A New AI-Driven Risk Assessment Tool for Investigating Insider Theft and Associated Maritime Crimes in a Southeast Asian Energy Company—A Case Study

Alex Martin
CEO & Co-Founder
Clearspeed, San Diego, CA, USA

Ben Smith
Clearspeed, San Diego, CA, USA

Keywords
Internal theft, maritime investigations, risk assessment, security, voice, artificial intelligence

Abstract
Ever since criminal networks have recognized the profit in oil and energy pipelines, the theft of hydrocarbon-based products has jeopardized the stability and security of global regions. Although numerous pipelines run across land and below the oceans, tankers serve as the most efficient way of transporting crude oil and natural gas between continents. This applied research study describes a novel AI-powered, a voice-based tool that identified human risk in a multi-national Southeast Asian energy company weakened by large-scale internal theft. 78.6 percent of completed automated interviews resulted in risk-positive evaluations. Ground truth from testimonial interviews and an internal investigation verified 92.6 percent of scrutinized flags. Previously undiscovered details were identified by the automated tool regarding the scope, size, and scale of crime issues, involving all job levels and local politicians. Analyses provided evidence of the technology's non-biased nature and demonstrated that its algorithm-generated outputs may be more dependable than observable behavioural cues. Findings (1) describe a potential decision support tool for detecting risk in situ, (2) contribute to employee fraud and internal theft literature, and (3) indicate that in the southeast Asian energy industry, approval for the approach described and recognition of its contribution are overwhelming.

Corresponding author: Alex Martin
Email address for the corresponding author: alex.martin@clearspeed.com
The first submission received: 1st April 2022
Revised submission received: 25th August 2022
Accepted: 5th September 2022

Introduction

In 2016, executives of a major Southeast Asian energy company became aware of losses amounting to $8 million of company revenue at one of their plants resulting from insider theft of petroleum products. Their internal audit of maritime operations and survey of personnel not only revealed the losses had transpired over five years (between 2012-2016), but also that theft had overwhelmingly occurred in three ways: skimming of mooring fees, manipulation of sludge production, and siphoning of fuel during bunkering procedures. Further, the allegation of an attempted felony also required investigating. To glean additional information and details as to the size, scale, and scope of the problem(s), the company contracted a crisis management firm.

As part of its crime review methodology, the firm incorporated Clearspeed technology as an AI-assisted risk screening tool. The subsequent evaluation was ideal for retrospective analysis as (1) the environment and population reflected the real-life conditions where the tool was designed to be used, (2) the sample size afforded high statistical power, (3) multiple variables were available for analysis, and (4) objective verification was available for flagged results.

Literature Review

Since the late nineteenth century, when the first internal combustion engine was created (Wang *et al*., 2020), humans have depended on oil, making it one of the modern world's most precious commodities. Today, daily crude oil use remains at over 95 million barrels globally, with up to 20 percent consumed in the U.S. alone ("Frequently asked…", 2021). The multiple industries heavily reliant on it include agriculture, aerospace, automobiles, aviation, chemicals, healthcare, homoeopathy, the military, plastics, and shipping (Dalby, 2014; Desjardins, 2016).

Considering the value of oil's production and downstream sector products (e.g., gasoline, kerosene, jet fuel, diesel oil, heating oil, lubricants, waxes, asphalt, natural gas, liquefied petroleum gas, and a plethora of petrochemicals; EPA, 2021), this natural resource serves as a steady source of wealth for commercial and government entities (Ralby, 2020). When comparing the world's largest companies by revenue, the most lucrative industry is oil and gas (Noyan, 2021).

A reality for those in the security field is that criminal networks have also recognized the profit in oil and energy pipelines. Globally, black market offenders steal oil to avoid taxes and take advantage of regional price differences (Osborne, 2020; Ralby, 2020). Although pipeline theft in places like Nigeria and Mexico is standard and high profile enough to make headlines (Ambituuni *et al.,* 2015; Desjardins, 2017; Jones and Sullivan, 2019), the problem is equally systemic in places like Cyprus, Turkey (Hadjicostis *et al*., 2020), Iran, Iraq (Dadpay, 2020), the United Kingdom (Gordon, 2017), Brazil (Terzian *et al.,* 2020), and the Philippines (Olchondra, 2014).

Although hydrocarbon product theft can impact operations upstream (for example, exploration and production), midstream (for example, processing, storing, transporting, and marketing), and downstream (for example, refining the raw materials into different energy products) levels, the latter is where most deleterious instances of it occur (Ralby, 2020; Soud, 2020). There are different types of fuel theft, including tapping pipelines (such as diverting refined oil and/or products via illicit taps), unauthorized ship-to-ship transfers, armed theft (for example, piracy and armed robbery at sea), bribery, smuggling/laundering (for example, bringing oil products to different jurisdictions, and selling them at discount), skimming mooring fees (that is, stealing from time-based charges applied for berthing marine vessels), illegal bunkering (such as pumping off to oil to small barges that are then illegally delivered to tankers), and adulteration (in other words, increasing profit margins by placing additives in oil or refined products which are then sold at full price; Desjardins, 2017; "Mooring lines…", 2021).

The impact of hydrocarbon theft is serious, jeopardizing the financial welfare, stability, and security of the regions and countries where it is committed (Ralby, 2020). It costs legitimate oil and gas organizations $133 billion annually in profits and investigative costs (Nazarova, 2019). A 2011 survey estimated that the majority (57.1 percent) of all fraud in the oil and gas industry was linked to grand corruption schemes, in other words, improper contributions made to high-level public officials and politicians (Hardoon and Heinrich, 2011; Ernst and Young, 2014). The practice is so deeply infiltrated in many operations that top officials, military service members, and law enforcement officers have been associated (Ralby, 2020). Hydrocarbon theft not only enables some international governments to haemorrhage oil production royalties and tax revenues but also helps to fund certain terrorist groups

(Gordts, 2014). Economically, these issues also contribute to higher gas prices for consumers (Desjardins, 2017).

Authorities have employed specific countermeasure methodologies to slow down the rate of oil and fuel theft. As an example, fuel dyes and pigments provide telltale tints to petroleum wares, which enables ease of identification. Also, the use of invisible molecular markers, a few parts per million of which are added to fuel, aids in discriminating between legitimate versus illicit hydrocarbon sources (Desjardins, 2017; Soud, 2020). Further, at the time of writing (April 2022), fuel tank level monitoring and alert systems using float sensor potentiometers, analogue systems, microcontrollers and/or Global System for Mobiles (GSM) and Global Positioning System (GPS) components are being tested and improved in India, Romania, and Pakistan (Lepcha *et al.,* 2015; Khuwaja *et al.,* 2018; Sharik *et al.,* 2019). The downside of doing these things is their high expense (Khuwaja *et al.,* 2018) and ability to counter (e.g., criminals in Ireland recently replicated and removed dyes from fuels; MacNamee, 2015).

As a result of the dramatic growth in international pipeline fuel theft cases since 2012 (Daugherty, 2014; Desjardins, 2017; Ralby, 2020; Terzian *et al.,* 2020), it is recognized that conventional countermeasures cannot adequately thwart thieves and help leaders gain control. In fact, theft-derived, annual crude oil and refined product financial losses often involve insider elements (Khartukov, 2021). Further, a global fraud survey conducted from 2017 to 2018 revealed that to meet business targets, 43 percent of oil, gas, and mining industry employees would engage in some sort of financial barratry (Ernst and Young, 2018). Considering the latter, it is logical for hydrocarbon companies to enhance human resource vetting approaches, including the implementation of cutting-edge software programs that use machine learning paradigms.

Currently, in the energy sector, machine learning and AI platforms are already used for operational purposes including interpretations of geology, geophysics, history, reservoir projects, and sensor data for knowledge graphs and safety evaluations (Brun et al., 2018; "Exploring the impact…", 2019). Although the field of artificial intelligence (AI) is relatively new (Buchanan, 2005; Brynjolfsson and McAfee, 2017; Rigano, 2018), it has proven to work in detecting employee theft (Quest *et al.,* 2018). In fact, multiple machine-learning systems have surpassed human-level performance in evaluating a person's cognitive state, based on facial expression or voice outputs (Brynjolfsson and McAfee, 2017; Junoh *et al.,* 2013; Rigano, 2018).

Research in this area shows that there are specific outputs of the human voice that are known to (1) be influenced by the speaker's internal and external environment and (2) carry real-time information about a person's psychological and physiological state, associations, and intentions (Hansen and Patil, 2007; Singh, 2019). Some of this information is embedded at such fine levels in the vocal signal, it is not discernible by the human ear. However, specific AI-powered machines can detect and map the signals, which serve as biomarkers of psychophysiological states, including specific threat-reactions (Hubculture, 2018; World Economic Forum, 2018). This paper describes the real-world application of one such AI-driven automated technology that detects human risk.

The primary purpose of this retrospective study was to determine whether the AI-driven automated technology used in a risk screening evaluation brought value in (1) providing insights and data not available through existing investigative methods to assist executive corporate personnel in making better-informed decisions about personnel, and (2) precisely identifying human-based risk concerning knowledge or involvement in internal theft and attempted poisoning. The secondary purpose of the study was to contribute to the literature, by revealing trends associated with the investigation of internal theft in a large organization.

Derived from the primary study purpose, the following hypotheses were tested: H1a: The automated technology would produce variation in risk assessment outputs; H1b: The flagging performance of the automated technology would be precise, H1c: No individual pertinent question posed would outperform the others in effectiveness, H1d: The automated technology would classify most admissions as risk-negative responses, and H1e: Admission numbers would be negatively associated with the consequences of pertinent questions posed.

Additionally, derived from the secondary purpose, to examine the robustness of the technology, the following hypotheses were tested: H2a: Interview outcomes would not be predicted by individual difference measures of job type, supervisory role, self-reported anxiety, and health status, and H2b: Observable behavioural tells would provide convergent validity for risk, anxiety, and age.

Research Methodology

Descriptive Statistics of Participants

The evaluation samples consisted of two sets of interviews (totalling *n*=56) conducted onsite on a Southeast Asian energy company's employees of a Malay archipelagic-situated facility (the exact company name and location of which must remain anonymous in this paper) on four days in 2016. The *n*=30 interviews conducted in the summer of 2016 (*n*=14 on August 2, *n*=16 on August 3) were focused on the theft of petroleum products and the attempted poisoning of an employee. The separate set of *n*=26 interviews executed in the late fall of 2016 (*n*=21 on December 8, *n*=5 on December 9) was focused on the theft of petroleum products and other corporate assets and resources. Of the ten interview time ranges represented, the frequencies (and respective *n*) distributed as follows: 29.79% (*n*=15) 10-11 AM, 17.86% (*n*=10) 11 AM-12 PM, 14.29% (*n*=8) 9-10 AM, 10.71% (*n*=6) 2-3 PM, 8.93% (*n*=5) 3-4 PM, 7.14% (*n*=4) 4-5 PM, 5.36% (*n*=3) each for 1-2 PM and 5-6 PM, and 1.79% (*n*=1) each for 12-1 PM and 8-9 PM.

A total of *n*=40 employees voluntarily participated in the automated interviews (*n*=14 only in August, *n*=10 only in December, and *n*=16 in both August and December). None of the participants who were scheduled to take an automated interview underwent attrition (e.g., as a result of refusing to participate, medical or mental faculty reasons, etc.). All interviewees represented in this evaluation voluntarily consented to take the automated interview as part of the ongoing investigation.

With respect to all completed automated interviews (all of which were executed in the Tagalog language), males represented 100% of the *n*=40 participants; the average age was 42.29 years ± 11.61 S.D. (range 25 to 60, *n*=26 unknown). For all *n*=56 automated interviews given, 62.5% (*n*=35) interviewees did not self-report anxiety, stress, or trauma, while 37.5% (*n*=21) did. Additionally, 55.4% (*n*=31) did not self-report medical states or conditions while 44.6% (*n*=25) did (at an average rate of 1.64 conditions/interviewee ± 0.91 S.D.). Of the 16 unique medical conditions or complaints self-reported, the frequencies (and respective *n*) distributed as follows: 19.51% (*n*=8) sleep deprivation; 17.07% (*n*=8) cold or flu; 12.20% (*n*=5) hypertension; 9.76% (*n*=4) headache; 7.32% (*n*=3) prescribed medications (for chronic conditions); 4.88% (*n*=2) each for cough, fatigue and sore throat; and 2.44% (*n*=1) each for alcohol use, back pain, diabetes, ear pain, hearing loss, hunger, shoulder pain, and sinus issues.

Of the 19 job positions represented, the frequencies (and respective *n*) distributed as follows: 20.0% (*n*=8) Second Engineer, 17.5% (*n*=7) Oiler, 15.0% (*n*=6) Wiper, 5% (*n*=2) Electrical Engineer, 5% (*n*=2) Operations Manager, 5% (*n*=2) Third Engineer, 2.5% (*n*=1) Boatswain, 2.5% (*n*=1) Common Maintenance Foreman (Engine Specialist), 2.5% (*n*=1) Head of Security, 2.5% (*n*=1) Head Shift Engineer, 2.5% (*n*=1) Instrument Engineer, 2.5% (*n*=1) Operations Engineering Supervisor, 2.5% (*n*=1) Mechanical Maintenance Supervisor, 2.5% (*n*=1) Procurement Officer (Fuel Specialist), 2.5% (*n*=1) Purchasing Officer, 2.5% (*n*=1) Security and Safety Health Specialist, 2.5% (*n*=1) Security Officer, 2.5% (*n*=1) Security Senior Manager, and 2.5% (*n*=1) Senior Electrical Engineer. Regarding the management of employees, 62.5% (*n*=25) of participants did not supervise others, while 37.5% (n=15) of participants held a supervisory role. The average tenure at the company was 9.63 years ± 7.63 S.D. (range 1.67 to 22, n=23 unknown).

Ethical Standards and Informed Consent

All procedures followed were in accordance with the ethical standards of the responsible committee on human experimentation [institutional and national] and with the Helsinki Declaration of 1975, as revised in 2000 (Snežana, 2001). Not originally collected for research purposes, the data from the evaluation represented real-world research. In keeping with Convention 108+, minimal data collection methods were employed (Greenleaf, 2014). The privacy of participants was respected and protected according to internationally established ethical guidelines (Robson and McCartan, 2016), including strict compliance with the Data Privacy Act (DPA) of 2012 (Ching *et al.,* 2018), with personally identifiable information either anonymized, kept strictly confidential (via legally binding agreements), or not collected at all.

Materials

Decision Support Tool

Clearspeed Verbal™, henceforth referred to as "the automated technology", is an enterprise-level, automated voice analytics tool that quickly assesses an individual's risk association relative to explicit themes or issues by means of an automated telephone interview. By evaluating specific vocal articulation outputs, this automated system detects and quantifies the presence or absence of voice-based risk reactions to client-defined questions. The AI-enabled technology leverages validated voice analytics and technical processes to evaluate responses to specific questions asked during the interview.

Unique Aspects of the Technology

The automated technology enables precise risk alerts based on an individual's vocal responses in any language, without the need to store personal identifiable information (PII). The tool incorporates the use of issue-specific questions asked during an automated telephonic interview to evaluate the presence or absence of risk signatures in the voice. Researchers have provided evidence that perceptions, cognitions, and emotional arousal are communicated through the voice (Cowen *et al.,* 2019; Simon-Thomas *et al.,* 2009). In the automated process, the voice characteristics evaluated are the result of distinct neurocognitive reactions to specific screening questions and have neural correlates (Dedovic *et al.,* 2009; Farrow *et al.,* 2013; Muehlhan *et al.,* 2013). There is ample evidence that specific information in human voice outputs can indicate the presence or absence, and intensity, of Central Nervous System and Autonomic Nervous System driven reactions in real-world environments wherein the perception of high stakes is involved (e.g., Brenner *et al.,* 1994; Laukka *et al.,* 2008; Scherer, 2003; Sondhi *et al.,* 2015; Van Puyvelde *et al.*, 2018; Williams and Stevens, 1972). Further, the link between linguistics, vocal cues, and risk markers of fraud detection has been established (Throckmorton *et al.,* 2015). The automated technology described here creates a model of the human voice in any language, for "yes" or "no" responses to risk-focused questions. The voice model is transformed, processed, analyzed, and quantified using a series of proprietary methodologies which evaluate and classify specific features of vocal responses. Once the voice input completes the processing cycle, a risk level for each response to specific questions is calculated and assigned, from low-to-high. At the time of this evaluation, the expected turnaround time for results was within 24 hours of interview completion.

Technical Process

The typical automated telephonic interview process employs Session Initiation Protocol (SIP) capable of securely conducting hundreds of simultaneous telephonic interviews from anywhere in the world.

Although the primary Cloud-based enterprise system uses SIP, for the evaluation described, Mobile systems specifically developed for austere applications were employed. These systems are housed in rugged MILSPEC notebook computers and equipped with specially configured VOIP telephone instruments to conduct interviews.

Upon interview completion, an encryption system packages the user responses, which are securely transferred to an AI-driven risk evaluation system, trained over several years via supervised-learning using labelled data. Additionally, multiple Quality Control processes are used to ensure the precision and accuracy of each evaluated response. A report of the evaluation is then automatically created and transferred to the client in the desired format. To adhere to privacy by design foundational principles, all data are encrypted both at rest and in transit. Interview results are typically accessed via a secure online dashboard, based on user role and permissions (i.e., the account owner can control and define permissions and restrict information only to those who need to see/use it). Further, the data of each interview is stored in a secure sound file format (instead of a commercially standard sound file), which requires proprietary software to decrypt.

Continuum of Individual Responses and Overall Results

The automated technology's risk framework boundaries have been pre-established (i.e., remain constant) such that evaluation output results fall into one of four risk determinations along a continuum: low risk (LR) which equates to no appreciable risk, average risk (AR) defined as negligible risk, potential risk (PR), which equates to a mid-level of risk, and high risk (HR). In this particular evaluation, due to the seven pertinent questions (PQs) asked, each interview produced a total of seven risk-reaction results, with one of four AI-generated risk scores per question. The highest risk score among all questions determined the overall risk assessment for each interview.

Interview Outcome Categories

Following the automated process, each interview was associated with an outcome result along a continuum: low risk (LR), average risk (AR), potential risk (PR), and high risk (HR). Further, three additional outcomes are admission (AD), suspected countermeasure (CM), and not completed (NC). The latter three are the result of QC evaluation and scoring. An AD was scored when the interviewee provided a "yes" response to any PQ asked. A CM was the result of an interviewee showcasing specific behavioural characteristics (e.g., inaudible whispers to one or more questions or refusing to respond to any PQ) that were presumed to be attempts at subverting accurate AI evaluations (Hughes, 2017; Navarro and Karlins, 2015; Nierenberg, 2010). An NC interview was the result of a technical or similar issue (e.g., bad telephone connection) occurring during the interview. Risk-negative interviews were those in the LR and AR ranges. Risk-positive interviews with outcomes of PR, HR, AD, or CM are typically recommended for follow-up.

Procedure

In August and December 2016, a Clearspeed specialist was deployed overseas to support a crisis management firm hired by a large Southeast Asian energy company after a severe and damaging insider theft problem was discovered. Some $8 million of company revenues had been lost to insider theft of petroleum products over the course of five years. The client sought to use the automated technology within the framework of the ongoing investigation that had already begun in July 2016. At that time, the problem-solving focus was on the human, mechanical, and maritime elements that enabled theft at the energy plant to be committed. Initial investigative steps included an internal audit and survey of personnel and logistical components to determine how resources were being stolen. Concurrent to the latter, the investigative team partnered services provided by Clearspeed over the course of ten total days (inclusive of four days of onsite automated interviews) to determine the extent of human risk at the site. From those initial interviews and subsequent follow-up analyses, the management firm continued to develop human threat assessments and learn more about the known network of criminality.

Employees who were not cleared by the preliminary investigative process underwent a Clearspeed interview. For each of the *n*=40 total participants from the personnel pool who were queried by the automated technology (*n*=14 in August 2016, *n*=10 in December 2016, and *n*=16 in both August and December 2016), the interview (1) was comprised of seven PQs posed in one language (Tagalog), (2) took less than 10 minutes to complete (i.e., an average of 9.40 minutes ± 0.41 S.D., range 7 to 12, *n*=12 unknown) and (3) had results processed within 24 hours.

Based on prior statements, details gleaned, and evidence gathered as part of the investigation (that started in July 2016), a subset of participants underwent *n*=20 subsequent testimonial interviews on six different days, between August 2016 and January 2017. To ensure the fairest and highest ethical codes of practice were exercised in keeping with Rule 15.07 of the Southeast Asian country's Labor Code (American Bar Association Asian Law Initiative, 2007), each interview was conducted in the presence of an attorney, an HR representative, and at least one investigative expert. Since most of these interviews involved employees who were also flagged for risk during the automated interview, the objective (factual) ground truth details involving these personnel were used to calculate the precision (positive predictive value) of the automated interview technology.

Design

Phase 1 vs. Phase 2

In Phase 1 (August 2016), the automated technology was implemented to complement the client's investigative methodology of procuring (1) physical evidence, (2) digital evidence, and (3) witness statements linking specific employees to the theft of petroleum products and the attempted poisoning of a co-worker. Specifically, *n*=30 employees were requested to take Phase 1 automated interviews by the investigative team and assigned attorney because of preliminary evidence and/or witness statements.

In Phase 2 (December 2016), the automated technology was implemented to complement the client's procurement of (1) physical evidence, (2) digital evidence, and (3) witness statements linking specific employees to the theft of petroleum products and other corporate assets and resources. Specifically, *n*=26 employees were requested to take Phase 2 automated interviews by the investigative team and assigned attorney because of preliminary evidence and/or witness statements.

Regardless of phase, under the guidance of a representative attorney, each participant voluntarily consented to take the automated interview as part of the ongoing investigation. The selection criteria for progressing to onsite testimonial interviews were: (1) connection with the most incriminating evidence, (2) witness statements, and (3) flags on the automated interview process. Of the *n*=20 Phase 1 and Phase 2 follow-up testimonial interviews, whereas the investigative experts were unblinded to the evidence and witness statements, they were blinded to 45 percent (*n*=6 Phase 1, *n*=3 Phase 2) of the automated risk evaluation results before the testimonial interviews. The steps of the client's investigative process (inclusive of pre-screening and automated interview insertion) are detailed in Figure 1.

Figure 1: Steps of Evidence Gathering and Screening Process

Step 1: Preliminary Due Diligence *(Onsite)*

- **Within first month of being hired, the crisis management team probes for evidence (physical, digital) and witness statements.**
- **HR and corporate attorney requests that specific employees take automated interview as part of ongoing investigation.**
- **Schedule and logistics provided.**

Step 2: Prescreen and Pretest Introduction *(Interview Room)*

- **Upon entering room, employee reads, and signs non-disclosure/disclaimer consent agreement provided and explained by local AP attorney.**
- **Employee is then greeted by Clearspeed supervisor and US-cleared linguist, who record basic demographic details and initiate pre-screening query process re self-reported alcohol, medicines, conditions, and anxiety (Y, N).**
- **Employee is then given instructions and list of pertinent questions.**

Step 3: Automated Interview *(Interview Room)*

- **To protect privacy and data confidentiality, a unique one-time 9-digit numeric code is assigned to each consenting employee, which does not provide any personally identifiable information to interview administrators, the system, or personnel.**
- **Clearspeed supervisor initiates automated interview (via Mobile system) in Tagalog**
- **To ensure quality control, supervisor monitors the candidate during the interview (~10 minutes of 2-3 iterations) and takes behavioral notes.**
- ***Risk results turnaround ~ 24 hours.***

Step 4: Investigative Testimonial Interview *(Private room)*

- **Consenting employees who flag for risk and/or evidential process are followed-up by team of experts (with HR and attorney present) for testimonial interview.**
- **Explanations for flags and associations to persons of interest and key criminal factors are explored in depth.**

Interview Foci

The automated interview approaches and questions were different for Phase 1 and Phase 2 interviewees. To minimize threats to validity, only the most salient issues were queried, reflected in the limited number—of seven—PQs approved by the client for each Phase. These questions were based on a combinatorial approach of *a priori* knowledge--client-specified question themes and issues—and information gleaned from in-depth communications with the client's executive team, regarding the history, pervasiveness, and impact of the most critical thematic areas. All final questions were approved in writing by the client.

Since the posed PQs reflected the high-stakes nature of the themes represented, the goal of the automated interview was to determine the presence or absence of risk reaction(s) to knowledge and/or personal involvement questions. Specifically, the foci of the automated interviews were on knowledge and/or involvement in the theft or misappropriation of company funds or resources, the theft of company fuel, the sale of stolen fuel, and the attempted poisoning of a coworker. The interviews were conducted to obtain potentially actionable information to resolve operational issues (See Table 1).

Table 1: Pertinent Questions used in Insider Theft and Attempted Felony Investigation

Phase 1 (August 2016): Fuel Theft and Poisoning

1. **Knowledge:** Do you have direct knowledge of company funds being stolen in any way?
2. **Knowledge:** Do you know who is stealing any fuel from the company?
3. **Knowledge:** Do you know who is selling any of the company's stolen fuel?
4. **Knowledge:** Do you know who is buying any of the company's stolen fuel?
5. **Involvement:** Have you ever stolen any fuel from the company?
6. **Involvement:** Have you ever sold any fuel stolen from the company?
7. **Knowledge:** Do you know who was involved in the attempted poisoning of a coworker?

Phase 2 (December 2016): Internal Theft

1. **Knowledge:** Do you know anyone involved with the theft or misappropriation of company resources worth more than 1,000 Pesos?
2. **Involvement:** Have you ever been intimidated not to report the theft of company resources?
3. **Knowledge-Involvement:** Have you ever been asked to help with the theft of any company resources worth more than 1,000 Pesos?
4. **Involvement:** Have you ever helped with the theft of any company resources worth more than 1,000 Pesos?
5. **Involvement:** Have you ever been directly involved in the theft of any company resources worth more than 1,000 Pesos?
6. **Involvement:** Have you ever received money not to report the theft of any company resources?
7. **Involvement:** Have you ever received any money as a result of the theft of any company resources?

Data Sample

Specifically, to address whether onsite employees of a Southeast Asian energy company had ever known about or participated in the theft or misappropriation of company funds, resources, fuel, and the attempted poisoning of a coworker, data not originally collected for research purposes was obtained from a total of n=40 employees of one geographical location (in other words, one plant work site) in the Tagalog language. The comprehensive data from Phase 1 originated from n=30 completed automated interviews (that transpired over two days between August 3-4, 2016). The data from Phase 2 originated from n=26 completed automated interviews (that transpired over two days between December 8-9, 2016). A total of n=20 follow-up testimonial interviews were conducted (n=4 on August 3, 2016; n=4 on August 4, 2016; n=2 on November 23, 2016; n=2 on November 24, 2016; n=4 on January 24, 2016; and n=4 on January 25, 2016).

When testing most hypotheses, the PQs (independent variables) posed by the automated tool were the stimuli. Other than tool-generated risk outputs (dependent variables), the additional data collected consisted of descriptive details (extraneous variables), such as interview times, self-reported medical status and anxiety, age, job position, supervisory role, tenure, overt behavioural cues, admissions of either knowing or being involved in the types of criminal activities queried and confirmatory details produced by the independent investigative team.

Interviewees' agitative and calm real-time behavioural cues were detected and hand-recorded by a risk and vulnerability assessment expert, with extensive intelligence community training in observable activity recognition and collection. The unique cues were sorted and quantified during the retrospective analysis. Intercoder reliability measures were not executed.

Establishing Ground Truth

During the investigative process, information was collected that was objective and real, supported by empirical (physical and digital) evidence. These facts were later used for testimonial interview verification of flagged automated interview assessment results. For instance, if a participant flagged on an automated interview internal fuel theft involvement question, and digital records indicated he had falsified fuel intake records, this would be considered a confirmation of risk. If admission or disclosure was made, factual verification of details relayed was a requirement of validation.

Statistical Tests

For hypotheses checking, the following analyses were executed via IBM SPSS Statistics 23 software: the non-parametric Chi-Square Goodness of Fit Test (to test the likelihood of a sample coming from a specified population distribution), the non-parametric rank-based Friedman analysis of variance (to test of whether more than two related groups differ), the non-parametric rank correlation Spearman's test (to test the strength and direction of association between two variables), cumulative odds ordinal logistic regression (to predict an ordinal dependent variable with a set of predictor variables), and multiple regression (to test the relationship between a single dependent variable and several predictor variables). Additionally, the following analyses were executed via Microsoft Office Excel 2016 software: frequencies of risk score outcomes, Student's *t*-test, and positive predictive value (PPV) proportion tests. *NOTE: For the first two analysis types, omitted interviews included those that were assessed as incompletes admissions and/or suspected countermeasures.*

Results

Descriptive Statistics

Individual Response Risk Results

In Phase 1 of our study (August 2016), the n=30 completed interviews resulted in a total of n=210 distinct responses (i.e., LR, AR, PR, HR, AD, CM). Of the n=210 distinct responses, the PQ risk evaluation frequencies distributed as follows: 24.8% (n =52) LR; 34.8% (n =73) AR; 16.7% (n =35) PR; 15.2% (n =32) HR; 5.2% (n =11) AD; and 3.3% (n =7) CM.

In Phase 2 (December 2016), the n=26 completed interviews resulted in a total of n=182 distinct responses (i.e., LR, AR, PR, HR, AD, CM). Of the n=182 distinct responses, the PQ risk evaluation frequencies distributed as follows: 31.3% (n =57) LR; 42.3% (n =77) AR; 8.8% (n =16) PR; 8.2% (n =15) HR; 1.6% (n =3) AD; and 7.7% (n =14) CM (Table 2A).

In both missions, based on the automated evaluation and quality control review, questions that resulted in admissions were separately analyzed for their associated risk levels, the results of which are presented in a separate section.

Interview Outcome Results

In Phase 1 (August 2016), of the n=30 completed interviews, the overall interview assessment outcomes (i.e., based on highest risk rating across the seven PQs) distributed as follows: 3.3% (n=1) LR; 10.0% (n=3) AR; 20.0% (n=6) PR; 46.7% (n=14) HR; 16.7% (n=5) AD; and 3.3% (n=1) CM.

In Phase 2 (December 2016), of the n=26 completed interviews, the overall interview assessment outcomes distributed as follows: 0.0% (n=0) LR; 30.8% (n=8) AR; 19.2% (n=5) PR; 30.8% (n=8) HR; 11.5% (n=3) AD; and 7.7% (n=2) CM (Table 2B).

Table 2: Comparison of result type frequencies of individual responses and interviews

A. Distinct PQ Responses

Evaluated risk	Phase 1 (*n* = 210) Fuel Theft and Poisoning	Phase 2 (*n* = 182) Internal Theft
Low Risk (LR)	24.8% (*n* = 52)	31.3% (*n* = 57)
Average Risk (AR)	34.8% (*n* = 73)	42.3% (*n* = 77)
Potential Risk (PR)	16.7% (*n* = 35)	8.8% (*n* = 16)
High Risk (HR)	15.2% (*n* = 32)	8.2% (*n* = 15)
Admission (AD)	5.2% (*n* = 11)	1.6% (*n* = 3)
Countermeasure (CM)	3.3% (*n* = 7)	7.7% (*n* = 14)

B. Completed Interview Outcomes

Evaluated risk	Phase 1 (*n* = 30) Fuel Theft and Poisoning	Phase 2 (*n* = 26) Internal Theft
Low Risk (LR)	3.3% (*n* = 1)	0.0% (*n* = 0)
Average Risk (AR)	10.0% (*n* = 3)	30.8% (*n* = 8)
Potential Risk (PR)	20.0% (*n* = 6)	19.2% (*n* = 5)
High Risk (HR)	46.7% (*n* = 14)	30.8% (*n* = 8)
Admission (AD)	16.7% (*n* = 5)	11.5% (*n* = 3)
Countermeasure (CM)	3.3% (*n* = 1)	7.7% (*n* = 2)

Pertinent Question Risk Results

In Phase 1 (August 2016), involvement-based PQ6 and knowledge-based PQ7 elicited the highest number of HR responses (18.8 percent each). Involvement-based PQ5 elicited the highest number of PR responses (22.9 percent).

In Phase 2 (December 2016), involvement based PQ4 elicited the highest number of HR responses (26.7 percent). Knowledge-based PQ1 and involvement-based PQ4 elicited the highest number of PR responses (25.0 percent each).

Observable Behavioral (Voice and Body) Indicator Results

Most participant volunteers (*n*=45, 80.4 percent) who completed *n*=56 automated interviews demonstrated a total of *n*=146 perceptible behavioural "tells" (Collett, 2004; Hughes, 2017) observed by the Clearspeed supervisor. Of the noted tell instances, *n*=66 were audible cues of voice/speech tells, and *n*=80 were visual cues of the body tells. Among the *n*=45 interviewees who showcased behavioural tells, the average number noted was 3.24 ± 2.08 (range 1-9).

For the *n*=18 distinct audible indicator types represented, their frequencies (and respective *n*) distributed as follows: 19.70% (*n*=13) changing pitch/intonation, 18.18% (*n*=12) low volume responses, 9.09% (*n*=6) protracted (drawn-out) responses, 9.09% (*n*=6) repeat replies, 9.09% (*n*=6) wrong language used, 4.55% (*n*=3) inappropriately responding "no" to neutral questions, 4.55% (*n*=3) changing volume, 3.03% (*n*=2) coughing, 3.03% (*n*=2) grumbling/groaning, 3.03% (*n*=2) laughing, 3.03% (*n*=2) muttering, 3.03% (*n*=2) not responding, 3.03% (*n*=2) stuttering, 1.52% (*n*=1) burping, 1.52% (*n*=1) slurring, 1.52% (*n*=1) voice-cracking, 1.52% (*n*=1) whistling, and 1.52% (*n*=1) yawning.

For the *n*=33 distinct visual indicator types represented, their frequencies (and respective *n*) distributed as follows: 13.75% (*n*=11) feet fidgeting, 10% (*n*=8) hand fidgeting, 7.5% (*n*=6) body rocking, 7.5% (*n*=6) brow furrowing or raising, 6.25% (*n*=5) crossing/uncrossing of legs, 5% (*n*=4) gaze shifting, 5% (*n*=4) increased blinking, 3.75% (*n*=3) toe-tapping, 2.5% (*n*=2) angry facial expression, 2.5% (*n*=2) exasperated facial expression, 2.5% (*n*=2) fingernail drumming, 2.5% (*n*=2) polyuria, 2.5% (*n*=2) leg-rubbing, 2.5% (*n*=2) head nodding, 2.5% (*n*=2) knee pumping, 2.5% (*n*=2) scratching, 1.25% (*n*=1) skittishness, 1.25% (*n*=1) chewing, 1.25% (*n*=1) eye-closing, 1.25% (*n*=1) crossing of arms, 1.25% (*n*=1) slitting of eyes, 1.25% (*n*=1) apprehensive facial expression, 1.25% (*n*=1) finger splaying, 1.25% (*n*=1) flitting eye gaze, 1.25% (*n*=1) forced/fake smiling, 1.25% (*n*=1) increased swallowing, 1.25% (*n*=1) chair leaning, 1.25% (*n*=1) lip-chewing, 1.25% (*n*=1) lip-smacking, 1.25% (*n*=1) taking off eyeglasses, 1.25% (*n*=1) blank-staring, 1.25% (*n*=1) sucking, and 1.25% (*n*=1) table gripping.

Inferential Statistics

Observed vs. Expected Interview Response Frequencies

To test H1a, we conducted a Chi-Square Goodness of Fit test for each operational phase. In Phase 1 (August 2016), of the *n*=30 completed interviews, *n*=24 participants completed interviews consisting of *n*=168 responses that were assessed as LR, AR, PR, or HR (as a result of "yes" or "no" responses to the seven PQs relevant to internal theft or the attempted poisoning of a co-worker). The statistical test revealed the assessments distributed unequally across the different risk levels, $\chi 2$ (3, N = 168) = 17.76, p = 0.00049. In Phase 2 (December 2016), of the *n*=26 completed interviews, *n*=21 participants completed interviews consisting of *n*=147 responses that were "purely" assessed as LR, AR, PR, or HR (as a result of "yes" or "no" responses to the seven focused PQs relevant to internal theft of company resources). As opposed to chance probability, the statistical test revealed the assessments distributed unequally across the different risk levels, $\chi 2$ (3, N = 147) = 60.59, $p < 0.00001$. Specifically, for both phases, low-risk evaluations were most common.

Flagging Precision

To test H1b, we executed a positive predictive value analysis, whereby the following definitions applied: "not confirmed" meant justification for the risk-positive score was not found, "confirmed" meant justification for the risk-positive score was found, "confirmed-validated" meant verification of risk or admission was due to factually-confirmed knowledge and/or involvement, and "confirmed-mitigated" meant verification of the risk-positive reaction resulted from associative memories, confusion (e.g., language barrier) or deliberate attempts to manipulate results.

Although the fourteen admissions among seven employees were made during the actual automated interviews of Phase 1 (August 2016, *n*=11 admissions) and Phase 2 (December 2016, *n*=3 admissions), all disclosure details were solely provided during the follow-up testimonial interviews. When considering both phases, *n*=44 total interviews flagged for risk (i.e., *n*=11 PR, *n*=22 HR, *n*=3 CM, *n*=8 AD). Due to evidence stemming from the company's internal investigation and *n*=20 follow-up testimonial interviews, *n*=27 of all flagged automated interviews were further reviewed. Of these, *n*=2 was not confirmed (i.e., potential Type I errors). Of the *n*=25 that confirmed, 92.0 percent (*n*=23) were confirmed-validated, and 0.8 percent (*n*=2) were confirmed-mitigated. The overall automated output and results production process showcased a high positive predictive value (PPV) of 92.6 percent.

Differences in Pertinent Questions

In Phase 1 (August 2016), for *n*=24 automated interviews assessed as LR, AR, PR, or HR, the application of the Friedman test revealed no evidence of stochastic dominance between the PQs for score outputs, $\chi 2(6) = 5.65$, $p = .46$.

In Phase 2 (December 2016), for *n*=21 interviews, the application of the Friedman test revealed no evidence of stochastic dominance between the PQs for score outputs, $\chi 2(6) = 3.28$, $p = .77$.

Consistent with H1c, then, in each evaluation phase, there was no evidence that any individual question outperformed the others.

Relationship Between Admissions and Risk Negative Outputs

When pooling the *n*=14 admissions from both phases of the evaluation, 100 percent of the *n*=14 admissions made were classified by the automated technology as risk-negative responses (57.1 percent or *n*=8 were associated with LR results, 42.9 percent, or *n*=6 were associated with AR results). Consistent with H1d, no admissions were associated with risk-positive (PR or HR) results.

Admission Frequencies and Correlation with Pertinent Question Stakes

In Phase 1 (August 2016), 16.7 percent (*n*=5) of a total of 30 completed automated interviews resulted in admissions during the automated interview phase. With *n*=11 admissions among *n*=5 admitters, the average rate was 2.2 admissions/admitter. Of all *n*=11 admission responses made, the rank order and frequency of admissions by PQ were: PQ1 = PQ2 (36.4 percent) > PQ3 (18.2 percent) > PQ4 (9.1 percent) > PQ5 = PQ6 = PQ7 (0 percent). Knowledge-themed (lower consequence) questions PQ1 (Do you have direct knowledge of company funds being stolen in any way?) and PQ2 (Do you know who is stealing any fuel from the company?) posed produced the most admissions (*n*=4 each). The last three PQs (Have you ever stolen any fuel from the company? Have you ever sold any fuel stolen from the company? Do you know who was involved in the attempted poisoning of a coworker?) did not result in any admissions. Considering the Phase 1 results of *n*=11 AD risk responses from *n*=5 interviews, a Spearman's rank correlation test produced a scatterplot, that visually revealed a monotonic relationship, $r(6) = -0.954$, $p = 0.001$. In support of H1e, a very strong, negative correlation was found between the perceived stakes of pertinent questions (ranked from most to least consequential, by crisis management experts) and admission numbers; this relationship was significant at the 99.99 percent confidence level.

In Phase 2 (December 2016), 11.5 percent (*n*=3) of a total of 26 completed interviews resulted in admissions during the automated interview phase. With *n*=3 admissions among *n*=3 admitters, the average rate was 1.0 admissions/admitter. Of all *n*=3 admission responses made the rank order and frequency of admissions by PQ were: PQ1 (100 percent) > PQ2 = PQ3 = PQ4 = PQ5 = PQ6 = PQ7 (0 percent). The first PQ (Do you know anyone involved with the theft or misappropriation of company resources worth more than 1,000 Philippine Pesos?) was the only one associated with all admissions made (*n*=3). All other pertinent questions posed resulted in no admissions. Due to the low sample number, a Spearman's rank correlation test was not attempted with Phase 2 results.

Effects of Supervisory Role, Job Tenure, Self-Reported Health Status and Anxiety on Interview Risk Outcomes

To test H2a a cumulative odds ordinal logistic regression with proportional odds was run to determine the effect of age, tenure, supervisory role, self-reported health status, and anxiety on the overall four types of risk outcomes (as determined by the automated technology) in interviews. To assure the independence of observations assumption was not violated, results from Phase 1 (when all interviewees were evaluation naïve) were analyzed.

In Phase 1, there were proportional odds, as assessed by a full likelihood ratio test. A generalization of the residual sum of squares used to test the suitability of the mathematical representation of interest, the deviance goodness-of-fit test indicated that the model was a good fit to the observed data, $\chi2(16) = 11.718$, $p = .763$. However, the final model did not statistically predict the dependent variable significantly over and above the intercept-only model, $\chi2(4) = 8.444$, $p = .077$. Having a supervisory role had no statistically significant effect on interview risk outcome, $\chi2(1) = .018$, $p = 0.894$. Also, self-reported health status (not including trauma) of an interviewee had no statistically significant effect, $\chi2(1) = .018$, $p = 0.894$. Self-reported anxiety status had no significant effect, $\chi2(1) = .001$, $p = 0.999$. Further, job tenure (expressed in years) was not significantly associated with a change in interview risk outcome, with an odds ratio of .737 (95% CI, 0.515 to 1.076), $\chi2(1) = 2.798$, $p = 0.094$. In other words, supervisory role, job tenure, self-reported health status and anxiety had no statistically significant effect on the final risk outcome.

Predictors of Observable Behavioral Tells Numbers in Interviewees

To test H2b, a multiple regression analysis was run to predict total "tell" numbers (body and voice) in interviews from interview evaluation risk outcomes, self-reported anxiety, and age. To assure the independence of observations assumption was not violated, results from Phase 1 (when all interviewees were evaluation naïve) were analyzed.

There was linearity as assessed by partial regression plots and a plot of studentized residuals against the predicted values. There was independence of residuals, as assessed by a Durbin-Watson statistic of 2.296. There was homoscedasticity, as assessed by visual inspection of a plot of studentized residuals versus unstandardized predicted values. There was no evidence of multicollinearity, as assessed by tolerance values greater than 0.1. There were no studentized deleted residuals greater than ±3 standard deviations, no leverage values greater than 0.2, and no values for Cook's distance above 1. The assumption of normality was met, as assessed by a Q-Q Plot. The multiple regression model statistically predicted total "tell" number at a significant level, $F(3, 3) = 13.68$, $p < .03$, adj. $R^2 = .864$—a large effect size, according to Cohen's (1988) classification. Risk level and self-reported anxiety were positively associated with observable behavioural tells, as assessed by the Clearspeed supervisor. Regression coefficients and standard errors can be found in Table 3.

Table 3: Regression results for a total number of Phase 1 interview body and voice tells

# risk+ responses	*B*	95% CI for *B*		*SE B*	β	R^2	ΔR^2
		LL	*UL*				
Model						0.932	0.864*
Constant	0.391	-3.839	4.620	1.329			
Risk Level	4.558*	1.945	7.171	0.821	1.944*		
Anxiety	-6.463	-10.367	-2.559	1.227	-1.949*		
Age	-0.008*	-0.085	0.069	0.024	-0.059		

Note. Model = "Enter" method in SPSS Statistics; *B* = unstandardized regression coefficient; CI = confidence level; *LL* = lower limit; *UL* = upper limit; *SE B* = standard error of coefficient; β = standardized coefficient; R^2 = coefficient of determination; ΔR^2 = adjusted R^2.
*$p \leq 0.05$, **$p < 0.01$, ***$p < 0.001$.

Discussion

In this retrospective use-case study conducted in a Southeast Asian Energy Company as part of a professional investigative inquest into insider theft and an allegation of an attempted felony, the automated tool precisely alerted the investigative team to risk identification flags of knowledge and involvement. This study involved a unique, real-world data set from a high-stakes internal theft investigation, with multiple control variables (i.e., using a single: organization, Clearspeed supervisor, and team of third-party experts who conducted subsequent interviews) to produce findings that balanced both internal and external validity.

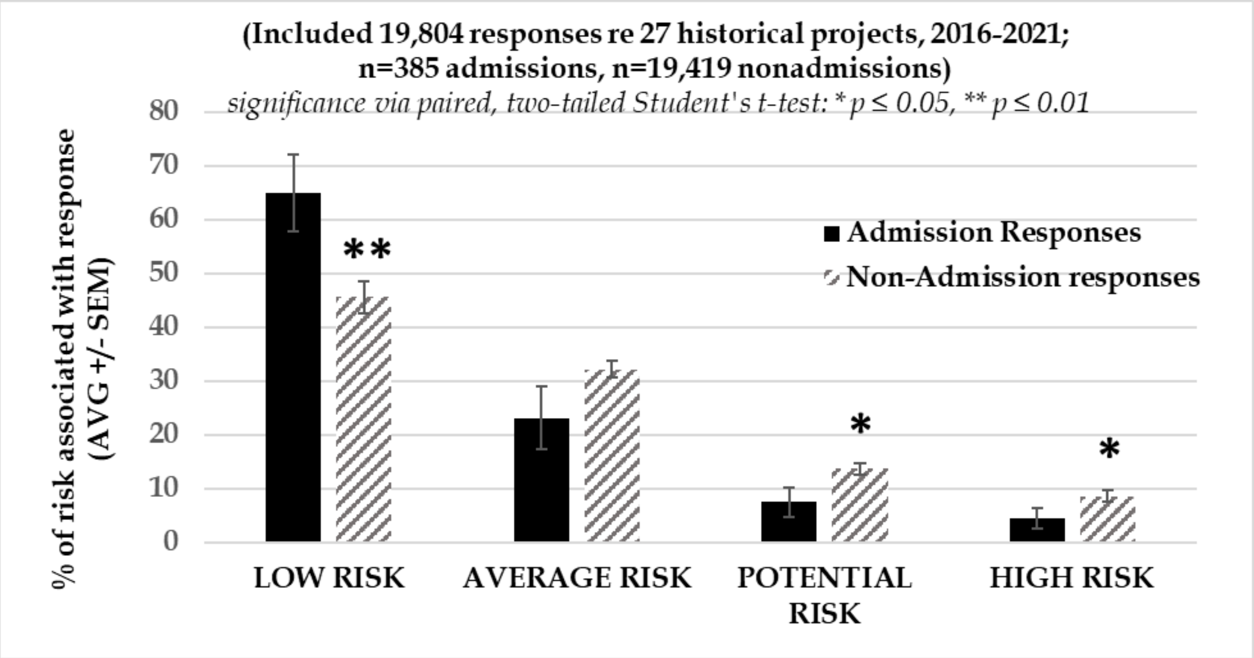
Notable Findings

The high percentage of HR interview outcomes (46.7 percent for Phase 1 and 30.8 percent for Phase 2) did not distribute normally, based on general historic statistical trends (that showcase lower frequencies of HR end results). Although the latter may be a manifestation of the tool's non-randomized implementation in this use-case (i.e., employees who were not initially cleared by the concurrent investigative process underwent a Clearspeed interview), it could also indicate an elevated risk status of this study's population.

Beyond the statistical evidence that the technology operated significantly better than chance and the pertinent questions used in the automated interviews were equally effective (in eliciting risk reactions), there was ~ 93 percent probability (precision) that flags would be confirmed (i.e., as a result of evidence discerned during the concurrent investigation and follow-up testimonial interviews). This result is close to the >94 percent PPV the AI-driven technology has achieved in similar historical operational settings, wherein verification details of flags were available (Martin *et al.,* 2021). The findings are also consistent with research suggesting that high-stakes environments translate to vocal signals that are easier to discern, more robust and more reliable than those derived in the lab (Brenner *et al.,* 1994; Mendoza and Carballo, 1998).

Regarding disclosures, in all instances, the first admission happened either during or after the execution of the automated interviews. Knowledge-based questions were associated with more admissions than involvement-based questions. Also, the more a question was perceived as consequential, the less likely an admission would be made; this information might be instructional for investigative experts. Further, considering both phases of the evaluation, it was notable that all of the *n*=14 admissions made during the automated interview were associated with the absence of risk detected in voice outputs (i.e., no detectable threat reactions). This trend is not new. When averaging other operational data spanning 27 similar projects (*n*=385 admissions, *n*=19,419 non-admissions), disclosures were consistently and significantly associated with higher frequencies of risk-negative and lower frequencies of risk-positive vocal outputs than their non-admission counterparts (Figure 2). These findings corroborate the literature that disclosures serve as points of release, which are reflected in attenuated, acute alert reactions (Farrow *et al.*, 2013; Suchotzki *et al.*, 2017; Verschuere *et al.*, 2018).

Figure 2: Historical Clearspeed Verbal™ risk frequencies for admissions vs. non-admissions



As predicted, it was determined that risk results were not influenced by supervisory role, job tenure, self-reported medical status, or anxiety. This serves as evidence against modelling bias by the technology's algorithm. It also substantiates similar evidence discerned from the automated risk results of a prior US-based client who had also experienced internal theft; in that use-case, neither gender, self-reported anxiety, shift worked, nor the date of interview impacted the interview outcome (Martin *et al.*, 2021).

Finally, while not correlated with age, the total number of overt behavioural (body and voice) tells exhibited by interviewees were directly linked to risk level and self-reported anxiety. The implication is that observable displays of risk relating to knowledge or involvement of an illicit act may vary according to apprehension levels and that using an automated risk assessment tool not influenced by anxiety and not requiring tête-à-tête interactions may be the wiser, more reliable choice. Further, the direct and significant association of behavioural tells with risk levels in this study provide convergent validity for the automated technology described.

Limitations and Future Research

There are limitations to this field evaluation that should be addressed. First, due to the necessary ethical constraints on the amount of sensitive information collected (Robson and McCartan 2016), specific characteristics of the client's operations and participants (e.g., prior criminal history, work history, education level) were not explored, which could have further enriched the analyses.

The relatively low frequency (7.4 percent) of apparent false positives may have reflected the objective nature of the automated technology, the prevalence of risk in the population sampled, and/or the skills of the following experts and investigative team. Nevertheless, even a low frequency of false positives can be associated with issues for an organization that adopts a novel risk management strategy. The consequences of making a Type I error can equate to an intervention that is unnecessary. Further, repeated overestimations of risk can, over time, devalue the reputation of an organization. If an automated decision support tool is deployed from a disciplinary human resource perspective, the privileges and status of employees who produce flagged interviews may be compromised up to the point of and after they are cleared (Kaminski and Schonert, 2017; Kemshall *et al.,* 2011; Murphy *et al.,* 2014)

A screening metrics constraint of this study was the inability to confirm risk-negative results (that is, true versus false negatives) using objective measures (i.e., the known blind spot of field research; Fiedler *et al.,* 2012). However, as is usually the case with field studies of real-world research, the implementation of known positives and known negatives (in randomly selected control vs. experimental groups) was not a viable option. Although challenging, it will be important for future studies to replicate high-stakes conditions, while also employing (1) random assignment of control vs experimental groups, (2) appropriate blinding procedures, and (3) the application of known ground truth to all results.

Finally, although the use of AI-driven systems holds great promise to boost profitability, productivity, and morale in energy sector organizations that seek enhanced risk identification strategies, there are also potential ethical concerns to consider, including (1) machines replacing humans in jobs devoted to risk detection, (2) lack of transparency of technological complexity and issues like concept drift, and (3) integration challenges with existing organizational risk management protocols and tools. (Chui and Manyika, 2018; NI Business Info, 2020). Future studies that address these issues would certainly help the field.

Conclusions

One of this study's implications is that AI-enabled automated technologies like the one described can serve as powerful additions to the investigative arsenal already in place for energy sector organizations. Although AI applications are underutilized in most industries (Hoadley and Sayler, 2020; Knight, 2020), resilient organizations tend to be more amenable to complementary risk assessment strategies. Approaches that are most readily adopted assess risk along a continuum, such that actionable steps can be taken to identify, mitigate, isolate, monitor, avoid, transfer, or escalate flagged issues (Gius *et al.,* 2018, Meyer *et al.,* 2011; Tselyutina and Vlasova, 2019). The most rigorous of correctly implemented contemporary risk assessment systems are holistic, layered, redundant, technology-enabled, interdisciplinary, and serve the purpose of helping end-users identify risk in order to make better decisions faster while allocating their precious resources accordingly (Gius *et al.,* 2018).

Another study implication is that the automated tool can help energy sector security managers gain realistic insights pertinent to the high-stakes issues in their establishments, such that resources are suitably allocated. In the evaluation described, the organization's leaders' most immediate objective was not to identify and arrest bad actors in their organization. Rather, their main goal was to understand the size, scale, and scope of their internal theft problem. Within a year, multiple arrests were made, and a more comprehensive picture of the criminal enterprise within the organization came to light.

In guiding international energy sector organizations on where to target resources, the automated tool described only distinguishes between risk-positive and risk-negative voice responses; it does not expose the rationale behind the signals. As even the best AI-risk assessment machines are limited and periodically err in determinations of "flag" or "clear", in decision support situations, they should identify, not adjudicate. It would be ethically problematic to exclusively use risk results as "evidence" in employee terminations or legal cases. Neurocognitive reactions that affect vocal outputs (translated into risk) can be due to a variety of reasons other than malfeasance (such as auxiliary memories, associations, knowledge, and individual variability). Ultimately, the investigative team needs to shoulder the burden of discovering what resultant risk flags indicate. Decisions of interpretation and procession should always be deferred to experts (people), who comprehend nuances. As a result, robots and AI must not be allowed to supplant human judgment and intelligence.

Acknowledgements

The authors wish to thank (1) Dr Leanne ten Brinke for her thoughtful review of an earlier draft of the manuscript. Her comments and suggestions helped improve and clarify the final version, and (2) the Clearspeed Research and Development Department for their technical expertise, guidance and input on the project and paper.

References

American Bar Association Asian Law Initiative, ABA-ALI (2007). "Legal Ethics Manual." *Scribd*. Scribd, 2007. Available at: https://www.scribd.com/doc/37897865

Ambituuni, A. *et al.* (2015) "Risk assessment of a petroleum product pipeline in Nigeria: The realities of managing problems of theft/sabotage," *WIT Transactions on The Built Environment* [Preprint]. Available at: https://doi.org/10.2495/safe150051.

Brenner, M., Doherty, T. and Shipp, T. (1994). "Speech measures indicating workload demand." *Aviation, Space, and Environmental Medicine*, 65, pp. 21-26. Available at: https://psycnet.apa.org/record/199420587001

Brun, A., Trench, M. and Vermaat, T. (2018) *Why oil and gas companies must act on analytics*, *McKinsey & Company*. McKinsey & Company. Available at: https://www.mckinsey.com/industries/oil-and-gas/our-insights/why oil-and-gas-companies-must-act-on-analytics.

Brynjolfsson, E. and McAfee, A. (2017) "The Business of Artificial Intelligence," *Harvard Business Review*, 24 June. Available at: https://hbr.org/2017/07/the-business-of-artificial-intelligence.

Buchanan, B. (2005) "A (Very) Brief History of Artificial Intelligence." *AI Magazine*. 26(4), pp. 53-60. Available at: https://ojs.aaai.org/index.php/aimagazine/article/view/1848

Ching, M., Fabito, B. and Celis, N. (2018) "Data Privacy Act of 2012: A case study." *Advanced Science Letters*, 24(10), pp. 7042–7046. Available at: https://www.researchgate.net/publication/327281053_Data_Privacy_Act_of_2012

Chui, M. and Manyika, J. (2018) *The real-world potential and limitations of Artificial Intelligence*, *McKinsey & Company*. McKinsey & Company. Podcast. Available at: https://www.mckinsey.com/featured-insights/artificial intelligence/the-real-world-potential-and-limitations-of-artificial-intelligence.

Cohen, J. (1988) *Statistical Power Analysis for the behavioral sciences*. Hillsdale, NJ: L. Erlbaum Associates.

Collett, P. (2004) *Book of tells*. TRANSWORLD PUBLISHERS LTD.

Cowen, A. *et al.* (2019) "Mapping 24 emotions conveyed by brief human vocalization.," *American Psychologist*, 74(6), pp. 698–712. Available at: https://doi.org/10.1037/amp0000399.

Dadpay, A. (2020) "An analysis of fuel smuggling in the Middle East as a single multinational market," *Journal of Industry, Competition and Trade*, 20(4), pp. 643–656. Available at: https://doi.org/10.1007/s10842-020-00340-6.

Dalby, C. (2014) *5 industries worried about Peak oil*, *OilPrice.com*. Available at: https://oilprice.com/Energy/Crude Oil/5-Industries-Worried-About-Peak-Oil.html.

Daugherty, A. (2014) *Rising oil theft in Mexico points to collusion with foreign businesses*, *InSight Crime*. Available at: https://insightcrime.org/news/brief/rising-oil-theft-points-to-criminal-collusion-with-foreign-businesses/.

Dedovic, K. *et al*. (2009) "Neural correlates of processing stressful information: An event-related fmri study," *Brain*

Research, 1293, pp. 49–60. Available at: https://doi.org/10.1016/j.brainres.2009.06.044.

Desjardins, J., Grandi, W.B.T.D. and Article & Editing Katie Jones (2017) *Crude awakening: The global black market for oil*, *Visual Capitalist*. Available at: https://www.visualcapitalist.com/global-black-market-fuel-theft/.

Desjardins, J. (2017) *Fuel theft is a big problem*, *Business Insider*. Business Insider. Available at: https://www.businessinsider.com/fuel-theft-is-a-big-problem-2017-5.

Desjardins, J., Grandi, W.B.T.D. and Article & Editing Katie Jones (2016) *The oil market is bigger than all metal markets combined*, *Visual Capitalist*. Available at: https://www.visualcapitalist.com/size-oil-market/.

EPA. Environmental Protection Agency (2021) *Types of Refined Petroleum Products. EPA.* Available at: https://www.epa.gov/emergency-response/types-refined-petroleum-products.

Ernst and Young (2014) *Managing Bribery and Corruption Risks in the Oil and Gas Industry, EY.* Available at: https://www.ey.com/Publication/vwLUAssets/EY-Managing-briberyand-corruption-risk-in-the-oil-and-gas-industry/$FILE/EY-Managing-bribery-and-corruption-risk-in-the-oil-and-gas-industry.pdf

Ernst and Young (2018) *Integrity in the Spotlight - the Future of Compliance 15th Global Fraud Survey. ReadkonG.com.* Available at: https://www.readkong.com/page/integrity-in-the-spotlight-the-future-of-compliance-4216989

Exploring the impact of artificial intelligence on offshore oil and Gas (2019). Available at: https://www.offshore technology.com/analysis/application-of-artificial-intelligence-in-oil-and-gas-industry/.

Farrow, T. *et al.* (2013) "Neural correlates of the behavioral-autonomic interaction response to potentially threatening stimuli," *Frontiers in Human Neuroscience*, 6(34), pp. 1-17. Available at: https://doi.org/10.3389/fnhum.2012.00349.

Fiedler, K., Kutzner, F. and Krueger, J. (2012) "The long way from α-error control to validity proper," *Perspectives on Psychological Science*, 7(6), pp. 661–669. Available at: https://doi.org/10.1177/1745691612462587.

Frequently asked questions (FAQs) - U.S. energy information administration (EIA) (2021) *Frequently Asked Questions (FAQs) U.S. Energy Information Administration (EIA)*. Available at: https://www.eia.gov/tools/faqs/faq.php?id=33&%3Bt=6.

Gius, D. *et al.* (2018) *Value and Resilience through Better Risk Management," McKinsey & Company*. McKinsey & Company. Available at: https://www.mckinsey.com.br/capabilities/risk-and-resilience/our-insights/value and-resilience-through-better-risk-management.

Gordon, A. (2017) *Father and son in scam to Syphon off £1m Hydrocarbon Fuel*, *Daily Mail Online*. Associated Newspapers. Available at: https://www.dailymail.co.uk/news/article-4600722/amp/Father-son-scam-syphon-1m hydrocarbon-fuel.html.

Gordts, E. (2014) *How ISIS is using oil to fund its Terror*, *HuffPost*. HuffPost. Available at: https://www.huffpost.com/entry/isis-oil_n_5877008.

Greenleaf, G. (2014) *Asian data privacy laws-trajectories, lessons, and optimism*, *OUP Academic*. Oxford University Press. Available at: https://doi.org/10.1093/acprof:oso/9780199679669.003.0020.

Hadjicostis, M. *et al.* (2020) *Cyprus blasts 'pirate state' Turkey's new gas drilling bid*, *The Times of Israel*. Available at: https://www.timesofisrael.com/cyprus-blasts-pirate-state-turkeys-new-gas-drilling-bid/.

Hansen, J.. and Patil, S. (2007) "Speech under stress: Analysis, Modeling and Recognition," *Lecture Notes in Computer Science*, pp. 108–137. Available at: https://doi.org/10.1007/978-3-540-74200-5_6.

Hardoon, D. and Heinrich, F. (2011) *Bribe payers index 2011 - publications*, *Transparency.org*. Available at: https://www.transparency.org/en/publications/bribe-payers-index-2011.

Hoadley, D. and Sayler, K. (2020) *Artificial Intelligence and National Security, CRS Report R45178. 2020.* Available at: https://crsreports.congress.gov/product/pdf/R/R45178/5

Hubculture (2018) *Hub culture Davos 2018 - Rita Singh, Research Faculty at Carnegie Mellon University, YouTube.* YouTube. Available at: https://www.youtube.com/watch?v=wvGP8I5UhHk.

Hughes, C. (2017) *The ellipsis manual: Analysis and engineering of human behavior*. Virginia Beach, VA: Evergreen Press.

Jones, N. and Sullivan, J. (2019) "Huachicoleros: Criminal cartels, fuel theft, and violence in Mexico," *Journal of Strategic Security*, 12(4), pp. 1–24. Available at: https://doi.org/10.5038/1944-0472.12.4.1742.

Junoh, A.K. *et al.* (2013) "Crime detection with DCT and artificial intelligent approach," *Advanced Materials Research*, 816-817, pp. 610–615. Available at: https://doi.org/10.4028/www.scientific.net/amr.816-817.610.

Kaminski, P. and Schonert, J. (2017) *The neglected art of risk detection*, *McKinsey & Company*. McKinsey & Company.

Available at: https://www.mckinsey.com/business-functions/risk-and-resilience/our-insights/the-neglected art-of-risk-detection.
Kemshall, H., Mackenzie, G. and Wilkinson, B. (2011). "6.1.2 False Positive and False Negatives in Risk Assessments." *6 Best Practice and Avoiding Errors | ROH V4.36. Risk of Harm Guidance and Training Resources*. Available at: http://www.nomsintranet.org.uk/roh/roh/6-best_practice/06_01_02.htm
Khartukov, E. (2021) *Oil theft: A frightening international perspective – Oilman Magazine.* Available at: https://oilmanmagazine.com/article/oil-theft-a-frightening-international-perspective/
Khuwaja, K.*et al.* (2018) "Automatic fuel tank monitoring, tracking and Theft Detection System," *ANNALS OF THE ORADEA UNIVERSITY. Fascicle of Management and Technological Engineering.*, Available at: https://doi.org/10.15660/auofmte.2018-1.3333.
Knight, W. (2020) *Ai is all the rage. so why aren't more businesses using it?*, *Wired*. Conde Nast. Available at: https://www.wired.com/story/ai-why-not-more-businesses-use/.
Laukka, P. *et al.* (2008) "In a nervous voice: Acoustic analysis and perception of anxiety in social phobics' speech," *Journal of Nonverbal Behavior*, 32(4), pp. 195–214. Available at: https://doi.org/10.1007/s10919-008-0055-9.
Lepcha N., Sherpa, T. and Tamang, J. (2015). "GSM Based Fuel Theft Detector Using Microcontroller." *International Journal of Advance Electrical and Electronics Engineering* 4(3), pp. 7-12. ISSN: 2278-8948.
MacNamee, G. (2015) *Fuel smuggling: Thugs taking courses on how to remove dye from Diesel*, *Irish Mirror*. Available at: https://www.irishmirror.ie/news/irish-news/crime/fuel-smuggling-thugs-taking-courses-5599589.
Martin, A., Zaal, J., Azevedo, M., and Dutra, C. (2021). A Novel Decision Support Risk Assessment Tool for Investigating Internal Theft in a Gaming Establishment: A Case Study. *International Journal of Gaming Hospitality and Tourism.* 1(1), pp. 1-20. Available at: https://ojs.stockton.edu/index.php/light/article/view/20/3
Mendoza, E. and Carballo, G. (1998) "Acoustic analysis of induced vocal stressby means of cognitive workload tasks," *Journal of Voice*, 12(3), pp. 263–273. Available at: https://doi.org/10.1016/s0892-1997(98)80017-9.
Meyer, M., Roodt, G. and Robbins, M. (2011) "Human Resources Risk Management: Governing people risks for improved performance," *SA Journal of Human Resource Management*, 9(1), pp 1-12. Available at: https://doi.org/10.4102/sajhrm.v9i1.366.
Mooring lines sample clauses (2021) *Law Insider*. Available at: https://www.lawinsider.com/clause/mooring-lines.
Muehlhan, M. *et al*. (2013) "Enhanced sympathetic arousal in response to fmri scanning correlates with task induced activations and deactivations," *PLoS ONE*, 8(8), pp. 1–11. Available at: https://doi.org/10.1371/journal.pone.0072576.
Murphy, K., Myors, B., and Wolach, A. (2014) *Statistical Power Analysis: A simple and general model for traditional and modern hypothesis tests*. London, England: Routledge.
Navarro, J. and Karlins, M. (2015) *What every body is saying: An ex-FBI agent's Guide to Speed-reading people*. New York, NY: Harper Collins.
NI Business Info (2020) *Risks and limitations of artificial intelligence in business*. Available at: https://www.nibusinessinfo.co.uk/content/risks-and-limitations-artificial-intelligence-business.
Nierenberg, G., Calero, H. and Grayson, G. (2010) *How to read a person like a book: Observing body language to know what people are thinking*. Garden City Park, NY: Square One Publishers.
Nazarova, A. (2019) *Dark world of oil theft and fraud*, *Fraud Magazine*. Available at: https://www.fraud magazine.com/article.aspx?id=4295006268.
Noyan, B. (2021) *Global 500 2016*, *Fortune*. Fortune. Available at: https://fortune.com/global500/2016/.
Olchondra, R. (2014) *With oil smuggled in, gov't waves tax revenues goodbye*, *INQUIRER.net*. Available at: https://business.inquirer.net/165510/with-oil-smuggled-in-govt-waves-tax-revenues-goodbye.
Osborne, J. (2020) *Can fuel theft be stopped once and for all? Houston Chronicle*. Houston Chronicle. Available at: https://www.houstonchronicle.com/business/energy/article/Can-fuel-theft-be-stopped-once-and-for-all 15277830.php.
Quest, L. *et al.* (2018) *The risks and benefits of using AI to detect crime*, *Harvard Business Review*. Available at: https://hbr.org/2018/08/the-risks-and-benefits-of-using-ai-to-detect-crime.
Ralby, I. (2020) *Downstream oil theft*, *Atlantic Council*. Available at: https://www.atlanticcouncil.org/programs/global energy-center/downstream-oil-theft/.

Rigano, C. (2018) *Using artificial intelligence to address criminal justice needs*, *National Institute of Justice*. Available at: https://nij.ojp.gov/topics/articles/using-artificial-intelligence-address-criminal-justice-needs.

Robson, C. and McCartan, K. (2016) *Real World Research*. Chichester: Wiley.

Scherer, K. (2003) "Vocal Communication of Emotion: A review of research paradigms," *Speech Communication*, 40(1 2), pp. 227–256. Available at: https://doi.org/10.1016/s0167-6393(02)00084-5.

Sharik, N., Chaitanya, D. and Shankar, L. (2019) "Fuel Level Monitoring & Alert System using IOT," *International Journal of Engineering and Advanced Technology*, 8(652), pp. 525–527. Available at: https://doi.org/10.35940/ijeat.f1153.0886s219.

Simon-Thomas, E. *et al.* (2009) "The voice conveys specific emotions: Evidence from vocal burst displays.," *Emotion*, 9(6), pp. 838–846. Available at: https://doi.org/10.1037/a0017810.

Singh, R. (2019) *Profiling humans from their voice*. Singapore: Springer.

Snežana, B. (2001). "The declaration of Helsinki: The cornerstone of research ethics". *Archives of Oncology*. 9(3): pp. 179–184. Available at: http://www.onk.ns.ac.rs/Archive/Vol9/PDFVol9/V9n3p179.pdf

Sondhi, S. *et al.* (2015) "Vocal indicators of emotional stress," *International Journal of Computer Applications*, 122(15), pp. 38–43. Available at: https://doi.org/10.5120/21780-5056.

Soud, D. (2020) *Downstream oil theft: Countermeasures and good practices*, *Atlantic Council*. Available at: https://www.atlanticcouncil.org/in-depth-research-reports/report/downstream-oil-theft-countermeasures and-good-practices/.

Suchotzki, K. *et al.* (2017) "Lying takes time: A meta-analysis on reaction time measures of deception.," *Psychological Bulletin*, 143(4), pp. 428–453. Available at: https://doi.org/10.1037/bul0000087.

Terzian, R. *et al.* (2020) "CCPD: A reactive information system to prevent fuel theft in oil pipelines," *Rio Oil and Gas Expo and Conference*, 20(2020), pp. 433–434. Available at: https://doi.org/10.48072/2525-7579.rog.2020.433.

Throckmorton, C. *et al.* (2015) "Financial fraud detection using vocal, linguistic and financial cues," *Decision Support Systems*, 74, pp. 78–87. Available at: https://doi.org/10.1016/j.dss.2015.04.006.

Tselyutina, T., and Vlasova, T. (2019). "Development of the Personnel Risks Assessment and Supply Chain Strategy as a Basis of the Risk Management System of Modern Organizations." *International Journal of Supply Chain Management*, 8(5), pp. 1030-1038. ISSN: 2050-7399.

Van Puyvelde, M. *et al.* (2018) "Voice stress analysis: A new framework for voice and effort in human performance," *Frontiers in Psychology*, 9. Available at: https://doi.org/10.3389/fpsyg.2018.01994.

Verschuere, B. *et al.* (2018) "Taxing the brain to uncover lying? meta-analyzing the effect of imposing cognitive load on the reaction-time costs of lying." *Journal of Applied Research in Memory and Cognition*. Available at: https://doi.org/10.31219/osf.io/uqvkt.

Wang, Z. *et al.* (2020) "The development trend of internal combustion engine," *Journal of Physics: Conference Series*, 1626(1), pp. 1–5. Available at: https://doi.org/10.1088/1742-6596/1626/1/012139.

Williams, C. and Stevens, K. (1972) "Emotions and speech: Some acoustical correlates," *The Journal of the Acoustical Society of America*, 52(4B), pp. 1238–1250. Available at: https://doi.org/10.1121/1.1913238.

World Economic Forum (2018) *Pinpointing people and places through Advanced Audio Analysis | Rita Singh*, *YouTube*. YouTube. Available at: https://www.youtube.com/watch?v=4HjcQjwKBWM.