DigitalTraces: Unveiling fraud through interactive user behaviour exploration

João Bernardo Narciso, Beatriz Feliciano, Rita Costa, and Pedro Bizarro

Feedzai

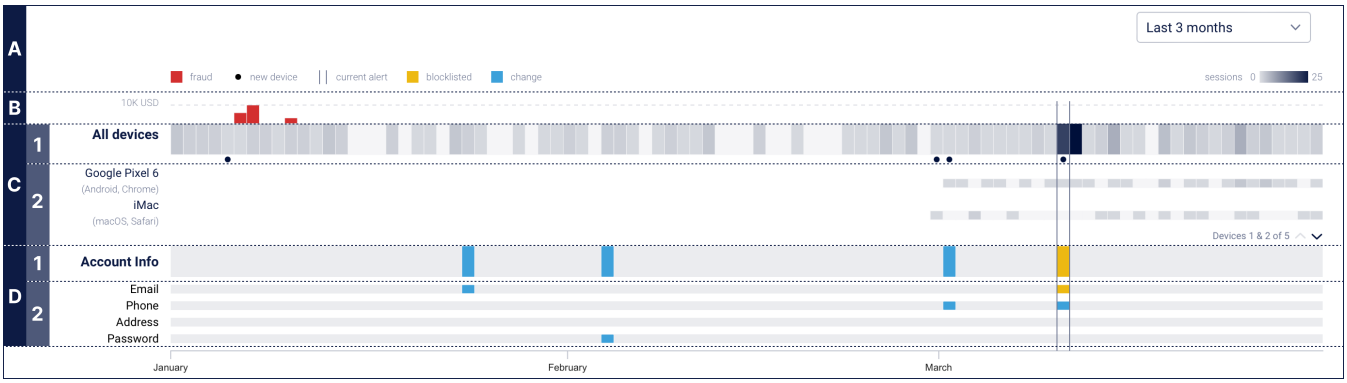

Figure 1: *Representation of the DigitalTraces' interface, depicting a scenario of a client account takeover (ATO) by fraudsters after the introduction of a new device. The central elements of the visualisation are the heatmaps that display the overall device information timeline* ***(C1)*** *and the device specific data* ***(C2)****. This information is enriched by account information* ***(D1)****, that signals changes in email, phone, password, and address* ***(D2)****. The interface is toped by a bar chart that represents the fraud records for that person* ***(B)*** *and an header with a legend and time window selector* ***(A)***

Abstract

Fraud detection teams in financial institutions face the challenge of identifying suspicious activity within user behaviour. However, existing tools often lack the ability to seamlessly integrate multiple dimensions of digital activity into a single, interactive visualisation, leading to increased cognitive load and preventing analysts from quickly spotting anomalies in varying sources of information. This paper introduces DigitalTraces, a visual analytics tool aimed at improving the detection of fraudulent patterns particularly in the dimensions tied with digital activity. The system combines several stacked timelines to offer an overview of multiple activity dimensions, integrating online banking session data, device identifiers, transactional activities, and account information. We validated our tool with a think-aloud experiment where two fraud analysts were tasked with detecting anomalies in a financial fraud scenario. Experts emphasised the tool's ability to provide intuitive insights and enhance understanding.

CCS Concepts

• ***Human-centered computing*** → ***visualisation systems and tools;***

1. Introduction

When using their computer, phones, or smartwatches to make payments, transfers or other financial transactions, people leave a trace. A digital footprint that indicates what are their habits and usual behaviours. A change in those might indicate fraudulent activity. For that reason, digital activity is one important dimension, among others, that analysts look into in the context of fraud detection. One key aspect of digital interactions is a session, which represents a continuous period of user activity on an online financial or banking platform. Sessions help track normal user behaviour and can be analysed for anomalies, as fraudulent activity often involves deviations from established patterns [SLT*20]. Detecting anomalous user behaviours is therefore essential for identifying potential financial fraud, as fraudulent activity is often preceded by user behaviour

© 2025 The Author(s).
Proceedings published by Eurographics - The European Association for Computer Graphics.
This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

that deviates from what was seen in the past [LGS*22]. Suspicious activity can include unusual login attempts from new devices or locations, sudden changes in transaction frequency, modifications to account details such as email or phone number, or high-velocity interactions with online platforms [SLT*20]. While each of these actions alone may not be inherently fraudulent, it's their combination that might indicate suspicious behaviour (e.g. a change in email from a new device combined with unusually high transactional activity).

Uncovering the combinations, sequences and patterns in data is a crucial task for fraud analysts, especially as financial crime rises. According to the Financial Trade Commission (FTC), fraud and identity theft have been growing continuously for the past decades, costing U.S. customers over $10 billion in 2023. Furthermore, digital activity-related crimes such as account takeovers (ATO) are among the most common types of financial crime [KDC*19, Fed24]. However, existing tools often lack the ability to seamlessly integrate multiple dimensions of digital activity into a single, interactive visualisation. This fragmentation forces analysts to manually cross-reference different data sources, increasing cognitive load and slowing down fraud detection efforts.

This paper introduces *DigitalTraces*, a visualisation tool designed to support fraud detection by offering a comprehensive and interactive overview of digital activity data. By integrating online banking session logins, device identifiers, transactional activity and account data information, the tool enables analysts to easily track and investigate patterns that may signify fraudulent behaviour. The interface allows for the dynamic analysis of multiple data dimensions and supports varying time windows. It helps analysts in identifying novel devices, detecting changes in user behaviour, and correlating digital activity with transactional fraud data.

This paper makes two key contributions. First, it identifies key requirements for analysts when reviewing digital activity data, based on our research and interviews with analysts. Secondly, it introduces a tool designed to enhance the analytical capabilities of fraud detection teams by providing an intuitive interface for detailed, multi-dimensional analysis of digital activity data.

2. Related Work

Several visualisations have been developed to analyse anomalous user behaviours across diverse domains, offering insights into complex digital interactions [SLT*20]. TargetVue [CSL*16] provides an interactive method to examine irregular communication patterns, while VASABI [NHC*20] uses hierarchical representations to summarize user activity at multiple levels. MOOCad [MXC*19] reveals irregular patterns in large-scale learning sequence data, highlighting the benefits of integrating temporal and categorical information. Additionally, #FluxFlow [ZCW*14] demonstrates how dynamic visualisations can capture abrupt changes in information spreading. These studies motivate the approach in *DigitalTraces*, which consolidates sessions data, device information, account information changes, and fraud history into a unified visualisation.

Focusing on fraud detection, research demonstrates that interactive dashboards and visual analytical tools significantly reduce the cognitive effort required to detect anomalies, helping analysts to quickly identify suspicious patterns in financial environments. Different studies have devised visual systems that not only highlight unusual activity but also integrate interactive features — such as tooltips, dynamic filtering, and the visualisation of multiple datasets at once — to enhance investigative processes [SMPM21, CLG*08, ZWW*23, FVS*23, FCA*24].

There has been considerable progress in integrating and summarising data from heterogeneous sources. Fraud analysis often requires the consolidation of session logs, device identifiers, account modifications, and fraud records into a single, coherent overview. Researchers have proposed methods for aggregating diverse data streams, ensuring that key signals — particularly those indicating shifts in user behaviour prior to transactional fraud — are preserved for analysis. This holistic approach enables analysts to detect anomalies more effectively and contextualize them within broader activity patterns [AAS23, NSH*18]. *DigitalTraces* draws inspiration from previous work while proposing a way to integrate in the visualisation different sets of data. The most significant advance is that this integration is done by stacking multiple timelines, with variable granularity, which enables the exploration and comparison of different kinds of data.

An effective strategy for visualising time-sensitive data is using timelines and heatmaps. CloudLines [KBK11] highlights temporal patterns across multiple time-series, providing a compact overview of event sequences. Heatmaps can draw attention to periods with unusually high activity, and interactive timelines allow for a detailed exploration of events within specific intervals [PMC*22]. Recent work has enhanced these methods by incorporating additional layers of context [CXC*24, BBC*25]. Our proposal leverages the power of heatmaps to highlight temporal patterns and spikes, directing the user to the suspicious activities that require attention, potentially uncovering fraud.

While the referenced previous studies and proposals are useful and serve their purpose, they do not fully address the specific requirements our proposal aims to meet in the context of digital activity for fraud detection. Our proposal offers a comprehensive approach to fraud detection in digital user behaviour by evaluating all key aspects of a user's behaviour simultaneously, providing a holistic view, reducing cognitive load and enabling analysts to quickly identify anomalies with greater ease and accuracy.

3. *DigitalTraces*

3.1. Design Requirements

Based on the related work and interviews with analysts, we developed a list of requirements that guided the design of *DigitalTraces*.

Fraud analysts rely on digital activity data to detect suspicious behaviour and assess fraud. Their workflow involves reviewing session login details, device identifiers (e.g., model name and operating system), and account identifiers (e.g., addresses, emails, phone numbers, and passwords). They focus on identifying deviations in user behaviour, such as previously unseen devices, blocklisted phone numbers or email accounts (in other words, previously flagged as fraudulent by analysts), and unusual changes in activity frequency like login spikes.

© 2025 The Author(s).
Proceedings published by Eurographics - The European Association for Computer Graphics.

Given their time constraints [SMPM21], analysts typically review only the past three months of data but may adjust the time window for broader trends (e.g., usage spikes during a couple of days) or more granular insights (e.g., abnormal behaviour within a day). Furthermore, they perform multi-factor analysis, integrating both transactional and non-transactional activities, as fraudulent digital behaviour often precedes actual fraudulent transactions [LGZ*20].

From these observations, we derived a set of design requirements to support analysts' workflows effectively:

- **R1:** Display comprehensive information for all device information and usage, and account information changes, providing an overview of user digital activity.
- **R2:** Enable the identification of new devices, highlighting when they were introduced and used.
- **R3:** Allow users to detect changes in behaviour frequency by comparing them against average client activity.
- **R4:** Integrate transactional fraud data history to support correlation analysis with digital activities.
- **R5:** Support variable time windows to accommodate different analytical needs.
- **R6:** Facilitate the analysis of changes across multiple digital data dimensions simultaneously.
- **R7:** Integrate with Feedzai's fraud detection software.

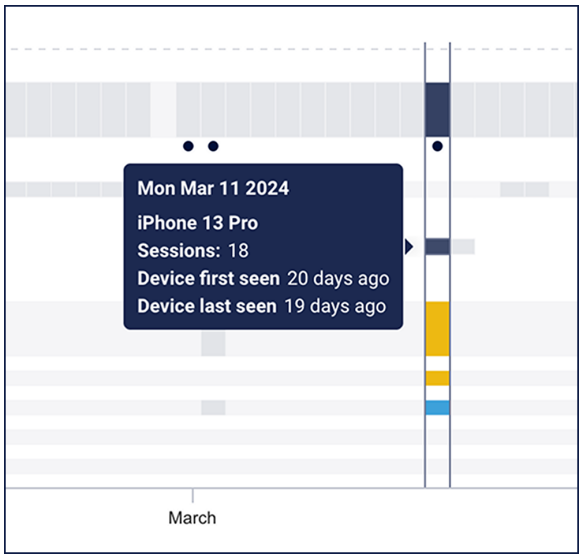

Figure 2: *Detail of a hover on a device timeline, triggering a tooltip with additional information, and highlighting the blocks of that single day, allowing for cross inspection of different data for the same period **[R6]**.*

3.2. Interface

The interface of *DigitalTraces* is composed of five sections, providing a complete overview for the different types of digital activity data (Figure 1): the header section **(A)**; the fraud bar chart **(B)**; the device information timelines **(C)**; and the account information timelines **(D)**. There is also a vertical *current alert* indicator that indicates the time block that raised the alert the analyst should review, intercepting sections **B**, **C**, and **D**. The data shown is ingested via parsers that map incoming fields to a common internal schema, enabling consistent alignment across timelines.

3.2.1. Header

The header **(A)** of the element is composed by two parts. Firstly, a legend that provides details on the various shapes and colour encodings, including the heatmap's colour range represented by the number of sessions, making sure that tool's insights are understood with ease. Secondly, a time window selector, that allows users to adjust the time window for their analysis, offering flexibility based on the needs of the investigation. Time windows can be set to predefined ranges based on the most common analysis intervals used by analysts. The granularity of the visualisation time blocks depends on the selected time window: one hour for the 24-hour time window, and one day for the rest **[R5]**.

3.2.2. Fraud

The fraud section **(B)** is composed of a bar chart that sits on top of the session summary timeline. Each bar represents the sum amount identified as fraud for the respective time block. It also includes a label with the maximum amount of the chart **[R4]**.

3.2.3. Device Information

The device information section **(C)** is composed of several horizontal timelines. The first element is the **device summary timeline (C1)** and it is a heatmap visualisation of number of logins for all the user devices where each time block is represented by a rectangle. The heatmap sequential colour scale goes from light grey, representing a time window with zero login sessions, to a darker shade, representing a time block with the most login sessions of the selected time window. When hovered, a tooltip appears with the number of sessions for the respective time block, amount of fraud, and characteristics of the device(s) added, if any **[R3]**.

Beneath **C1** there is a set of stacked heatmaps,the **device timelines (C2)**. There is one for each device with activity during the selected time window. These follow the same colour scale as the main heatmap. When a block is hovered, a tooltip displays the device characteristics, number of sessions for that device and when the device was first and last seen (Figure 2). Each timeline is named after the correspondent device's characteristics. When the name is clicked, the unique ID of the device is copied to the user's clipboard, allowing for further investigation in a fraud detection software **[R7]**. To keep a fixed height, a maximum of two device timelines, ordered from the most recent, can appear at the same time. To navigate between timelines, the user can either scroll or use the navigation buttons on the bottom right corner of this section. Next to the buttons is a text indicating the index of the devices currently visible and the total number of devices. Below each rectangle of **C1** there may be a circle indicating that a new device was introduced during that period. That circle can be clicked, resulting in the display of the respective device timeline without the analyst having to manually navigate through the rest of the timelines **[R1]**.

© 2025 The Author(s).
Proceedings published by Eurographics - The European Association for Computer Graphics.

3.2.4. Account information

The account information section **(D)** is also composed of several horizontal timelines. Like in Device Information **(C)**, this area includes a main timeline that aggregates all data and several more detailed ones. The **account information summary timeline (D1)** is coloured by rectangles that present the time blocks in which an account data change occurred. These rectangles can be either blue or yellow, representing an account data change or an account data change to a blocklisted mean, respectively. The **account information specific timelines (D2)**, following the same colour encoding, expands the information of the summary timeline, representing changes for each of the account information considered: email, phone number, password, and home address. When a block is hovered, a tooltip displays the previous and the new account data (does not apply to passwords) **[R2]**.

4. Experts Feedback

To evaluate the effectiveness and usability of *DigitalTraces* in real-world scenarios, we conducted a user study with two risk strategy analysts, with 15 and 10 years of experience each in fraud detection. They had no previous contact with *DigitalTrust*. Each session lasted approximately 40 minutes and was divided into two phases: an initial free exploration phase and a task-based evaluation.

The data shown in the interface is synthetic and mimics an ATO scenario (Figure 1). Synthetic **account change information** was generated using a conditional probabilistic model. Each account information is sampled from a categorical distribution with change states for each time block. For the same time block, if there was a change in any account information, a new conditional distribution is used for the rest of the attributes, increasing the likelihood of simultaneous changes. **Device activity**, or the number of events per day for each device follow a Poisson distribution with λ values dependent on device and date range. **Fraudulent events** are mimicked by randomly selecting a random number of dates and assigning them a fraud value.

During the free exploration phase, the analysts were given unrestricted time to navigate the interface and familiarise themselves with its interactive features, while employing the think-aloud protocol [vSBS94] to verbalize their thought processes. This approach enabled us to capture their immediate impressions and reasoning strategies. With minimal input from our side, the analysts understood what all elements in the visualisation represented by resorting to the legend. After this exploratory phase, the analysts were given a specific task aimed at identifying potential fraudulent behaviour in the digital activities of a client.

The task was for the analysts to identify every aspect of the user behaviour that might be considered suspicious. Using *DigitalTraces*, the analysts began by examining the summary of sessions **(C1)**. This summary presents a heatmap that is especially dense in a region consisting of two days. With that, the analysts instantly recognized a surge in activity during that particular period. The first day of that period coincides with the current alert indicator, which means that the fraud detection system flagged that event for review. Hovering over this section, the analysts resorted to the tooltip feature for detailed insights, confirming a spike in logins. By looking into the new device indicator, the analysts correlates the spike of logins with the introduction of a new device. This correlation is confirmed with the device specific timeline **(A2)**. Further investigation using the account information timelines unveils that, during this peak period, there was a change **(B1)** in the client's registered email to one previously flagged as blocklisted **(B2)**. The analysts also successfully identified previous fraud history that does not seem to be related with the current suspicious activity, because it happened prior to the introduction of the suspicious device and was associated with another device that seems to be often used by the client. In conclusion, experts successfully identified all signs of frauds and validated the scenario characterizing it as *realistic*.

After the exploration and task completion, the experts praised the ability to quickly identify suspicious activities and the ability to combine in one visualisation data from multiple sources. This integration of distinct data dimensions, such as account changes and device usage patterns, was highlighted as particularly beneficial. This feature assists analysts not only in identifying potential fraudulent behaviour but also in understanding the context of such behaviour, making the tool highly effective in supporting comprehensive fraud investigations.

Another highlight concerns the user interface in a general sense, which was complimented for its capability to present complex data in an intuitive manner. The interactive elements, such as clickable indicators that directly link to relevant device timelines, were appreciated for enhancing the exploratory experience, allowing users to delve deeper into specific devices with ease. While the overall feedback was very positive, the experts did provide constructive suggestions that could further optimize the user experience. Namely, simplifying or removing the account information summary timeline **(D1)** to reduce redundancy could help in maintaining the visualisation's effectiveness without overwhelming the user; and having the ability to visualize a bigger number of selected devices at the same time could also help in comparison tasks.

5. Discussion and Future Work

Experts have validated the usefulness of *DigitalTraces* in supporting anomaly detection in investigative workflows, confirming its potential to improve analysts' ability to identify and respond to potentially fraudulent activities. However, *DigitalTraces* has yet to undergo formal user testing. Future work should include usability studies to validate its effectiveness. Replacing synthetic data with real datasets would provide a more accurate assessment of its practical utility. Moreover, introducing a custom time selector would allow for a more flexible analysis.

In conclusion, *DigitalTraces* presents a novel approach to fraud detection by integrating and visualizing multiple dimensions of digital activity data in an interactive interface. By enabling analysts to quickly identify suspicious patterns, correlate digital behaviours with potential fraud, and streamline investigative workflows, the tool enhances analytical capabilities of fraud detection teams. Experts feedback has validated its usefulness. Ultimately, *DigitalTraces*' key contribution lies in its ability to combine multiple data sources — session activity, device information, account changes, and fraud history — into a unified visualisation tailored for fraud analysis.

© 2025 The Author(s).
Proceedings published by Eurographics - The European Association for Computer Graphics.

References

[AAS23] ANDRIENKO N., ANDRIENKO G., SHIRATO G.: Episodes and topics in multivariate temporal data. *Computer Graphics Forum 42*, 6 (Aug. 2023). URL: http://dx.doi.org/10.1111/cgf.14926, doi:10.1111/cgf.14926. 2

[BBC*25] BERNARD J., BARTH C.-M., CUBA E., MEIER A., PEIRIS Y., SHNEIDERMAN B.: Ivesa – visual analysis of time-stamped event sequences. *IEEE Transactions on Visualization and Computer Graphics 31*, 4 (Apr. 2025), 2235–2256. URL: http://dx.doi.org/10.1109/tvcg.2024.3382760, doi:10.1109/tvcg.2024.3382760. 2

[CLG*08] CHANG R., LEE A., GHONIEM M., KOSARA R., RIBARSKY W., YANG J., SUMA E., ZIEMKIEWICZ C., KERN D., SUDJIANTO A.: Scalable and interactive visual analysis of financial wire transactions for fraud detection. *Information Visualization 7*, 1 (Feb. 2008), 63–76. URL: http://dx.doi.org/10.1057/palgrave.ivs.9500172, doi:10.1057/palgrave.ivs.9500172. 2

[CSL*16] CAO N., SHI C., LIN S., LU J., LIN Y.-R., LIN C.-Y.: Targetvue: Visual analysis of anomalous user behaviors in online communication systems. *IEEE Transactions on Visualization and Computer Graphics 22*, 1 (Jan. 2016), 280–289. URL: http://dx.doi.org/10.1109/TVCG.2015.2467196, doi:10.1109/tvcg.2015.2467196. 2

[CXC*24] CANTAREIRA G. D., XING Y., COLE N., BORGO R., ABDUL-RAHMAN A.: Interactive hierarchical timeline for collaborative text negotiation in historical records. *IEEE Transactions on Visualization and Computer Graphics* (2024), 1–12. URL: http://dx.doi.org/10.1109/TVCG.2024.3376406, doi:10.1109/tvcg.2024.3376406. 2

[FCA*24] FELICIANO B., COSTA R., ALVES J., LIÉBANA J., DUARTE D., BIZARRO P.: ""Show Me What's Wrong!"": Combining charts and text to guide data analysis, 2024. URL: https://arxiv.org/abs/2410.00727, doi:10.48550/ARXIV.2410.00727. 2

[Fed24] FEDERAL TRADE COMMISSION: Consumer Sentinel Network Data Book 2023, February 2024. URL: https://www.ftc.gov/system/files/ftc_gov/pdf/CSN-Annual-Data-Book-2023.pdf. 2

[FVS*23] FIRAT E. E., VYTLA D., SINGH N. V., JIANG Z., LARAMEE R. S.: MoneyVis: Open Bank Transaction Data for Visualization and Beyond. In *EuroVis 2023 - Short Papers* (2023), Hoellt T., Aigner W., Wang B., (Eds.), The Eurographics Association. doi:10.2312/evs.20231052. 2

[KBK11] KRSTAJIC M., BERTINI E., KEIM D.: Cloudlines: Compact display of event episodes in multiple time-series. *IEEE Transactions on Visualization and Computer Graphics 17*, 12 (Dec. 2011), 2432–2439. URL: http://dx.doi.org/10.1109/TVCG.2011.179, doi:10.1109/tvcg.2011.179. 2

[KDC*19] KAWASE R., DIANA F., CZELADKA M., SCHÜLER M., FAUST M.: Internet fraud: The case of account takeover in online marketplace. In *Proceedings of the 30th ACM Conference on Hypertext and Social Media* (Sept. 2019), HT '19, ACM, p. 181–190. URL: http://dx.doi.org/10.1145/3342220.3343651, doi:10.1145/3342220.3343651. 2

[LGS*22] LIU C., GAO Y., SUN L., FENG J., YANG H., AO X.: User behavior pre-training for online fraud detection. In *Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining* (Aug. 2022), KDD '22, ACM, p. 3357–3365. URL: http://dx.doi.org/10.1145/3534678.3539126, doi:10.1145/3534678.3539126. 2

[LGZ*20] LIU G., GUO J., ZUO Y., WU J., GUO R.-Y.: Fraud detection via behavioral sequence embedding. *Knowledge and Information Systems 62*, 7 (Jan. 2020), 2685–2708. URL: http://dx.doi.org/10.1007/s10115-019-01433-3, doi:10.1007/s10115-019-01433-3. 3

[MXC*19] MU X., XU K., CHEN Q., DU F., WANG Y., QU H.: Moocad: Visual analysis of anomalous learning activities in massive open online courses. *EuroVis 2019 - Short Papers* (2019). URL: https://diglib.eg.org/handle/10.2312/evs20191176, doi:10.2312/EVS.20191176. 2

[NHC*20] NGUYEN P. H., HENKIN R., CHEN S., ANDRIENKO N., ANDRIENKO G., THONNARD O., TURKAY C.: Vasabi: Hierarchical user profiles for interactive visual user behaviour analytics. *IEEE Transactions on Visualization and Computer Graphics 26*, 1 (Jan. 2020), 77–86. URL: http://dx.doi.org/10.1109/TVCG.2019.2934609, doi:10.1109/tvcg.2019.2934609. 2

[NSH*18] NIEDERER C., STITZ H., HOURIEH R., GRASSINGER F., AIGNER W., STREIT M.: Taco: Visualizing changes in tables over time. *IEEE Transactions on Visualization and Computer Graphics 24*, 1 (Jan. 2018), 677–686. URL: http://dx.doi.org/10.1109/TVCG.2017.2745298, doi:10.1109/tvcg.2017.2745298. 2

[PMC*22] PALMEIRO J., MALVEIRO B., COSTA R., POLIDO D., MOREIRA R., BIZARRO P.: Data+shift: Supporting visual investigation of data distribution shifts by data scientists. URL: https://diglib.eg.org/handle/10.2312/evs20221097, doi:10.2312/EVS.20221097. 2

[SLT*20] SHI Y., LIU Y., TONG H., HE J., YAN G., CAO N.: Visual analytics of anomalous user behaviors: A survey. *IEEE Transactions on Big Data* (2020). URL: http://dx.doi.org/10.1109/TBDATA.2020.2964169, doi:10.1109/tbdata.2020.2964169. 1, 2

[SMPM21] SILVA P., MAÇÃS C., POLISCIUC E., MACHADO P.: Visualisation tool to support fraud detection. In *2021 25th International Conference Information Visualisation (IV)* (July 2021), IEEE, p. 77–87. URL: http://dx.doi.org/10.1109/IV53921.2021.00022, doi:10.1109/iv53921.2021.00022. 2, 3

[vSBS94] VAN SOMEREN M., BARNARD Y., SANDBERG J.: *The think aloud method: a practical approach to modelling cognitive processes*. Academic Press, 1994. 4

[ZCW*14] ZHAO J., CAO N., WEN Z., SONG Y., LIN Y.-R., COLLINS C.: #fluxflow: Visual analysis of anomalous information spreading on social media. *IEEE Transactions on Visualization and Computer Graphics 20*, 12 (Dec. 2014), 1773–1782. URL: http://dx.doi.org/10.1109/TVCG.2014.2346922, doi:10.1109/tvcg.2014.2346922. 2

[ZWW*23] ZHOU J., WANG X., WANG J., YE H., WANG H., ZHOU Z., HAN D., YING H., WU J., CHEN W.: Fraudauditor: A visual analytics approach for collusive fraud in health insurance. URL: https://arxiv.org/abs/2303.13491, doi:10.48550/ARXIV.2303.13491. 2

© 2025 The Author(s).
Proceedings published by Eurographics - The European Association for Computer Graphics.