MARCO SCHREYER
MARCEL BAUMGARTNER
T. FLEMMING RUUD
DAMIAN BORTH

ARTIFICIAL INTELLIGENCE IN INTERNAL AUDIT AS A CONTRIBUTION TO EFFECTIVE GOVERNANCE

Deep-learning enabled detection of anomalies in financial accounting data

The technological advances of Artificial Intelligence (AI) are increasingly perceived as a valuable tool for internal auditing. The following article is intended to highlight possible applications and challenges of Deep Learning, a comparatively young sub-discipline of AI, using a practical example from Nestlé S.A.*

1. INTERNAL AUDIT AND TECHNOLOGICAL ADVANCES

The recent developments in information technology, such as cloud computing, Artificial Intelligence (AI) and the Internet of Things (IoT), currently result in a large number of corporate efforts to gradually digitalise business processes. This transformation also affects Enterprise Resource Planning (ERP) systems and is changing the collection of audit evidence. Nowadays, ERP systems capture exhaustive volumes of audit relevant information, such as journal entries, process logs, or segregation of duties configurations. At the same time, the nature and scope of recorded corporate data enable the application of innovative digital auditing procedures, which promises to significantly increase the effectiveness and efficiency of internal audits.

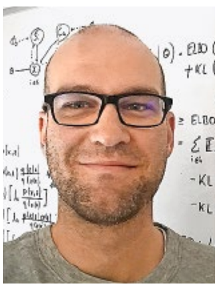
According to the definition of the Institute of Internal Auditors (IIA), internal audit provides "independent and objective assurance and consulting activity designed to add value and improve an organi[z]ation's operations" [1]. Furthermore, in exercising due professional care, auditors are required to consider the use of technology-based audit and other data analysis techniques [2]. In order to provide objective and comprehensive assurance in an increasingly digitalised economy, the audit profession is required to innovate and continuously improve its audit procedures [3]. As a result, auditors also invest in AI-enabled audit capabilities [4]. The use of AI-enabled audit techniques promises to contribute to objective and risk-oriented audit results [5].

2. ARTIFICIAL INTELLIGENCE, MACHINE LEARNING AND DEEP LEARNING

In general, AI aims to model the cognitive capabilities of human intelligence to accomplish intellectual tasks in a self-learning manner [6]. *Figure 1* shows an exemplary taxonomy of AI and associated sub-disciplines. A prominent sub-discipline of AI includes the field of machine learning (ML). The term ML refers to a methodology that provides computers with the ability to learn a model to solve a predefined task, i.e. without explicit human programming [7]. Such a learning characteristic describes the fact that the quality of the model might improve over time with new information or data [8]. In the course of such a learning setup, ML algorithms identify task-relevant correlations or patterns in the data to derive a desired solution.

In the context of traditional ML approaches, such as decision trees or Bayes classifiers, the learning success is determined by the amount of human expertise. Such knowledge is required to manually extract task-relevant features from the raw data (*Figure 2, top*). The extracted features are then utilised by ML algorithms to identify patterns or correlations and derive a solution to solve a predefined task (e.g. the classification of images). As a result, the manually extracted features directly determine the learning success and quality of the ML model.

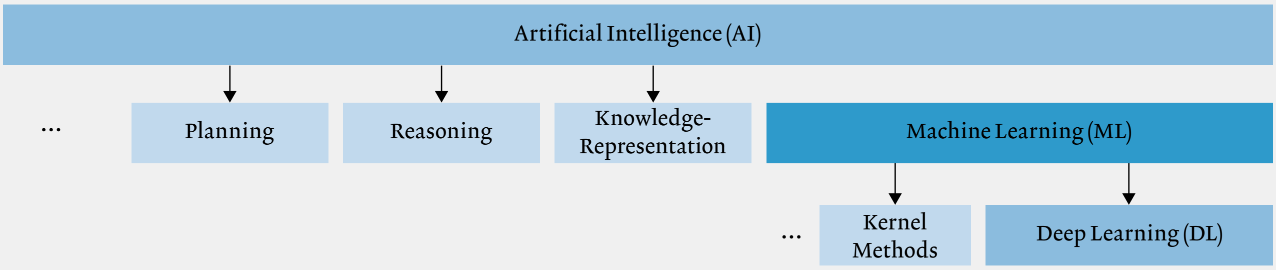
A successful sub-discipline of ML is referred to as Deep Learning (DL). In general, DL denotes a biologically inspired learning technique that is inspired by the structure and function of the human brain to imitate "intelligent" behaviour.



MARCO SCHREYER, RESEARCHER, CHAIR OF ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING, INSTITUTE OF COMPUTER SCIENCE (ICS), UNIVERSITY OF ST. GALLEN

MARCEL BAUMGARTNER, STATISTICIAN, DATA ANALYTICS EXPERT, NESTLÉ INTERNAL AUDIT, SOCIÉTÉ DES PRODUITS NESTLÉ S.A.

Figure 1: **SAMPLE TAXONOMY OF ARTIFICIAL INTELLIGENCE** [23]



Similar to the neurons of the human cortex, DL techniques comprise a series of interconnected artificial neurons [10]. The attribute "deep" refers to the high number of often up to several hundred layers of artificial neurons. The foundational ideas of DL are not new and originate from the early development of artificial neural networks [11] and artificial neurons [12] in the 1940s–1950s. The successful renaissance of deep autoencoder neural networks today can be attributed to:

→ the accessibility of exhaustive data volumes, e.g. image, speech and social media data;
→ the development of novel artificial neural network architectures, e.g. convolutional and long-short term memory neural networks;
→ the advances in high-performance computing, e.g. graphic processing units (GPUs).

In contrast to traditional ML approaches, DL techniques exhibit the ability to "autonomously" learn features from raw data to solve a predefined task (*Figure 2, bottom*). The manual extraction of relevant data features by humans becomes obsolete. The simultaneous learning of (i) relevant features and (ii) solutions for a given task is referred to as end-to-end learning [13]. In addition, when compared to the traditional hypothesis-driven analysis techniques used in internal auditing, DL methods allow for hypothesis-free unsupervised learning. In the context of auditing, the synthesis of both learning paradigms enables DL techniques to recognise patterns in exhaustive volumes of digital journal entries. Therefore, the target-oriented combination of both end-to-end and unsupervised learning represents a valuable addition to the toolbox of audit data analytics techniques [14].

An application of DL in internal auditing will be introduced below. The practical example outlines a DL-enabled technique to detect unusual journal entries, referred to as anomalies, in large-scale accounting data.

3. DEEP AUTOENCODER NEURAL NETWORKS IN THE INTERNAL AUDIT

In the context of internal auditing, it is generally assumed that erroneous or fraudulent activities represent exceptional cases that deviate from a company's usual behavioural pattern. Such deviating activities are recorded by a small fraction of journal entries that exhibit "anomalous" attribute values, e.g. by deviating vendor bank master data or posting times. Furthermore, when examining the journal entries recorded in ERP systems the following attribute characteristic can be observed [15]:

→ The extensive number of distinct journal entry attribute values, e.g. due to a large number of posted vendors, general ledger accounts or document types.
→ The extensive number of distinct journal entry attribute value correlations, e.g. a document type usually posted in combination with a particular posting key and general ledger account.

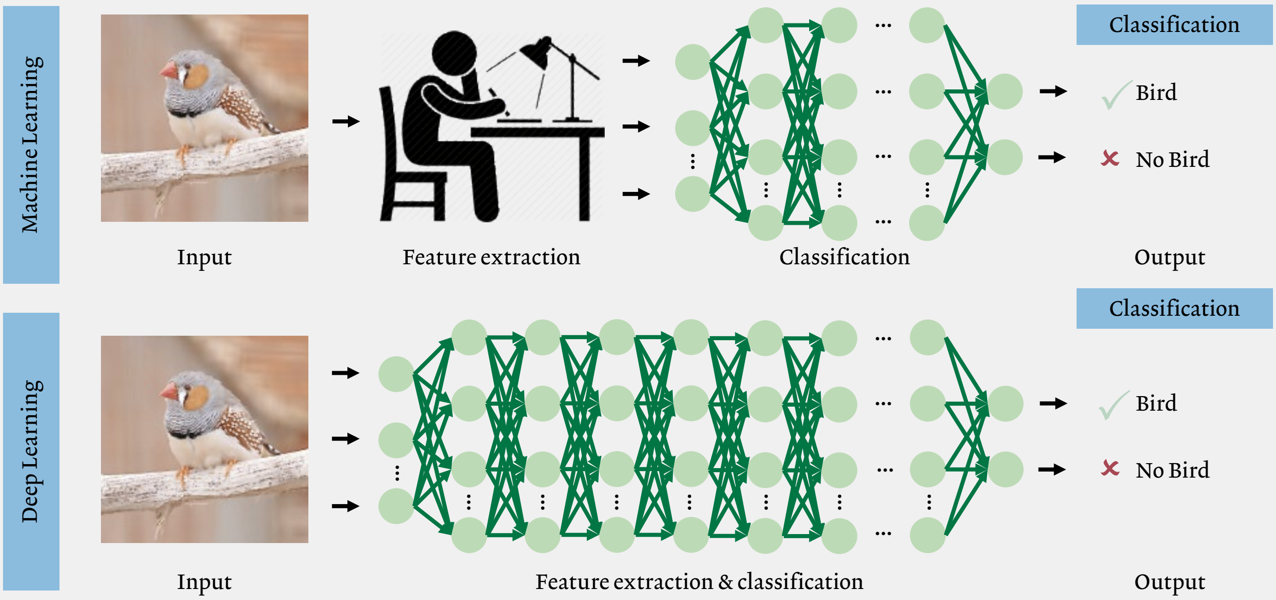
The application of DL techniques enables auditors to derive a precise, non-linear model of both journal entry characteristics. A network architecture particularly suitable for the unsupervised end-to-end learning of complex attribute characteristics are deep autoencoder neural networks (herafter referred to as autoencoders).

In general, an autoencoder refers to a special type of deep neural network that can be trained to reconstruct its input. Therefore, autoencoders consist of two interconnected neural networks, referred to as encoder and decoder networks, respectively [16]. In general, both networks exhibit a symmetric architecture and encompass several layers of artificial

T. FLEMMING RUUD,
PROF. EM., PH.D., CPA (NO),
CHAIR OF BUSINESS ADMINISTRATION (INTERNAL AUDIT, INTERNAL CONTROL),
UNIVERSITY OF ST. GALLEN

DAMIAN BORTH, PROF. DR.,
CHAIR OF ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING,
DIRECTOR AT THE INSTITUTE OF COMPUTER SCIENCE (ICS),
UNIVERSITY OF ST. GALLEN

Figure 2: **IMAGE CLASSIFICATION USING MACHINE LEARNING (TOP) AND DEEP LEARNING (BOTTOM)** [9]



neurons. *Figure 3* shows the schematic view of an autoencoder architecture. The training objective of autoencoders is to minimise the dissimilarity of a given input journal entry and its reconstruction as faithfully as possible. The difference between the original input and its reconstruction is thereby referred to as reconstruction error. To prevent the autoencoder from simply forwarding the journal entry attributes from the input layer to the output layer, the number of neurons of the inner network layers are reduced (usually referred to as "bottleneck"). Imposing such a constraint forces the autoencoder to learn a compressed model of the entries most prevalent attribute values and correlations.

As the autoencoder training progresses, the network learns the characteristic behavioural posting pattern evident in a given population of journal entries [17]. As a result, the autoencoder is increasingly able to reconstruct regular journal entries almost without errors. At the same time, anomalous journal entries, i. e. very rare attribute values and correlations cannot be learnt due to the limited capacity of the inner network layers. Ultimately, unusual postings result in an increased reconstruction error. For a given journal entry the error magnitude can then be interpreted as the degree of unexpected deviation from regular behavioural posting patterns. Based on this criterion, the reconstruction error enables auditors to distinguish regular journal entries from anomalies at the end of the training process.

The practical application of such a DL-enabled anomaly detection approach is described below in the context of the Nestlé S. A. internal audit function.

4. DETECTION OF ACCOUNTING ANOMALIES BY THE INTERNAL AUDIT OF NESTLÉ S. A.

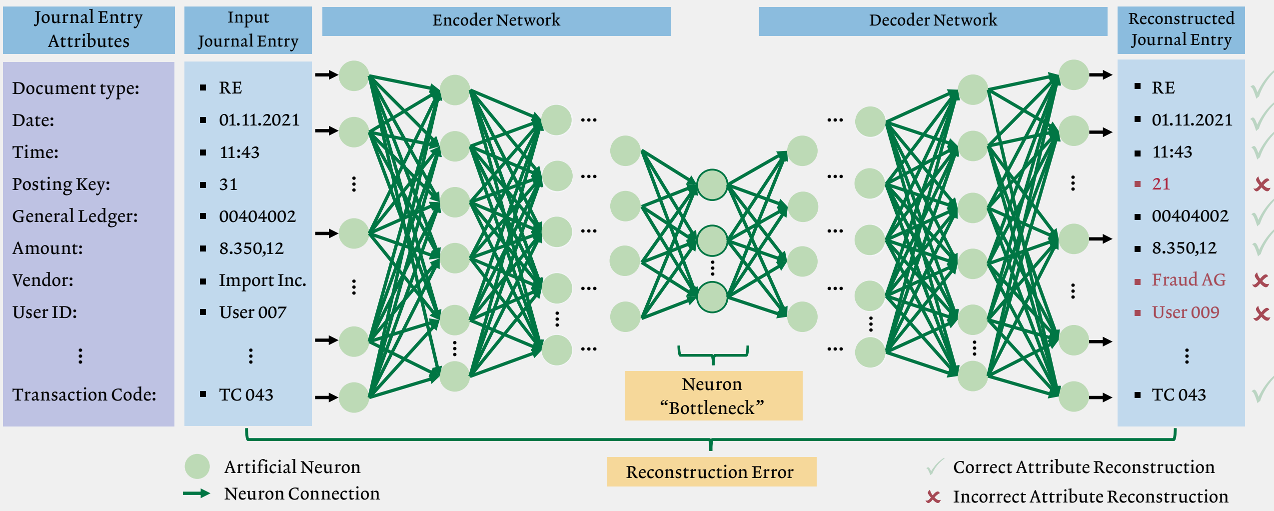
Within the internal audit function of Nestlé S. A. (hereafter referred to as Nestlé), data analytical audit procedures have been applied for more than ten years. There is naturally an interest in developing new analytical auditing methods based on AI. Due to the extensive use of SAP ERP systems within Nestlé, it is possible to extensively focus on the analysis of digital journal entries. In doing so, the postings are distinguished from a process point of view, e. g. into postings of the purchasing process (Accounts Payables, SAP T-Code: FBL1N) and the sales process (Accounts Receivables, SAP T-code: FBL5N). The volume of postings in the sales department of an entity such as Nestlé Switzerland comprises more than half a million journal entries per financial year.

Throughout the regular audit of selected Nestlé entities, data analytical audit procedures are performed prior to the assignment of the respective audit teams. The objective of the analyses is to identify anomalies in the journal entry population of the entity in-scope of the audit and derive a targeted audit sample. The audit teams are then mandated to conduct an on-site substantive audit of the anomalies. The deliberate process for identifying accounting anomalies follows the three distinct phases described below:

Phase 1: Extraction, validation, and preparation of accounting data

In the first phase, journal entry data of the entity in-scope of the audit is exported from the regional SAP ERP systems, validated, and prepared for subsequent analyses. Depending on the focus of the audit, the data is filtered for certain classes of internal and external suppliers and customers. The processing and subsequent analyses are mainly based on routines developed in open-source programming languages (such as R [18] and Python [19]). Based on these languages, a variety of analytical audit procedures have been developed in recent years within Nestlé's internal audit function. Nowadays,

Figure 3: **SCHEMATIC STRUCTURE OF A DEEP AUTOENCODER NEURAL NETWORK AND SAMPLE JOURNAL ENTRY RECONSTRUCTION** [17]



these developments enable auditors to efficiently export, prepare and analyse accounting data.

Phase 2: Univariate and multivariate statistical analyses
Based on selected journal entry attributes, univariate and multivariate statistical analyses are performed in the second phase. The focus here is, among other things, on the analysis of low-dimensional characteristic features and relationships. *Figure 4* shows an example of the result of a correlation analysis of the two journal entry attributes "Document Type" and "SAP Transaction Code". The representation as a heat map highlights regular and anomalous attribute value co-occurrences evident in the population of journal entries. Co-occurrences that have a comparatively low and thereby anomalous frequency are marked by dark red cells. The engagement team is then mandated to examine the corresponding journal entries and their underlying business transaction in terms of error, deviations, and fraud.

Phase 3: Artificial-Intelligence-enabled anomaly detection
In the third phase, the analyses of the second phase are extended by an AI-enabled audit process. The objective is to identify journal entries that exhibit unusual high-dimensional attribute values and attribute value combinations. In this context, the autoencoder neural networks described in section 3 are utilised. In the context of the example, journal entries that do not correspond to regular behavioural posting activities of the sales department result in a high reconstruction error. Consequently, this phase results in a report in which the analysed journal entries are ranked according to their error magnitude. The audit team is then again mandated to conduct an on-site audit of journal entries whose error magnitude exceeds a predefined threshold. Throughout the on-site audit, the audit team is required to examine the journal entries that have been flagged as anomalies by the autoencoder.

Figure 5 shows an overview of the autoencoder-based anomaly detection process of sales postings of a Nestlé entity. The autoencoder model training and inference was conducted using a virtual Microsoft Windows environment (Intel Xeon 2.3 GHz processor including 32 GB RAM). The encoder-decoder architecture encompasses 11 neural network layers consisting in total of 2,660 artificial neurons. Using the DL-enabled audit approach, the following (anonymised) cases could be identified as examples of accounting anomalies:

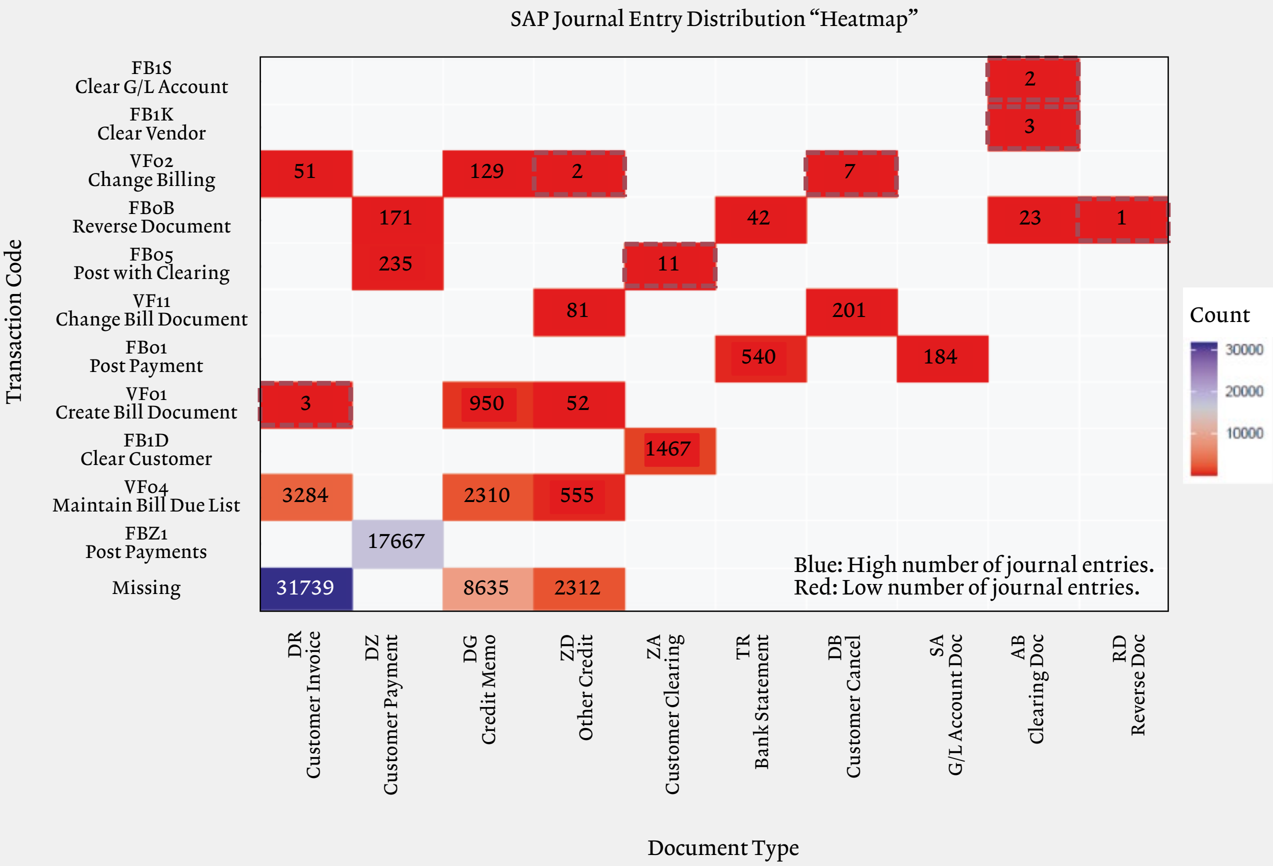
→ Within a Latin American entity, a credit posting in favour of a customer was identified. The posting, which was entered manually, affects a rarely used offset account. Furthermore, the posting exhibits a significantly high posting amount and was created by an unusual combination of posting key, customer number range and user ID.
→ Within an Asian entity, a series of credit entries were detected that have been posted by an unusual combination of user ID, posting key and clearing account. Using these credit entries, customer invoices with material amounts were manually cleared in the corresponding ERP system within a very short time interval.

5. CURRENT PRACTICAL CHALLENGES AND OUTLOOK

Nowadays, the application of AI and particular ML techniques is emerging as a valuable tool that can be used in various phases of the audit process. Furthermore, the gradual adoption of AI-enabled audit procedures can be observed by internal auditors [20]. For example, within the internal audit function at Nestlé, the anomalies identified by the autoencoder in interaction with human auditors are recorded in a designated database. The objective is to continuously transfer the detection of such anomalies into rule-based auditing routines in order to detect sources of error or fraud patterns even more efficiently in the future.

At the same time, the enhancement of audit data analytics by AI capabilities poses well-known challenges, which con-

Figure 4: **HEATMAP OF SELECTED JOURNAL ENTRY ATTRIBUTES: "DOCUMENT TYPE" AND "SAP TRANSACTION CODE"**

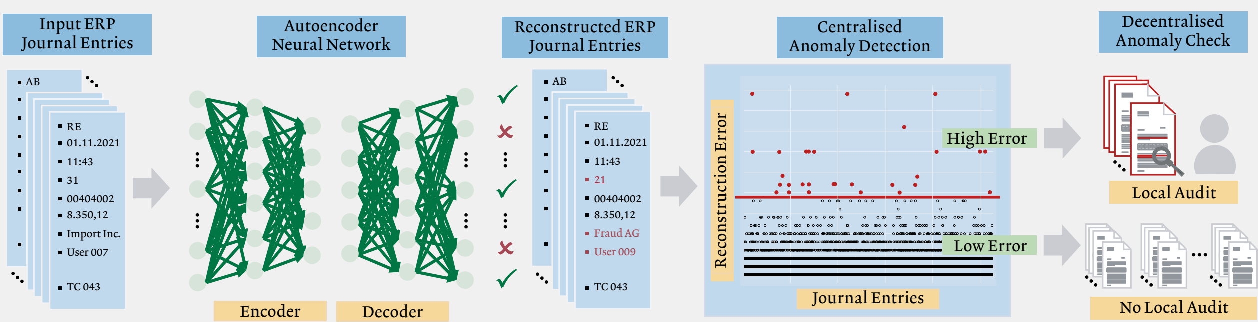


cern data extraction, data preparation, and data protection. In addition, as a result of the application of AI, internal audit is currently facing a number of new challenges, such as legal liability, data privacy, and model compliance [21]. In particular, it often remains unclear which members of an internal audit team should deploy and apply machine-learning methods. A combination of mature audit experience and technical expertise is required for beneficial use. Thus, within the internal audit at Nestlé, a major challenge is the interpretation of the results obtained by AI techniques, such as autoencoder neural networks. For an appropriate interpretation, auditors are required to have a solid understanding of:

- the analysed business processes, controls, and associated risks;
- the ERP systems of their organisation; and,
- the applied statistical analysis and machine-learning methods.

In the medium term, audit departments themselves will therefore be challenged to invest in the training of their au-

Figure 5: **SCHEMATIC PROCEDURE MODEL OF AUTOENCODER-BASED ANOMALY DETECTION WITHIN THE SAP ERP JOURNAL ENTRIES OF A NESTLÉ S. A. ENTITY**



ditors in order to build up expertise at the interface of technical, analytical, and other audit skills.

In conclusion, the question may certainly be asked whether the primary task of internal audit is exhausted in the continuous detection of fraud and error or whether the objective is not rather to assess whether processes function within predefined tolerance limits. Whether this is a purely philosophical question or one that will determine the future is open to debate. For some time now, auditors have been encouraged by the IIA to interpret and exercise their function in the sense of a "trusted advisor" [22]. In this role, internal auditing strives to "add value and improves an organi[z]ation's operations", as embodied in the initially stated IIA definition [24]. In the future, the application of AI can help internal auditors achieve this goal within their organisations. ■

Footnotes: *The authors would like to thank Yasmine Bensultana-Weiser, Monika Heyder and Michael Mommert for their numerous, valuable suggestions and critical review of the manuscript. The content of this article reflects the views only of the authors, which do not necessarily reflect the official position of Nestlé S.A.
1) The Institute of Internal Auditors. Research Foundation, 2017. International Professional Practices Framework (IPPF). **2)** The Institute of Internal Auditors, 2012. International Standards for the Professional Practice of Internal Auditing 1220.A2, p. 6. **3)** Dai, J. and Vasarhelyi, M.A., 2016. Imagineering Audit 4.0. Journal of Emerging Technologies in Accounting, 13/1, pp. 1–15. **4)** Kokina, J. and Davenport, T. H., 2017. The Emergence of Artificial Intelligence: How Automation is Changing Auditing. Journal of Emerging Technologies in Accounting, 14/1, pp. 115–122. **5)** Ruud, F., Schramm, K. and Allgaier, A., 2021. Leitlinien zum Internen Audit, IIA Switzerland, 4th edition, pp. 125–126. **6)** Russell, S. and Norvig, P., 2016. Artificial Intelligence: A Modern Approach, Third Edition, Pearson Education Ltd., pp. 1–5. **7)** Samuel, A.L., 1959. Some Studies in Machine Learning Using the Game of Checkers. IBM Journal of Research and Development, 3/3, pp. 210–229. **8)** Bishop, C. M., 2006. Pattern Recognition and Machine Learning, Springer Science LLC, pp. 3–4. **9)** Gierbl, A.S., Schreyer, M., Borth, D. and Leibfried, P., 2021. Deep Learning für die Wirtschaftsprüfung – Eine Darstellung von Theorie, Funktionsweise und Anwendungsmöglichkeiten. Zeitschrift für Internationale Rechnungslegung (IRZ), 7/8, pp. 349–355. **10)** LeCun, Y., Bengio, Y. and Hinton, G., 2015. Deep learning. Nature, 521/7553, pp. 436–444. **11)** Mc-Culloch, W.S. and Pitts, W., 1943. A Logical Calculus of the Ideas Immanent in Nervous Activity. The Bulletin of Mathematical Biophysics, 5/4, pp. 115–133. **12)** Rosenblatt, F., 1958. The Perceptron: A probabilistic Model for Information Storage and Organization in the Brain. Psychological review, 65/6, p. 386. **13)** Bengio, Y., Courville, A. and Vincent, P., 2013. Representation Learning: A Review and New Perspectives. IEEE Transactions on Pattern Analysis and Machine Intelligence, 35/8, pp. 1798–1828. **14)** Sun, T., 2019. Applying Deep Learning to Audit Procedures: An Illustrative Framework. Accounting Horizons, 33/3, pp. 89–109. **15)** Schreyer, M., Sattarov, T., Schulze, C., Reimer, B. and Borth, D., 2019. Detection of Accounting Anomalies in the Latent Space Using Adversarial Autoencoder Neural Networks. arXiv preprint arXiv: 1908.00734. **16)** Hinton, G.E. and Salakhutdinov, R. R., 2006. Reducing the Dimensionality of Data with Neural Networks. Nature, 313/5786, pp. 504–507. **17)** Schreyer, M., Sattarov, T., Borth, D., Dengel, A. and Reimer, B., 2017. Detection of Anomalies in Large Scale Accounting Data Using Deep Autoencoder Networks. arXiv preprint arXiv:1709.05254. **18)** The R Project for Statistical Computing (https://www.r-project.org, accessed on 7 November 2021). **19)** The Python Programming Language (https://www.python.org, accessed on 7 November 2021). **20)** Zhang, C. A., Dai, J. and Vasarhelyi, M.A., 2018. The Impact of Disruptive Technologies on Accounting and Auditing Education: How Should the Profession Adapt? The CPA Journal, 88/9, pp. 20–26. **21)** Gierbl, A.S., Schreyer, M., Borth, D. and Leibfried, P., 2020. Künstliche Intelligenz in der Prüfungspraxis – Eine Bestandsaufnahme aktueller Einsatzmöglichkeiten und Herausforderungen. Expert Focus, 2020/9, pp. 612–617. **22)** Ruud, F. and Schramm, K., 2019. Erwartungen und Mehrwert des Internen Audits – Die Rolle wirksamer Kommunikation. Expert Focus, 2019/10, pp. 49–54. **23)** Borth, D., 2020. Introduction to Machine Learning and Deep Learning. Lehrveranstaltung an der Universität St. Gallen (HSG). **24)** The Institute of Internal Auditors. Research Foundation, 2017. International Professional Practices Framework (IPPF).