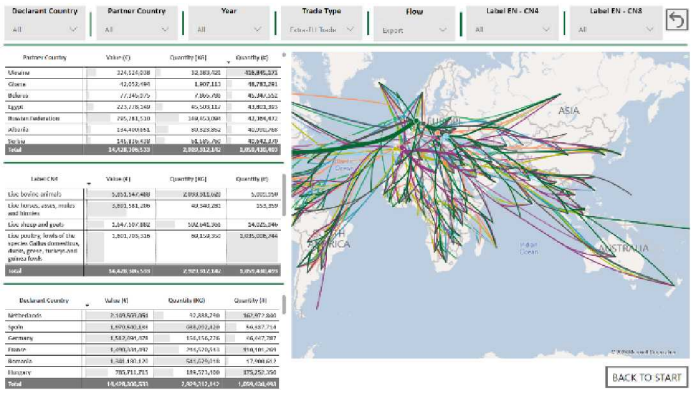
The interactive online dashboard was created to help visualise the data collected and analysed for the ECA report 'Transport of live animals in the EU: challenges and opportunities'.

During their work, the ECA auditors identified vast amounts of data scattered across various databases and websites. To make this information more accessible and easier to analyse, data scientists in the DATA team, with the support of auditors, developed an interactive online dashboard with facts and figures on the transport of cattle, pigs, poultry, and other animals.

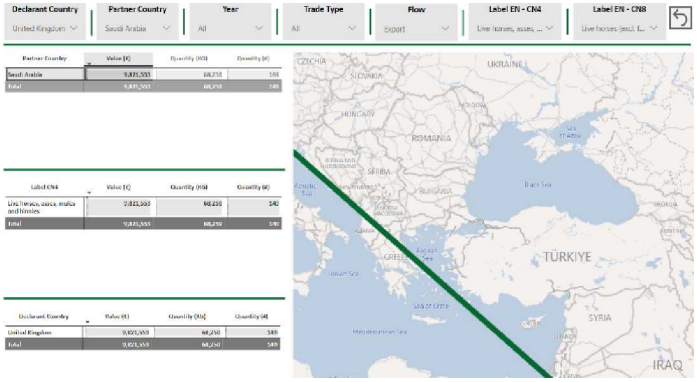
The **dashboard is a powerful research tool**. It compiles information from several sources and allows users to explore data on animal movements, such as the number of animals, their weight, and value of the trade for various animal categories. Users can also compare data on animal movements between different countries, animal categories, and economic values of movements. It also allows users to investigate the reasons behind animal transport, such as trade, fattening, or breeding.

This tool can uncover patterns and trends, providing valuable insights for auditors and the public alike. To facilitate understanding of the information and to allow users to draw meaningful conclusions, the DATA Team used visualisations such as graphs, charts, and maps, created with the Microsoft Power BI tool. Data shown in the dashboard was extracted in May 2022 from COMEXT, Eurostat's reference database for detailed statistics on international trade in goods.

Information compiled and displayed in the dashboard can sometimes be surprising. For instance, the dashboard revealed that between 2017 and 2021 the largest importer of live animals from EU countries by value was the US, which received exports with a value of €2.1 billion, whereas the largest importer from the EU by quantity was Ukraine (416 million animals).



The visualisation is made of data extracted from the COMEXT database.



The dashboard can help to identify the value of certain types of animal trades.

Another interesting fact that the dashboard allows to explore is the significant value of certain types of animal trades. For example, racehorses (i.e. live horses excluding those for slaughter, pure-bred for breeding) can be of high worth: between 2017 and 2021, the United Kingdom exported 149 racehorses to Saudi Arabia, for a total value of almost €10 million.

The dashboard also allows the comparison of main destinations of animal transport for two member states. For instance, users of the dashboard can compare the values of intra-EU exports, see which are the main partner countries, and the categories of animals traded. The comparator report also shows information for numbers of animals transported and their weight.

By harnessing the power of data science and creating this interactive dashboard, we have taken a significant step forward in **improving the accessibility and usability of audit data for investigations**. This tool not only benefitted the audit team but also empowers the readers of our reports to engage with the data and better understand the implications of animal transport in the EU.

As data science continues to advance, we can expect to see even more powerful tools and methodologies for auditing and investigating complex issues, ultimately leading to better-informed decision-making and more effective policies.

The dashboard allows comparisons between two state members.

If you want more information on this topic, please contact the DATA Team.

Lukasz Kolodziej, Data and Technology for Audit

CURIA RATIONUM

EUROPEAN COURT OF AUDITORS

*By **Andreas Mayrbaeurl***
Deputy head of Department for Information Technology, Digitalization, and Organization

Data Analysis in the Austrian Court of Audit The R-Community

The use of data analysis in the Austrian Court of Audit (ACA) has been constantly expanded in the recent years. In this context, the working group "R–Community" has provided an informal framework for the exchange of knowledge and information.

Like many other SAIs, the ACA has a privileged access to data. In contrast to the scientific and corporate sector, it has, in principle, the right to be given access to all data collected by public authorities. The initial objective of the R–Community was to fully harness the opportunities arising therefrom and to provide greater support for the use of data analysis tools.

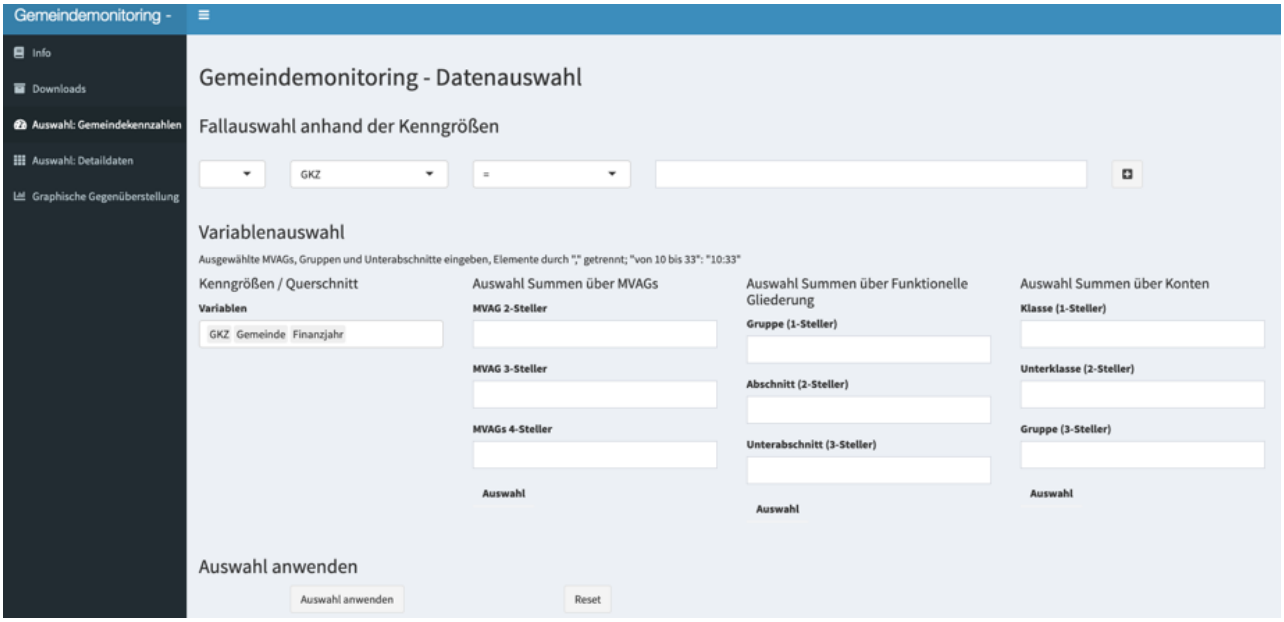
In 2004, when it was still in its infancy, the R–Community consisted of a small group of auditors with data analysis skills, who exchanged ideas about statistical methods and the statistics programme "R". Since then, the R–Community has served as a platform to facilitate networking. Its members meet bimonthly at least.

The R–Community team furthermore strives to render the topic of data analysis accessible to as many auditors as possible through in-house knowledge-sharing events.

In order to provide low-threshold learning opportunities for programming "R", the ACA also offers online courses whose curricula can be completed at an individual pace.

The overall goal is not to ensure that all staff members of the ACA will eventually be skilled in using "R", but rather to **generate sufficient knowledge about what data analysis can do and for what purposes / examinations it can be used.**

The field of data analysis has been greatly expanded in the recent years. "R" is now used in many audits to calculate more complex analyses or graphs, and it is also deployed for recurring tasks such as the monitoring of municipalities through the Municipality Monitoring Tool.



The Municipality Monitoring Tool of the ACA allows a risk-oriented audit selection.

With the Municipality Monitoring Tool, the ACA has access to the data presented in the financial statements of every Austrian municipality. Through the use of the application "Shiny", programmed with "R", the auditors can carry out extensive data evaluations.

In addition to fact sheets with key figures related to individual municipalities, tailor-made analyses can also be created interactively.

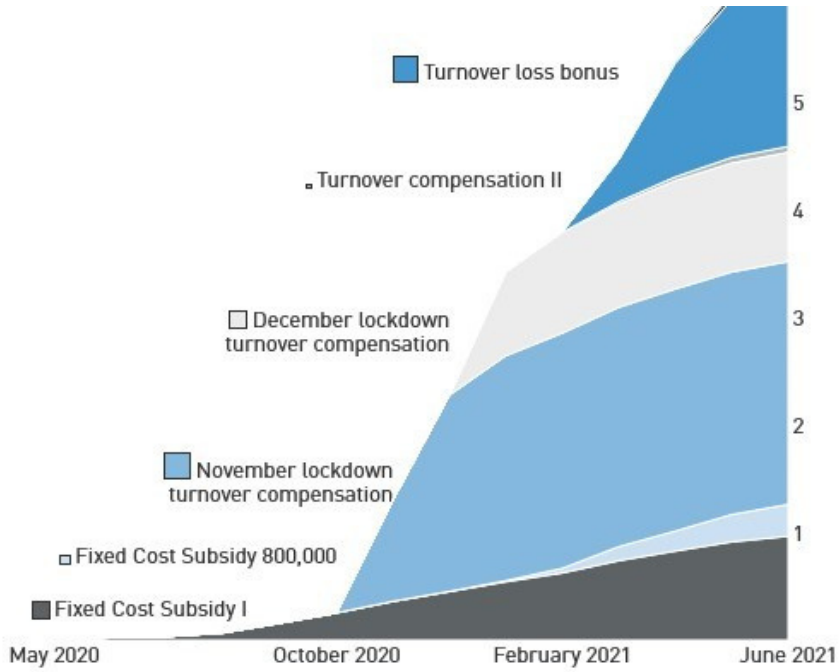
With this tool it is possible, for example, **to make a risk-oriented audit selection of municipalities or to create comparative statistics on the financial situation of municipalities**. The use of "R" has the great advantage that the entire programming – with the underlying know-how – lies within the ACA itself and can also be changed or extended almost at will.

The reporting on audit findings is another area that lends itself to data analysis. For example, in the report on the financial aid provided during the COVID-19 pandemic, data analysis methods were used to uncover and illustrate structural patterns. As for the visualization, a dynamic graphic has been created to ensure the best possible appropriation of the report's findings.

The successes of the past years and my personal experience with several scandinavian SAIs or the UK's show that **the use of data analysis will continue to play an increasingly important role** before, during and after audit activities (read the article from the turkisk SAI, page 9).

The ACA will continue to drive the use of modern data analysis methods with a strong team of data analysts and a broad range of support and training services. The R–Community is looking forward to being a part of this exciting task.

Read a summary of this report in English
or Listen to a podcast



Extract from the dynamic graph on the financial aid provided during the Covid-19 pandemic.

For more information on all these topics, please feel free to contact me.

Andreas Mayrbaeurl, MBA, Deputy Head of Department for Information Technology, Digitalization, and Organization

By Dr. ***Peter Koß***
Audit manager and project lead of the Bundesrechnungshof's data analytics project

Exploring tools and techniques to foster data analysis

Public authorities are striving for faster and more transparent administrative action. They are adapting their processes to growing digital interaction with citizens and businesses. To this end, they are not only digitising their existing processes, but are increasingly redesigning them to meet the requirements of a digital government. They generate and process digital products and use IT systems and applications for management and control.

In some areas, automated processes support human workers and replace them partially or completely. These developments increasingly impact on the *Bundesrechnungshof*'s core functions.

In response to these developments, the *Bundesrechnungshof*'s supreme decision-making body, as early as in March 2019, decided to adopt **an agile and structured approach to growing digitisation and emerging technologies encompassing all audit units**. The overall objective is to lay the groundwork for the "digital audit work of tomorrow", focusing on modern methods and technologies.

On the basis of the *Bundesrechnungshof*'s constitutional tasks of auditing, reporting and advising, the overall objective sets out the further strategic approach. The overarching purpose is to fulfil our core functions in a digital world – in the best way possible.

To this end, the *Bundesrechnungshof* started several activities, i. e. a data analytics project. The project explores the audit areas of artificial intelligence (AI), digitisation, data management and data analytics. The project members take an agile approach to identify potential areas of use and to test the benefits of new data analytics methods.

For this purpose, **we recruited multiple data analytics experts. These experts conducted data-driven pilot audits and applied new audit methods**. This approach enabled audit findings that cannot be reached by traditional means, or only at great expense.

In addition, the *Bundesrechnungshof* trained all members (audit directors and senior audit directors) in digitisation, data analytics and emerging technologies.

In 2022, the *Bundesrechnungshof* released **an internal audit guide for conducting audit work on emerging technologies**. This guide builds upon the efforts of five Supreme Audit Institutions (SAIs of Finland, Germany, Norway, the Netherlands and the United Kingdom) that joined forces under a Memorandum of Understanding on Data Analytics.

Schematic visualisation of data extraction from unstructured data

In 2020, they released their whitepaper on "auditing algorithms". The whitepaper comes with a helper tool to aid auditors. It follows a structured methodology and can be used by all auditors, regardless of their level of expertise in AI or IT.

Auditors can choose from a multitude of questions. Auditors with technical experience could choose to address AI-specific aspects such as model development and data engineering in their audits, while non-IT auditors might choose less technical topics such as business understanding or financial and legal aspects.

In January 2021, the paper was outlined in the International Journal of Government Auditing. In April 2023, the SAIs updated their whitepaper based on the lessons learned from their respective pilot audit work. The *Bundesrechnungshof* contributed to lessons learned from interdisciplinary audits of emerging technologies (cf. example 1, example 2) to this update.

For the pilot audits, we took a holistic approach and matched technical with substantive expertise. In doing so, we ensured that our findings and recommendations were in line with the subject matter context.

The data analytics experts within the *Bundesrechnungshof* support non-IT auditors in their daily work, for example by **extracting and transforming audit data from unstructured documents or databases**. We were able to successfully apply this technique in multiple use cases, such as audit work on travel cost reimbursement, or an interactive visualisation of planning data (budget, staff) of federal IT measures.

The *Bundesrechnungshof* is also exploring opportunities **to make a data analytics and multi-document search tool available to all auditors**. To this end, an Elasticsearch-driven database with an analytics backend that contains data and documents from a multitude of data sources was set up. With such a tool at their disposal, the *Bundesrechnungshof*'s staff should be able to quickly search thousands of documents of varying file type in a matter of seconds.

At the moment, the *Bundesrechnungshof* is examining further AI-driven use cases for the Elasticsearch-driven database, such as automated summarisation or "named entity recognition" to identify entities such as people or organisations within documents.

For more information on these topics and projects, please feel free to contact me.

Dr. ***Peter Koß****, audit manager and project lead of the Bundesrechnungshof's data analytics project.*

Bundesrechnungshof

This newsletter is published by the Project Group "Innovations in EUROSAI" and is part of EUROSAI's strategic goal 1 (SG1), which aims to **support effective, innovative and relevant audits by promoting and facilitating professional cooperation among European SAIs**.

Its first issue was released in December 2018 and the previous published issue dates back to December 2020.

It was originally driven by the German and Lithuanian SAIs, and the French SAI volunteered to take over in December 2022.

This issue is therefore the number 7, and is delivered to you in a **new layout** and with a **new editorial line**, which is to always offer readers **the most operational content possible**.

Three main goals are set, and every willing SAI will contribute to their achievements:

- Reinforce the concrete and practical dimension by including links and contacts to enable readers to extend their reading in an operational manner.
- Favour a short format, including an iconography.
- Diversify contributors to reflect the diversity of our SAIs.

A semi-annual publication is planned (Summer & Winter) and each issue will be focused on a single topic.

Our next issue will be focused on the following topic: "**engaging with society**".

Each SAI will soon receive a call for contribution with relevant details on how to join the team !

For any information regarding this newsletter, feel free to contact **Denis Gettliffe**, editor in chief of the EUROSAI Innovations newsletter.